SAS®-Anwendungsroutinen in SAS® Enterprise Guide® 8.4 und SAS® Add-In 8.4 for Microsoft Office

8.4*

* Dieses Dokument gilt möglicherweise für weitere Versionen der Software. Öffnen Sie dieses Dokument in SAS Help Center und klicken Sie auf die Version im Banner, um alle verfügbaren Versionen anzuzeigen.

SAS® Dokumentation
18. April 2024

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2024. *SAS®-Anwendungsroutinen in SAS® Enterprise Guide® 8.4 und SAS® Add-In 8.4 for Microsoft Office*. Cary, NC: SAS Institute Inc.

SAS®-Anwendungsroutinen in SAS® Enterprise Guide® 8.4 und SAS® Add-In 8.4 for Microsoft Office

Copyright © 2024, SAS Institute Inc., Cary, NC, USA

All Rights Reserved. Produced in the United States of America.

For a hard copy book: No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

For a web download or e-book: Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

The scanning, uploading, and distribution of this book via the Internet or any other means without the permission of the publisher is illegal and punishable by law. Please purchase only authorized electronic editions and do not participate in or encourage electronic piracy of copyrighted materials. Your support of others' rights is appreciated.

U.S. Government License Rights; Restricted Rights: The Software and its documentation is commercial computer software developed at private expense and is provided with RESTRICTED RIGHTS to the United States Government. Use, duplication, or disclosure of the Software by the United States Government is subject to the license terms of this Agreement pursuant to, as applicable, FAR 12.212, DFAR 227.7202-1(a), DFAR 227.7202-3(a), and DFAR 227.7202-4, and, to the extent required under U.S. federal law, the minimum restricted rights as set out in FAR 52.227-19 (DEC 2007). If FAR 52.227-19 is applicable, this provision serves as notice under clause (c) thereof and no other notice is required to be affixed to the Software or documentation. The Government's rights in Software and documentation shall be only those set forth in this Agreement.

SAS Institute Inc., SAS Campus Drive, Cary, NC 27513-2414

April 2024

SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.

8.4-P1:egamotasks

Inhalt

Kapitel 1 / Arbeiten mit Anwendungsroutinen **1**
Integrierte Anwendungsroutinen und SAS Studio-Tasks 1
Filtern der Daten für eine Task 3
Hinzufügen von Benutzercode zu einer Anwendungsroutine 5
Zuweisen von Spalteneigenschaften 6
Auswählen des Grafikausgabeformats 6
Arbeiten mit Anwendungsroutinenvorlagen 7

Kapitel 2 / Anwendungsroutine 'Tabelle anhängen' **11**
Allgemeines über die Anwendungsroutine 'Tabelle anhängen' 11
Tabelle anhängen: Erstellen einer Liste der Eingabetabellen 12
Tabelle anhängen: Speichern von Ergebnissen 12

Kapitel 3 / Flächendiagramm **15**
Allgemeines über die Anwendungsroutine 'Flächendiagramm' 15
Flächendiagramm: Auswählen eines Diagrammtyps 16
Flächendiagramm: Zuweisen von Variablen zu Analyserollen 16
Flächendiagramm: Festlegen von Darstellungsoptionen 18
Festlegen von Titeln und Fußnoten 19
Anzeigen von Eigenschaften 19

Kapitel 4 / Anwendungsroutine 'ARIMA-Modellierung und -Prognosen' **21**
Allgemeines über die Anwendungsroutine 'ARIMA-Modellierung und -Prognosen' 21
ARIMA-Modellierung und -Prognosen: Zuweisen von Variablen zu Analyserollen 22
ARIMA-Modellierung und -Prognosen: Festlegen von Identifizierungsoptionen 23
ARIMA-Modellierung und -Prognosen: Festlegen von Schätzoptionen 25
ARIMA-Modellierung und -Prognosen: Festlegen von Prognoseoptionen 28
Festlegen von Titeln und Fußnoten 30
Anzeigen von Eigenschaften 30

Kapitel 5 / Zuweisen einer Projektbibliothek **33**
Allgemeines über den Projektbibliotheks-Assistenten 33
Projektbibliothek zuweisen: Angeben eines Namens und eines Servers 34
Projektbibliothek zuweisen: Angeben einer Engine 34
Projektbibliothek zuweisen: Festlegen von Konfigurationsoptionen 37
Projektbibliothek zuweisen: Überprüfen der Zusammenfassungsinformationen 38

Kapitel 6 / Balkendiagramm **39**
Allgemeines über die Anwendungsroutine 'Balkendiagramm' 39
Balkendiagramm: Auswählen des Diagrammtyps 40
Balkendiagramm: Zuweisen von Variablen zu Analyserollen 42
Balkendiagramm: Festlegen von Darstellungsoptionen 43

Festlegen von Titeln und Fußnoten ... 44
Anzeigen von Eigenschaften ... 44

Kapitel 7 / Assistent für Balkendiagramme ... 47
Allgemeines über den Assistenten für Balkendiagramme ... 47
Assistent für Balkendiagramme: Auswählen von Daten ... 47
Assistent für Balkendiagramme: Zuweisen von Variablen zu Rollen ... 48
Assistent für Balkendiagramme: Festlegen der Darstellungsoptionen ... 49
Assistent für Balkendiagramme: Festlegen von Titeln und Fußnoten ... 50

Kapitel 8 / Balken-Linien-Diagramm ... 51
Allgemeines über die Anwendungsroutine 'Balken-Liniendiagramm' ... 51
Balken-Linien-Diagramm: Zuweisen von Variablen zu Analyserollen ... 52
Balken-Linien-Diagramm: Festlegen von Darstellungsoptionen ... 53
Festlegen von Titeln und Fußnoten ... 54
Anzeigen von Eigenschaften ... 55

Kapitel 9 / Einfache Prognosen ... 57
Allgemeines über die Anwendungsroutine 'Einfache Prognosen' ... 57
Einfache Prognosen: Zuweisen von Variablen zu Analyserollen ... 58
Einfache Prognosen: Festlegen von Prognoseoptionen ... 59
Einfache Prognosen: Erzeugen von Plots ... 62
Einfache Prognosen: Festlegen von Ergebnisoptionen ... 62
Festlegen von Titeln und Fußnoten ... 63
Anzeigen von Eigenschaften ... 63

Kapitel 10 / Box-Regelkarte ... 65
Allgemeines über die Anwendungsroutine 'Box-Regelkarte' ... 65
Regelkarten: Zuweisen von Variablen zu Analyserollen ... 66
Regelkarten: Festlegen von Regelgrenzen ... 68
Regelkarte: Auswählen der durchzuführenden Tests ... 69
Regelkarten: Festlegen von Optionen für die Plotdarstellung ... 70
Regelkarten: Speichern von Analyseergebnissen ... 72
Festlegen von Titeln und Fußnoten ... 73
Anzeigen von Eigenschaften ... 73

Kapitel 11 / Box-Plot ... 75
Allgemeines über die Anwendungsroutine 'Box-Plot' ... 75
Box-Plot: Auswählen eines Diagrammtyps ... 76
Box-Plot: Zuweisen von Variablen zu Analyserollen ... 77
Box-Plot: Festlegen von Darstellungsoptionen ... 78
Festlegen von Titeln und Fußnoten ... 80
Anzeigen von Eigenschaften ... 80

Kapitel 12 / Blasendiagramm ... 83
Allgemeines über die Anwendungsroutine 'Blasendiagramm' ... 83
Blasendiagramm: Zuweisen von Variablen zu Analyserollen ... 84
Blasendiagramm: Festlegen von Darstellungsoptionen ... 85
Festlegen von Titeln und Fußnoten ... 86
Anzeigen von Eigenschaften ... 86

Kapitel 13 / Kanonische Korrelation ... **89**
Allgemeines über die Anwendungsroutine 'Kanonische Korrelation' ... 89
Kanonische Korrelation: Zuweisen von Variablen zu Analyserollen ... 90
Kanonische Korrelation: Festlegen von Statistikoptionen ... 91
Kanonische Korrelation: Erzeugen von Diagrammen ... 93
Kanonische Korrelation: Festlegen von Ergebnisoptionen ... 93
Festlegen von Titeln und Fußnoten ... 94
Anzeigen von Eigenschaften ... 95

Kapitel 14 / c-Regelkarte ... **97**
Allgemeines über die Anwendungsroutine 'c-Regelkarte' ... 97
c-Regelkarte: Zuweisen von Variablen zu Analyserollen ... 98
Regelkarten: Festlegen von Regelgrenzen ... 99
Regelkarte: Auswählen der durchzuführenden Tests ... 101
Regelkarten: Festlegen von Optionen für die Plotdarstellung ... 102
Regelkarten: Speichern von Analyseergebnissen ... 102
Festlegen von Titeln und Fußnoten ... 103
Anzeigen von Eigenschaften ... 104

Kapitel 15 / CDF-Plots ... **105**
Allgemeines über die Anwendungsroutine 'CDF-Plots' ... 105
Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen ... 106
CDF-Plots: Auswählen der Verteilung ... 107
CDF-Plots: Anpassen der Verteilung ... 108
Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung ... 109
Prozessfähigkeitsanalyse: Festlegen der Analysetypen ... 113
Festlegen von Titeln und Fußnoten ... 114
Anzeigen von Eigenschaften ... 114

Kapitel 16 / Charakterisieren von Daten ... **117**
Allgemeines über den Assistenten zum Charakterisieren von Daten ... 117
Charakterisieren von Daten: Auswählen von Daten ... 118
Charakterisieren von Daten: Berichtoptionen auswählen ... 118
Charakterisieren von Daten: Angeben der Ausgabegrenzen ... 119

Kapitel 17 / Clusteranalyse ... **121**
Allgemeines über die Anwendungsroutine 'Clusteranalyse' ... 121
Clusteranalyse: Zuweisen von Variablen zu Analyserollen ... 122
Clusteranalyse: Festlegen von Clusteroptionen ... 123
Clusteranalyse: Erzeugen von Plots und Diagrammen ... 125
Clusteranalyse: Festlegen von Ergebnisoptionen ... 126
Festlegen von Titeln und Fußnoten ... 128
Anzeigen von Eigenschaften ... 129

Kapitel 18 / Vergleichen von Daten ... **131**
Allgemeines zum Assistenten zum Vergleichen von Daten ... 131
Daten vergleichen: Auswählen von Daten ... 132
Daten vergleichen: Auswählen der zu vergleichenden Variablen ... 132
Daten vergleichen: Angeben der Vergleichskriterien ... 133
Daten vergleichen: Festlegen der Ausgabeoptionen ... 133

Kapitel 19 / Konturdiagramm ... **137**
Allgemeines über die Anwendungsroutine 'Konturdiagramm' ... 137
Konturdiagramm: Auswählen eines Diagrammtyps ... 138
Konturdiagramm: Zuweisen von Variablen zu Analyserollen ... 139
Konturdiagramm: Festlegen von Darstellungsoptionen ... 140
Festlegen von Titeln und Fußnoten ... 141
Anzeigen von Eigenschaften ... 142

Kapitel 20 / Dateien kopieren ... **143**
Allgemeines über die Anwendungsroutine 'Dateien kopieren' ... 143
Kopieren von Dateien von ihrem lokalen Computer auf einen SAS Server ... 144

Kapitel 21 / Korrelationen ... **147**
Allgemeines über die Anwendungsroutine 'Korrelationen' ... 147
Korrelation: Zuweisen von Variablen zu Analyserollen ... 148
Korrelation: Festlegen von Korrelationsoptionen ... 149
Korrelation: Festlegen von Ergebnisoptionen ... 151
Korrelationen: Erstellen der Ausgabedaten ... 152
Festlegen von Titeln und Fußnoten ... 153
Anzeigen von Eigenschaften ... 153

Kapitel 22 / Erstellen eines Formats ... **155**
Allgemeines über die Anwendungsroutine 'Format erstellen' ... 155
Erstellen eines Formats: Festlegen der Formatoptionen ... 156
Erstellen eines Formats: Definieren des neuen Formats ... 157
Anzeigen von Eigenschaften ... 160

Kapitel 23 / Format aus Datei erstellen ... **163**
Allgemeines über die Anwendungsroutine 'Format aus Datei erstellen' ... 163
Erstellen eines Formats aus einer Datei ... 165

Kapitel 24 / Erstellen einer Feature-Tabelle für Karten ... **167**
Allgemeines über die Anwendungsroutine 'Feature-Tabelle für Karten erstellen' 167
Erstellen einer Feature-Tabelle für Karten: Zuweisen von Variablen zu Analyserollen ... 168
Erstellen einer Feature-Tabelle für Karten: Festlegen von Optionen ... 169
Erstellen einer Feature-Tabelle für Karten: Festlegen der Ergebnisse ... 171
Festlegen von Titeln und Fußnoten ... 171
Anzeigen von Eigenschaften ... 171

Kapitel 25 / Assistent zum Erstellen von Zeitreihendaten ... **173**
Allgemeines über den Assistenten zum Erstellen von Zeitreihendaten ... 173
Assistent zum Erstellen von Zeitreihendaten: Auswählen der Daten ... 174
Assistent zum Erstellen von Zeitreihendaten: Zuweisen von Variablen zu Rollen ... 174
Assistent zum Erstellen von Zeitreihendaten: Festlegen der Optionen für Akkumulation, fehlende Werte und Null-Werte ... 175
Assistent zum Erstellen von Zeitreihendaten: Zusätzliche Ausgaben auswählen ... 176
Assistent zum Erstellen von Zeitreihendaten: Festlegen von Titeln und Fußnoten ... 176

Kapitel 26 / Dateiattribute ... **179**
Allgemeines über den Assistenten für Dateiattribute ... 179
Dateiattribute: Auswählen von Daten ... 180
Dateiattribute: Festlegen der Ausgabe ... 180

Kapitel 27 / Löschen von Dateien und Formaten ... **183**
Allgemeines über die Anwendungsroutine 'Dateien und Formate löschen' ... 183
Löschen von Dateien, Ansichten und benutzerdefinierten Formaten ... 184
Anzeigen von Eigenschaften ... 184

Kapitel 28 / Diskriminanzanalyse ... **187**
Allgemeines über die Anwendungsroutine 'Diskriminanzanalyse' ... 187
Diskriminanzanalyse: Zuweisen von Variablen zu Analyserollen ... 188
Diskriminanzanalyse: Festlegen der Analyseoptionen ... 189
Diskriminanzanalyse: Klassifizieren neuer Daten ... 191
Diskriminanzanalyse: Festlegen von Ergebnisoptionen ... 193
Festlegen von Titeln und Fußnoten ... 194
Anzeigen von Eigenschaften ... 194

Kapitel 29 / Verteilungsanalyse ... **197**
Allgemeines über die Anwendungsroutine 'Verteilungsanalyse' ... 197
Verteilungsanalyse: Zuweisen von Variablen zu Analyserollen ... 198
Verteilungsanalyse: Auswählen der Verteilungen ... 199
Verteilungsanalyse: Anpassen der Verteilung ... 200
Verteilungsanalyse: Erzeugen von Plots ... 201
Verteilungsanalyse: Einfügungen hinzufügen ... 203
Verteilungsanalyse: Festlegen von Ergebnisoptionen ... 203
Festlegen von Titeln und Fußnoten ... 204
Anzeigen von Eigenschaften ... 204

Kapitel 30 / Ringdiagramm ... **207**
Allgemeines über die Anwendungsroutine 'Ringdiagramm' ... 207
Ringdiagramm: Auswählen des Diagrammtyps ... 208
Ringdiagramm: Zuweisen von Variablen zu Analyserollen ... 208
Ringdiagramm: Festlegen von Darstellungsoptionen ... 209
Festlegen von Titeln und Fußnoten ... 212
Anzeigen von Eigenschaften ... 212

Kapitel 31 / Herunterladen von Dateien auf den PC ... **215**
Allgemeines über die Anwendungsroutine 'Dateien auf PC herunterladen' ... 215
Dateien auf PC herunterladen: Wählen Sie die herunterzuladenden SAS-Dateien aus ... 216
Dateien auf PC herunterladen: Angeben der Optionen zum Herunterladen ... 216

Kapitel 32 / Faktoranalyse ... **219**
Allgemeines über die Anwendungsroutine 'Faktoranalyse' ... 219
Faktoranalyse: Zuweisen von Variablen zu Analyserollen ... 220
Faktoranalyse: Festlegen von Optionen für Faktormethoden ... 221
Faktoranalyse: Festlegen von Kommunalitätsoptionen ... 223
Faktoranalyse: Festlegen von Rotations- und Plotoptionen ... 225
Faktoranalyse: Festlegen von Ergebnisoptionen ... 227
Festlegen von Titeln und Fußnoten ... 229

Kapitel 33 / Forecast Studio - Projekt erstellen ... **231**
Allgemeines über den Assistenten 'Forecast Studio - Projekt erstellen' ... 231
Forecast Studio - Projekt erstellen: Überprüfen der Daten ... 232
Forecast Studio - Projekt erstellen: Auswählen der Prognosevariablen ... 232
Forecast Studio - Projekt erstellen: Auswählen der Klassifizierungsvariablen ... 233
Forecast Studio - Projekt erstellen: Speichern des Projekts ... 235

Kapitel 34 / Forecast Studio - Projekt öffnen ... **237**
Allgemeines über den Assistenten 'Forecast Studio - Projekt öffnen' ... 237
Forecast Studio - Projekt öffnen: Auswählen des Prognoseprojekts ... 238
Forecast Studio - Projekt öffnen: Auswählen der Prognosezeitreihe ... 238
Forecast Studio - Projekt öffnen: Auswählen des Ergebnistyps ... 239
Forecast Studio - Projekt öffnen: Festlegen von Titeln und Fußnoten ... 239

Kapitel 35 / Forecast Studio - Wertekorrektur für Projekt ... **241**
Allgemeines über den Assistenten 'Forecast Studio - Wertekorrekturen für Projekt' ... 241
Forecast Studio - Wertekorrektur für Projekt: Auswählen des Prognoseprojekts 242
Forecast Studio - Wertekorrektur für Projekt: Auswählen der Prognosezeitreihe und Übertragen der Wertekorrekturen ... 243

Kapitel 36 / Verallgemeinerte lineare Modelle ... **247**
Allgemeines über die Anwendungsroutine 'Verallgemeinerte lineare Modelle' ... 247
Verallgemeinerte lineare Modelle: Zuweisen von Variablen zu Analyserollen ... 248
Verallgemeinerte lineare Modelle: Erstellen eines Modells ... 250
Verallgemeinerte lineare Modelle: Festlegen von Modelloptionen ... 250
Verallgemeinerte lineare Modelle: Auswählen von Post-Hoc-Tests ... 252
Verallgemeinerte lineare Modelle: Erzeugen von Plots ... 252
Verallgemeinerte lineare Modelle: Festlegen von Vorhersageoptionen ... 254
Festlegen von Titeln und Fußnoten ... 255
Anzeigen von Eigenschaften ... 255

Kapitel 37 / High-Performance - Lineare Regression ... **257**
Allgemeines über die Anwendungsroutine 'High Performance - Lineare Regression' ... 257
High-Performance - Lineare Regression: Zuweisen von Variablen zu Analyserollen ... 258
High-Performance - Lineare Regression: Erstellen eines Modells ... 260
High-Performance - Lineare Regression: Festlegen von Modelloptionen ... 263
High-Performance - Lineare Regression: Festlegen der Partitionsoptionen ... 264
High-Performance - Regression: Festlegen von Performance-Optionen ... 265
High-Performance - Regression: Festlegen von Vorhersageoptionen ... 266
Festlegen von Titeln und Fußnoten ... 267

Kapitel 38 / High-Performance - Logistische Regression ... **269**
Allgemeines über die Anwendungsroutine 'High Performance - Logistische Regression' ... 269
High-Performance - Logistische Regression: Zuweisen von Variablen zu Analyserollen ... 270
High-Performance - Logistische Regression: Erstellen eines Modells ... 271
High-Performance - Logistische Regression: Festlegen von Modelloptionen ... 276
High-Performance - Regression: Festlegen von Performance-Optionen ... 277

High-Performance - Regression: Festlegen von Vorhersageoptionen ... 278
Festlegen von Titeln und Fußnoten ... 279

Kapitel 39 / Histogramme ... **281**
Allgemeines über die Anwendungsroutine 'Histogramme' ... 281
Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen ... 282
Histogramme: Auswählen der Verteilung ... 284
Histogramme: Anpassen der Verteilung ... 285
Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung ... 286
Prozessfähigkeitsanalyse: Festlegen der Analysetypen ... 290
Festlegen von Titeln und Fußnoten ... 291
Anzeigen von Eigenschaften ... 291

Kapitel 40 / Urwertkarte ... **293**
Allgemeines über die 'Urwertkarte' ... 293
Urwertkarte: Zuweisen von Variablen zu Analyserollen ... 294
Regelkarten: Festlegen von Regelgrenzen ... 295
Regelkarte: Auswählen der durchzuführenden Tests ... 297
Regelkarten: Festlegen von Optionen für die Plotdarstellung ... 298
Regelkarten: Speichern von Analyseergebnissen ... 298
Festlegen von Titeln und Fußnoten ... 299
Anzeigen von Eigenschaften ... 300

Kapitel 41 / Sterbetafeln ... **301**
Allgemeines über die Anwendungsroutine 'Sterbetafeln' ... 301
Sterbetafeln: Zuweisen von Variablen zu Analyserollen ... 302
Sterbetafeln: Festlegen von Methodenoptionen ... 304
Sterbetafeln: Erzeugen von Plots ... 305
Sterbetafeln: Festlegen von Ergebnisoptionen ... 306
Festlegen von Titeln und Fußnoten ... 307
Anzeigen von Eigenschaften ... 308

Kapitel 42 / Liniendiagramm ... **309**
Allgemeines über die Anwendungsroutine 'Liniendiagramm' ... 309
Liniendiagramm: Auswählen eines Diagrammtyps ... 310
Liniendiagramm: Zuweisen von Variablen zu Analyserollen ... 311
Liniendiagramm: Festlegen von Darstellungsoptionen ... 313
Festlegen von Titeln und Fußnoten ... 314
Anzeigen von Eigenschaften ... 315

Kapitel 43 / Assistent für Liniendiagramme ... **317**
Allgemeines über die Anwendungsroutine 'Liniendiagramm - Assistent' ... 317
Assistent für Liniendiagramme: Auswählen der Daten ... 317
Assistent für Liniendiagramme: Zuweisen von Variablen zu Rollen ... 318
Assistent für Liniendiagramme: Festlegen der Plotoptionen ... 319
Assistent für Liniendiagramme: Festlegen der Darstellungsoptionen ... 320
Assistent für Liniendiagramme: Festlegen von Titeln und Fußnoten ... 320

Kapitel 44 / Lineare Modelle ... **321**
Allgemeines über die Anwendungsroutine 'Lineare Modelle' ... 321
Lineare Modelle: Zuweisen von Variablen zu Analyserollen ... 322
Lineare Modelle: Erstellen eines Modells ... 323

Lineare Modelle: Festlegen von Modelloptionen ... 324
Lineare Modelle: Festlegen erweiterter Optionen ... 325
Lineare Modelle: Auswählen von Post-Hoc-Tests ... 326
Lineare Modelle: Erzeugen von Plots ... 330
Lineare Modelle: Festlegen von Vorhersageoptionen ... 331
Festlegen von Titeln und Fußnoten ... 332
Anzeigen von Eigenschaften ... 332

Kapitel 45 / Lineare Regression ... 335
Allgemeines über die Anwendungsroutine 'Lineare Regression' ... 335
Lineare Regression: Zuweisen von Variablen zu Analyserollen ... 336
Lineare Regression: Festlegen von Modelloptionen ... 337
Lineare Regression: Festlegen von Statistikoptionen ... 340
Lineare Regression: Erzeugen von Plots ... 341
Lineare Regression: Festlegen von Vorhersageoptionen ... 343
Festlegen von Titeln und Fußnoten ... 344
Anzeigen von Eigenschaften ... 345

Kapitel 46 / Listenbericht ... 347
Allgemeines über die Anwendungsroutine 'Listenbericht' ... 347
Listenbericht: Zuweisen von Variablen zu Analyserollen ... 348
Listenbericht: Festlegen von Listenoptionen ... 349
Festlegen von Titeln und Fußnoten ... 351
Anzeigen von Eigenschaften ... 351

Kapitel 47 / Listenberichtassistent ... 353
Allgemeines über den Listenberichtassistenten ... 353
Listenberichtassistent: Auswählen von Daten ... 354
Listenberichtassistent: Definieren der Liste ... 354
Listenberichtassistent: Anzeigen von Summen ... 355
Listenberichtassistent: Festlegen von Titeln und Fußnoten ... 356

Kapitel 48 / Logistische Regression ... 357
Allgemeines über die Anwendungsroutine 'Logistische Regression' ... 357
Logistische Regression: Zuweisen von Variablen zu Analyserollen ... 358
Logistische Regression: Erstellen eines Modells ... 360
Logistische Regression: Festlegen von Modelloptionen ... 364
Logistische Regression: Erzeugen von Plots ... 366
Logistische Regression: Festlegen von Vorhersageoptionen ... 368
Festlegen von Titeln und Fußnoten ... 369
Anzeigen von Eigenschaften ... 369

Kapitel 49 / Kartendiagramm ... 371
Allgemeines über die Anwendungsroutine 'Kartendiagramm' ... 371
Kartendiagramm: Auswählen des Kartentyps ... 372
Kartendiagramm: Zuweisen von Variablen zu Analyserollen ... 373
Kartendiagramm: Festlegen von Darstellungsoptionen ... 374
Festlegen von Titeln und Fußnoten ... 376
Kartendiagramm: Ergebnisvorschau ... 377
Anzeigen von Eigenschaften ... 377

Kapitel 50 / X-quer- und R-Regelkarte ... **379**
Allgemeines über die Anwendungsroutine 'X-quer- und R-Regelkarte' ... 379
X-quer- und R-Regelkarte: Zuweisen von Variablen zu Analyserollen ... 380
Regelkarten: Festlegen von Regelgrenzen ... 381
Regelkarte: Auswählen der durchzuführenden Tests ... 383
Regelkarten: Festlegen von Optionen für die Plotdarstellung ... 384
Regelkarten: Speichern von Analyseergebnissen ... 384
Festlegen von Titeln und Fußnoten ... 385
Anzeigen von Eigenschaften ... 386

Kapitel 51 / X-quer- und S-Regelkarte ... **387**
Allgemeines über die 'X-quer- und S-Regelkarte' ... 387
X-quer- und S-Regelkarte: Zuweisen von Variablen zu Analyserollen ... 388
Regelkarten: Festlegen von Regelgrenzen ... 389
Regelkarte: Auswählen der durchzuführenden Tests ... 391
Regelkarten: Festlegen von Optionen für die Plotdarstellung ... 392
Regelkarten: Speichern von Analyseergebnissen ... 392
Festlegen von Titeln und Fußnoten ... 393
Anzeigen von Eigenschaften ... 394

Kapitel 52 / Gemischte Modelle ... **395**
Allgemeines über die Anwendungsroutine 'Gemischte Modelle' ... 395
Gemischte Modelle: Zuweisen von Variablen zu Analyserollen ... 396
Gemischte Modelle: Bilden von Modellen mit festen Effekten ... 397
Gemischte Modelle: Festlegen von Optionen für Modelle mit festen Effekten ... 397
Gemischte Modelle: Angeben zufälliger Effekte ... 399
Gemischte Modelle: Angeben wiederholter Effekte ... 401
Gemischte Modelle: Angeben der Mittelwerte der kleinsten Quadrate ... 403
Gemischte Modelle: Erzeugen von Plots ... 405
Gemischte Modelle: Festlegen von Vorhersageoptionen ... 406
Festlegen von Titeln und Fußnoten ... 407
Anzeigen von Eigenschaften ... 407

Kapitel 53 / Modell-Scoring ... **409**
Allgemeines über den Modell-Scoring-Assistenten ... 409
Modell-Scoring: Auswählen der Daten ... 410
Modell-Scoring: Auswählen eines Scoring-Modells ... 410
Modell-Scoring: Zuweisen von Datenvariablen zu erforderlichen Modelleingabewerten ... 411
Modell-Scoring: Auswählen der Ausgabedatenspalten ... 411
Modell-Scoring: Speichern der Ausgabedaten ... 412
Modell-Scoring: Bestätigen Ihrer Auswahl ... 412
Modell-Scoring: Modell-Scoring: Interpretieren der Ergebnisse ... 413

Kapitel 54 / Nichtlineare Regression ... **419**
Allgemeines über die Anwendungsroutine 'Nichtlineare Regression' ... 419
Nichtlineare Regression: Zuweisen von Variablen zu Analyserollen ... 420
Nichtlineare Regression: Festlegen von Modelloptionen ... 421
Nichtlineare Regression: Berechnungsoptionen festlegen ... 421
Nichtlineare Regression: Erzeugen von Plots ... 423
Nichtlineare Regression: Festlegen von Ergebnisoptionen ... 424
Festlegen von Titeln und Fußnoten ... 425

Anzeigen von Eigenschaften ... 425

Kapitel 55 / Nichtparametrische einfache ANOVA ... 427
Allgemeines über die Anwendungsroutine 'Nichtparametrische einfache ANOVA' ... 427
Nichtparametrische einfache ANOVA: Zuweisen von Variablen zu Analyserollen 428
Nichtparametrische einfache ANOVA: Auswählen der Analyse ... 429
Nichtparametrische einfache ANOVA: Berechnen exakter p-Werte ... 430
Nichtparametrische einfache ANOVA: Festlegen von Ergebnisoptionen ... 430
Festlegen von Titeln und Fußnoten ... 431
Anzeigen von Eigenschaften ... 432

Kapitel 56 / np-Regelkarte ... 433
Allgemeines über die Anwendungsroutine 'np-Regelkarte' ... 433
np-Regelkarte: Zuweisen von Variablen zu Analyserollen ... 434
Regelkarten: Festlegen von Regelgrenzen ... 436
Regelkarte: Auswählen der durchzuführenden Tests ... 437
Regelkarten: Festlegen von Optionen für die Plotdarstellung ... 438
Regelkarten: Speichern von Analyseergebnissen ... 439
Festlegen von Titeln und Fußnoten ... 440
Anzeigen von Eigenschaften ... 440

Kapitel 57 / Einfache ANOVA ... 443
Allgemeines über die Anwendungsroutine 'Einfache ANOVA' ... 443
Einfache ANOVA: Zuweisen von Variablen zu Analyserollen ... 444
Einfache ANOVA: Auswählen von Tests ... 445
Einfache ANOVA: Vergleichen von Gruppenmittelwerten ... 446
Einfache ANOVA: Auswählen von Statistiken für quantitative Variablen ... 446
Einfache ANOVA: Erzeugen von Plots ... 447
Einfache ANOVA: Festlegen von Ergebnisoptionen ... 448
Festlegen von Titeln und Fußnoten ... 449
Anzeigen von Eigenschaften ... 450

Kapitel 58 / Einfache Häufigkeiten ... 451
Allgemeines über die Anwendungsroutine 'Einfache Häufigkeiten' ... 451
Einfache Häufigkeiten: Zuweisen von Variablen zu Analyserollen ... 452
Einfache Häufigkeiten: Festlegen von Statistikoptionen ... 453
Einfache Häufigkeiten: Erzeugen von Plots ... 454
Einfache Häufigkeiten: Festlegen von Ergebnisoptionen ... 454
Festlegen von Titeln und Fußnoten ... 456
Anzeigen von Eigenschaften ... 456

Kapitel 59 / p-Regelkarte ... 459
Allgemeines über die 'p-Regelkarte' ... 459
p-Regelkarte: Zuweisen von Variablen zu Analyserollen ... 460
Regelkarten: Festlegen von Regelgrenzen ... 462
Regelkarte: Auswählen der durchzuführenden Tests ... 463
Regelkarten: Festlegen von Optionen für die Plotdarstellung ... 464
Regelkarten: Speichern von Analyseergebnissen ... 465
Festlegen von Titeln und Fußnoten ... 466
Anzeigen von Eigenschaften ... 466

Kapitel 60 / Pareto-Diagramm ... **469**
Allgemeines über die Anwendungsroutine 'Pareto-Diagramm' ... 469
Pareto-Diagramm: Zuweisen von Variablen zu Analyserollen ... 470
Pareto-Diagramm: Beschränken der Kategorieanzahl ... 471
Pareto-Diagramm: Festlegen von Optionen für die Plotdarstellung ... 473
Festlegen von Titeln und Fußnoten ... 476
Anzeigen von Eigenschaften ... 477

Kapitel 61 / Kreisdiagramm ... **479**
Allgemeines über die Anwendungsroutine 'Kreisdiagramm' ... 479
Kreisdiagramm: Auswählen des Diagrammtyps ... 480
Kreisdiagramm: Zuweisen von Variablen zu Analyserollen ... 481
Kreisdiagramm: Festlegen von Darstellungsoptionen ... 482
Festlegen von Titeln und Fußnoten ... 484
Anzeigen von Eigenschaften ... 485

Kapitel 62 / Assistent für Kreisdiagramme ... **487**
Allgemeines über die Anwendungsroutine 'Assistent für Kreisdiagramme' ... 487
Assistent für Kreisdiagramme: Auswählen der Daten ... 487
Assistent für Kreisdiagramme: Zuweisen von Variablen zu Rollen ... 488
Assistent für Kreisdiagramme: Festlegen der Darstellungsoptionen ... 489
Assistent für Kreisdiagramme: Festlegen von Titeln und Fußnoten ... 489

Kapitel 63 / P-P-Plots ... **491**
Allgemeines über die Anwendungsroutine 'P-P-Plots' ... 491
Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen ... 492
P-P-Plots: Auswählen der Verteilung ... 494
P-P-Plots: Anpassen der Verteilung ... 495
Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung ... 495
Prozessfähigkeitsanalyse: Festlegen der Analysetypen ... 499
Festlegen von Titeln und Fußnoten ... 500
Anzeigen von Eigenschaften ... 500

Kapitel 64 / Zeitreihendaten vorbereiten ... **503**
Allgemeines über die Anwendungsroutine 'Zeitreihendaten vorbereiten' ... 503
Zeitreihendaten vorbereiten: Zuweisen von Variablen zu Analyserollen ... 504
Zeitreihendaten vorbereiten: Festlegen der Ausgabehäufigkeit ... 505
Zeitreihendaten vorbereiten: Festlegen von Zeitreihenmerkmalen ... 506
Zeitreihendaten vorbereiten: Angeben der Transformationen für jede Zeitreihenvariable ... 508
Zeitreihendaten vorbereiten: Festlegen von Ergebnisoptionen ... 509
Anzeigen von Eigenschaften ... 509
Erstellen der Beobachtungszeilen für fehlende Zeilen in der Datei ... 510
Konvertieren von Zeitreihendaten von einer Stichprobenhäufigkeit in eine andere ... 511
Ändern der Beobachtungsmerkmale der Eingabedaten ... 512
Konvertieren aperiodischer Reihen in periodische Schätzer ... 513
Interpolieren fehlender Werte ... 513
Anwenden generischer Transformationen auf die Daten ... 515

Kapitel 65 / Hauptkomponenten ... **517**
Allgemeines über die Anwendungsroutine 'Hauptkomponenten' ... 517

Hauptkomponenten: Zuweisen von Variablen zu Analyserollen ... 518
Hauptkomponenten: Festlegen der Analyseoptionen ... 519
Hauptkomponenten: Erzeugen von Plots ... 520
Hauptkomponenten: Festlegen von Ergebnisoptionen ... 521
Festlegen von Titeln und Fußnoten ... 522
Anzeigen von Eigenschaften ... 523

Kapitel 66 / Wahrscheinlichkeitsplots ... 525
Allgemeines über die Anwendungsroutine 'Wahrscheinlichkeitsplots' ... 525
Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen ... 526
Wahrscheinlichkeitsplots: Auswählen der Verteilung ... 528
Wahrscheinlichkeitsplots: Anpassen der Verteilung ... 529
Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung ... 530
Prozessfähigkeitsanalyse: Festlegen der Analysetypen ... 534
Festlegen von Titeln und Fußnoten ... 535
Anzeigen von Eigenschaften ... 535

Kapitel 67 / Proportionale Hazards ... 537
Allgemeines über die Anwendungsroutine 'Proportionale Hazards' ... 537
Proportionale Hazards: Zuweisen von Variablen zu Analyserollen ... 538
Proportionale Hazards: Festlegen von Modelloptionen ... 540
Proportionale Hazards: Optionen für das Festlegen der Methoden ... 542
Proportionale Hazards: Erzeugen von Plots ... 543
Proportionale Hazards: Festlegen von Ergebnisoptionen ... 544
Festlegen von Titeln und Fußnoten ... 545
Anzeigen von Eigenschaften ... 546

Kapitel 68 / Q-Q-Plots ... 547
Allgemeines über die Anwendungsroutine 'Q-Q-Plots' ... 547
Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen ... 548
Q-Q-Plots: Auswählen der Verteilung ... 550
Q-Q-Plots: Anpassen der Verteilung ... 551
Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung ... 551
Prozessfähigkeitsanalyse: Festlegen der Analysetypen ... 555
Festlegen von Titeln und Fußnoten ... 556
Anzeigen von Eigenschaften ... 556

Kapitel 69 / Radardiagramm ... 559
Allgemeines über die Anwendungsroutine 'Radardiagramm' ... 559
Radardiagramm: Zuweisen von Variablen zu Analyserollen ... 560
Radardiagramm: Festlegen von Darstellungsoptionen ... 561
Festlegen von Titeln und Fußnoten ... 562
Anzeigen von Eigenschaften ... 562

Kapitel 70 / Zufällige Stichprobe ... 565
Allgemeines über die Anwendungsroutine 'Zufällige Stichprobe' ... 565
Zufällige Stichprobe: Erstellen einer zufälligen Stichprobe ... 566

Kapitel 71 / Rangzuweisung für Daten ... 569
Allgemeines zur Anwendungsroutine 'Rangzuweisung für Daten' ... 569
Rangzuweisung für Daten: Zuweisen von Variablen zu Analyserollen ... 570
Rangzuweisung für Daten: Festlegen der Analyseoptionen ... 570

Rangwert: Festlegen von Ergebnisoptionen . . . 572
Anzeigen von Eigenschaften . . . 573

Kapitel 72 / Recency-, Frequency- und Monetary-Analyse . . . 575
Allgemeines zur Anwendungsroutine 'RFM-Analyse' . . . 575
Recency-, Frequency- und Monetary-Analyse: Auswählen des zu analysierenden Datentyps . . . 576
Recency-, Frequency- und Monetary-Analyse: Zuweisen von Variablen zu Analyserollen . . . 577
Recency-, Frequency- und Monetary-Analyse: Kategorisieren von Daten nach Intervallen . . . 579
Recency-, Frequency- und Monetary-Analyse: Speichern von Ergebnissen . . . 580
RFM-Analyse: Auswählen von Plots . . . 580

Kapitel 73 / Regressionsanalyse mit autoregressiven Fehlern . . . 581
Allgemeines über die Anwendungsroutine 'Regressionsanalyse mit autoregressiven Fehlern' . . . 581
Regressionsanalyse mit autoregressiven Fehlern: Zuweisen von Variablen zu Analyserollen . . . 582
Regressionsanalyse mit autoregressiven Fehlern: Festlegen von Modelloptionen 583
Regressionsanalyse mit autoregressiven Fehlern: Festlegen von Statistikoptionen . . . 584
Regressionsanalyse mit autoregressiven Fehlern: Erzeugen von Plots . . . 587
Regressionsanalyse mit autoregressiven Fehlern: Festlegen von Ergebnisoptionen . . . 588
Festlegen von Titeln und Fußnoten . . . 590
Anzeigen von Eigenschaften . . . 590

Kapitel 74 / Regressionsanalyse von Paneldaten . . . 593
Allgemeines über die Anwendungsroutine 'Regressionsanalyse von Paneldaten' 593
Regressionsanalyse von Paneldaten: Zuweisen von Variablen zu Analyserollen . 594
Regressionsanalyse von Paneldaten: Festlegen von Modelloptionen . . . 595
Regressionsanalyse von Paneldaten: Festlegen von Ergebnisoptionen . . . 596
Festlegen von Titeln und Fußnoten . . . 597
Anzeigen von Eigenschaften . . . 597

Kapitel 75 / SAS ODS Statistical Graphics . . . 599
Allgemeines zu "SAS ODS Statistical Graphics" . . . 599
Verwenden des ODS Graphics Designer . . . 600
ODS Statistical Graph anzeigen . . . 600

Kapitel 76 / SAS Rapid Predictive Modeler . . . 603
Allgemeines über den SAS Rapid Predictive Modeler . . . 603
Sampling-Strategien für den SAS Rapid Predictive Modeler . . . 604
SAS Rapid Predictive Modeler: Organisieren von Daten . . . 605
SAS Rapid Predictive Modeler: Zuweisen von Variablen zu Analyserollen . . . 607
SAS Rapid Predictive Modeler: Auswählen eines Modells . . . 608
SAS Rapid Predictive Modeler: Festlegen von Berichtoptionen . . . 613
SAS Rapid Predictive Modeler: Festlegen von Ausgabeoptionen . . . 615
SAS Rapid Predictive Modeler: Registrieren Ihres Modells . . . 616

Kapitel 77 / Streuungsdiagramm ... **619**
Allgemeines über die Anwendungsroutine 'Streuungsdiagramm' ... 619
Streuungsdiagramm: Auswählen eines Diagrammtyps ... 620
Streuungsdiagramm: Zuweisen von Variablen zu Analyserollen ... 620
Streuungsdiagramm: Festlegen von Darstellungsoptionen ... 622
Festlegen von Titeln und Fußnoten ... 624
Anzeigen von Eigenschaften ... 624

Kapitel 78 / Streuungsplot-Matrix ... **627**
Allgemeines über die Anwendungsroutine 'Streuungsplot-Matrix' ... 627
Streuungsplot-Matrix: Erstellen einer Matrix mit Streuungsplots ... 628

Kapitel 79 / Daten sortieren ... **631**
Allgemeines über die Anwendungsroutine 'Daten sortieren' ... 631
Daten sortieren: Zuweisen von Variablen zu Analyserollen ... 632
Daten sortieren: Optionen für "Daten sortieren" festlegen ... 632
Daten sortieren: Festlegen von Ergebnisoptionen ... 635
Anzeigen von Eigenschaften ... 635

Kapitel 80 / Teilen von Spalten ... **637**
Allgemeines über die Anwendungsroutine 'Spalten teilen' ... 637
Spalten teilen: Zuweisen von Variablen zu Analyserollen ... 639
Spalten teilen: Festlegen von Ergebnisoptionen ... 640
Anzeigen von Eigenschaften ... 640

Kapitel 81 / Spalten stapeln ... **643**
Allgemeines über die Anwendungsroutine 'Spalten stapeln' ... 643
Spalten stapeln: Zuweisen von Variablen zu Analyserollen ... 645
Spalten stapeln: Festlegen von Optionen für Spaltenbenennungen ... 646
Spalten stapeln: Festlegen von Ergebnisoptionen ... 647
Anzeigen von Eigenschaften ... 647

Kapitel 82 / Standardisieren von Daten ... **649**
Allgemeines über die Anwendungsroutine 'Daten standardisieren' ... 649
Daten standardisieren: Zuweisen von Variablen zu Analyserollen ... 650
Daten standardisieren: Standardisierungsoptionen festlegen ... 651
Daten standardisieren: Festlegen von Ergebnisoptionen ... 651
Anzeigen von Eigenschaften ... 652

Kapitel 83 / Beschreibende Statistiken ... **653**
Allgemeines über die Anwendungsroutine 'Beschreibende Statistiken' ... 653
Beschreibende Statistiken: Zuweisen von Variablen zu Analyserollen ... 654
Beschreibende Statistiken: Auswählen beschreibender Statistiken ... 655
Beschreibende Statistiken: Festlegen der Plotoptionen ... 657
Beschreibende Statistiken: Festlegen von Ergebnisoptionen ... 658
Festlegen von Titeln und Fußnoten ... 660
Anzeigen von Eigenschaften ... 661

Kapitel 84 / Assistent für beschreibende Statistiken ... **663**
Allgemeines über den Assistenten für beschreibende Statistiken ... 663
Assistent für beschreibende Statistiken: Auswählen der Daten ... 663
Assistent für beschreibende Statistiken: Zuweisen von Variablen zu Rollen ... 664

Assistent für beschreibende Statistiken: Festlegen von Berichtsstatistiken und Ausgabergebnissen ... 665
Assistent für beschreibende Statistiken: Festlegen von Titeln und Fußnoten ... 666

Kapitel 85 / Zusammenfassungstabellen ... 669
Allgemeines über die Anwendungsroutine 'Zusammenfassungstabellen' ... 669
Zusammenfassungstabellen: Zuweisen von Variablen zu Analyserollen ... 670
Zusammenfassungstabellen: Einrichten der Zusammenfassungstabelle ... 671
Zusammenfassungstabellen: Festlegen von Ergebnisoptionen ... 674
Festlegen von Titeln und Fußnoten ... 674
Anzeigen von Eigenschaften ... 675

Kapitel 86 / Zusammenfassungstabellen-Assistent ... 677
Allgemeines über den Assistenten für Zusammenfassungstabellen ... 677
Zusammenfassungstabellen-Assistent: Auswählen der Daten ... 677
Zusammenfassungstabellen-Assistent: Zuweisen von Variablen zu Rollen ... 678
Zusammenfassungstabellen-Assistent: Auswählen der Klassifizierungsvariablen ... 679
Zusammenfassungstabellen-Assistent: Angeben der Summen ... 679
Zusammenfassungstabellen-Assistent: Auswählen zusätzlicher Ausgaben ... 680
Zusammenfassungstabellen-Assistent: Festlegen von Titeln und Fußnoten ... 680

Kapitel 87 / Oberflächenplot ... 683
Allgemeines über die Anwendungsroutine 'Oberflächenplot' ... 683
Oberflächenplot: Auswählen eines Diagrammtyps ... 684
Oberflächenplot: Zuweisen von Variablen zu Analyserollen ... 685
Oberflächenplot: Festlegen von Darstellungsoptionen ... 686
Festlegen von Titeln und Fußnoten ... 687
Anzeigen von Eigenschaften ... 688

Kapitel 88 / t-Test ... 689
Allgemeines über die Anwendungsroutine 't-Test' ... 689
t-Test: Auswahl des t-Test-Typs ... 690
t-Test: Zuweisen von Variablen zu Analyserollen ... 690
t-Test: Festlegen der Analyseoptionen ... 692
t Test: Erzeugen von Plots ... 692
Festlegen von Titeln und Fußnoten ... 693
Anzeigen von Eigenschaften ... 694

Kapitel 89 / Tabellenanalyse ... 695
Allgemeines über die Anwendungsroutine 'Tabellenanalyse' ... 695
Tabellenanalyse: Zuweisen von Variablen zu Analyserollen ... 696
Tabellenanalyse: Definieren von Tabellen ... 697
Tabellenanalyse: Auswählen von Zellenstatistiken ... 698
Tabellenanalyse: Angeben von Tabellenstatistiken ... 699
Tabellenanalyse: Festlegen von Ergebnisoptionen ... 704
Festlegen von Titeln und Fußnoten ... 706
Anzeigen von Eigenschaften ... 706

Kapitel 90 / Kacheldiagramm ... 709
Allgemeines über die Anwendungsroutine 'Kacheldiagramm' ... 709
Kacheldiagramm: Zuweisen von Variablen zu Analyserollen ... 710

Kacheldiagramm: Festlegen von Darstellungsoptionen ... 711
Kacheldiagramm: Festlegen erweiterter Optionen ... 714
Festlegen von Titeln und Fußnoten ... 715

Kapitel 91 / Transponieren ... 717
Allgemeines über die Anwendungsroutine 'Transponieren' ... 717
Transponieren: Zuweisen von Variablen zu Analyserollen ... 718
Transponieren: Festlegen der Analyseoptionen ... 719
Transponieren: Festlegen von Ergebnisoptionen ... 720
Anzeigen von Eigenschaften ... 720

Kapitel 92 / u-Regelkarte ... 723
Allgemeines über die 'u-Regelkarte' ... 723
u-Regelkarte: Zuweisen von Variablen zu Analyserollen ... 724
Regelkarten: Festlegen von Regelgrenzen ... 726
Regelkarte: Auswählen der durchzuführenden Tests ... 727
Regelkarten: Festlegen von Optionen für die Plotdarstellung ... 728
Regelkarten: Speichern von Analyseergebnissen ... 729
Festlegen von Titeln und Fußnoten ... 730
Anzeigen von Eigenschaften ... 730

Kapitel 93 / Aktualisieren von Bibliotheksmetadaten ... 733
Allgemeines über den Assistenten 'Bibliotheksmetadaten aktualisieren' ... 733
Bibliotheksmetadaten aktualisieren: Auswählen des SAS-Servers und der SAS-Bibliothek ... 734
Bibliotheksmetadaten aktualisieren: Ausführen eines Berichts für die Metadaten oder Aktualisieren der Metadaten ... 734

Kapitel 94 / Zu CAS hochladen ... 737
Allgemeines über die Anwendungsroutine 'Zu CAS hochladen' ... 737
Hochladen von Daten auf einen SAS CAS Analytic Server ... 738

Kapitel 95 / Dateien auf Server hochladen ... 739
Allgemeines über die Anwendungsroutine 'Dateien auf Server hochladen' ... 739
Dateien auf Server hochladen: Auswählen der hochzuladenden SAS-Datendateien ... 740
Dateien auf Server hochladen: Angeben der Optionen zum Hochladen ... 740

Kapitel 96 / Auf LASR hochladen ... 743
Allgemeines über die Anwendungsroutine 'Auf LASR hochladen' ... 743
Hochladen von Daten auf einen SAS LASR Analytic Server ... 744

Anhang 1 / Allgemeine Darstellungsoptionen für Grafik-Anwendungsroutinen ... 745
Festlegen von Achsenoptionen ... 745
Hinzufügen von Referenzlinien ... 748
Festlegen von Legendenoptionen ... 749
Anpassen des Diagrammbereichs ... 750
Festlegen der Balkendarstellung ... 751
Festlegen von Balkenoptionen ... 753
Festlegen der Liniendiagramm-Optionen ... 753
Auswählen der Interpolationsmethode ... 754
Festlegen erweiterter Optionen ... 756

Anhang 2 / Allgemeine Darstellungsoptionen für Regelkarten-Anwendungsroutinen ... 759
Festlegen von Achsenoptionen ... 759
Hinzufügen von Referenzlinien ... 760
Einfügungen hinzufügen ... 760
Festlegen weiterer Plotoptionen ... 760
Festlegen von Blockvariablenoptionen ... 762

Anhang 3 / Erstellen eines Modells ... 765
Allgemeines zu Modellen ... 765
Angeben von Haupteffekten ... 766
Festlegen von Kreuzeffekten (Interaktionen) ... 766
Angeben von Verschachtelungseffekten ... 766
Erstellen mehrerer Interaktionen für zwei oder mehr Variablen ... 767
Erstellen polynomialer Interaktionen quantitativer Variablen ... 767
Entfernen der Effekte der Konstante aus der Analyse ... 767

Anhang 4 / Integrierte Anwendungsroutinen ... 769
Integrierte SAS-Anwendungsroutinen und dazugehörige SAS-Prozeduren und -Lizenzen ... 769

1

Arbeiten mit Anwendungsroutinen

Integrierte Anwendungsroutinen und SAS Studio-Tasks ... 1
Siehe auch ... 3
Filtern der Daten für eine Task ... 3
Hinzufügen von Benutzercode zu einer Anwendungsroutine ... 5
Zuweisen von Spalteneigenschaften ... 6
Auswählen des Grafikausgabeformats ... 6
Arbeiten mit Anwendungsroutinenvorlagen ... 7
Allgemeines über Anwendungsroutinenvorlagen ... 7
Erstellen einer Anwendungsroutinenvorlage ... 8
Bearbeiten einer Anwendungsroutinenvorlage ... 8
Ausführen von Anwendungsroutinenvorlagen ... 9
Importieren einer Anwendungsroutinenvorlage ... 9
Exportieren einer Anwendungsroutinenvorlage ... 9

Integrierte Anwendungsroutinen und SAS Studio-Tasks

SAS Enterprise Guide und SAS Add-In für Microsoft Office werden mit vielen integrierten SAS-Anwendungsroutinen ausgeliefert. Der Zugriff auf diese Anwendungsroutinen erfolgt im Fenster **Anwendungsroutinen**. Die Liste der Anwendungsroutinen kann auch benutzerdefinierte Anwendungsroutinen und SAS Studio-Tasks enthalten.

So greifen Sie aus SAS Enterprise Guide oder SAS Add-In for Microsoft Office auf Tasks in SAS Studio zu:

1 Vergewissern Sie sich, dass SAS Studio in Ihrer Site installiert ist.

2 Stellen Sie sicher, dass der Standardserver des aktiven Profils der SAS Studio-Server ist.

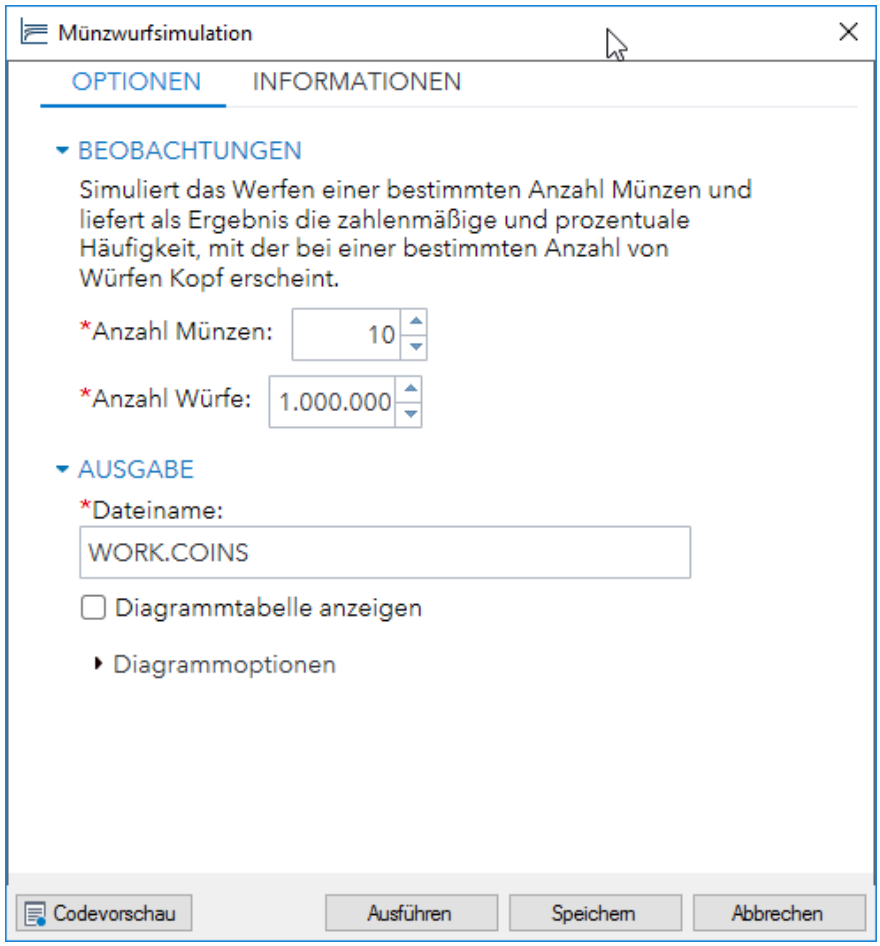
3 Wählen Sie in den Einstellungen von SAS Enterprise Guide oder SAS Add-In for Microsoft Office die Option **SAS Studio-Tasks anzeigen** aus.

- Wenn Sie SAS Studio 3.4 oder höher ausführen, können Sie entweder auf die SAS Studio Enterprise Edition oder die SAS Studio Single User Edition zugreifen. SAS Studio Single User Edition und SAS Enterprise Guide bzw. SAS Add-In for Microsoft Office müssen auf demselben Rechner installiert sein.
- In SAS Enterprise Guide können Sie auch auf Tasks aus SAS Studio 5,2 zugreifen. Geben Sie die URL und die Anmeldedaten für Ihr Release von SAS Studio an.

  Hinweis: SAS Studio 5.2-Tasks sind in SAS-Add-In for Microsoft Office nicht verfügbar.

Dadurch stehen Ihnen die SAS Studio-Tasks im Fensterabschnitt **Anwendungsroutinen** und im Fenster **Tasks** in SAS Enterprise Guide bzw. SAS Add-In for Microsoft Office zur Verfügung. Um nur die Tasks für **SAS Studio** im Fensterabschnitt **Tasks** bzw. der Task-Galerie anzuzeigen, wählen Sie **SAS Studio** als Anwendungsroutinentyp aus.

Die Bildschirmoberfläche für Tasks in SAS Studio unterscheidet sich leicht von den Dialogfeldern der integrierten Anwendungsroutinen in SAS Enterprise Guide und SAS Add-In for Microsoft Office. Hier sehen Sie beispielhaft die Benutzeroberfläche der Task **Münzwurfsimulation**.



Die Verfügbarkeit bestimmter SAS Studio-Tasks hängt davon ab, welche Lizenzen Sie besitzen und welche Software in Ihrer Site tatsächlich installiert ist. Einige SAS Studio-Tasks ermöglichen beispielsweise das Arbeiten in der SAS Viya-Umgebung. Solange Sie SAS Viya jedoch nicht lizenziert und installiert haben, stehen Ihnen diese Tasks nicht zur Verfügung.

Siehe auch

- Anhang 4, “Integrierte Anwendungsroutinen” auf Seite 769
- “SAS Studio Tasks and Associated Procedures and Licenses” in *SAS Studio: Task Reference Guide*

Filtern der Daten für eine Task

Mit Hilfe der Dropdown-Listen im Dialogfeld **Daten und Filter bearbeiten** können Sie die Filter festlegen, die Sie auf die aktuelle Datenquelle anwenden möchten.

Nachdem Sie einen Filter erstellt haben, können Sie die Makrovariable _CLIENTTASKFILTER in den Titeln und Fußnoten der Anwendungsroutine verwenden, sodass die Filterinformationen im Titel oder in der Fußnote Ihrer Ergebnisse angezeigt werden.

Hinweis: Wenn Sie das SAS Add-In for Microsoft Office verwenden, werden die Filter, die Sie im Fenster **Datenquelle ändern** festgelegt haben, zuerst angewendet. Anschließend verwendet das SAS Add-In die Filter, die Sie im Dialogfeld **Daten und Filter bearbeiten** erstellt haben.

So erstellen Sie einen Filter für Ihre Daten:

1. Klicken Sie im Fensterabschnitt **Daten** der Anwendungsroutine bzw. im Schritt **Daten überprüfen** des Assistenten auf **Bearbeiten**. Das Dialogfeld **Daten und Filter bearbeiten** wird geöffnet.
2. (Optional) Wählen Sie eine Eingabedatenquelle in SAS Enterprise Guide. Wenn Sie eine Anwendungsroutine öffnen, verwendet die Anwendungsroutine standardmäßig die aktive Datenquelle zum Einlesen. Der Name der aktiven Datenquelle wird in der Dropdown-Liste **Datenquelle** angezeigt.
3. Wählen Sie in der ersten Dropdown-Liste die zu filternde Variable aus. Die Dropdown-Liste enthält alle Variablen in Ihren Daten.

   Hinweis: Standardmäßig werden in dieser Dropdown-Liste die Variablennamen angezeigt. Wenn in dieser Dropdown-Liste die Variablenetiketten angezeigt werden sollen, aktivieren Sie das Kontrollkästchen **Etiketten statt Variablennamen anzeigen**.
4. Wählen Sie in der zweiten Dropdown-Liste einen Operator.
5. Geben Sie im dritten Feld den Wert der Variablen auf eine der folgenden Arten ein:
   - Tippen Sie den Wert der Variablen in das Textfeld ein.
   - Klicken Sie auf die Schaltfläche Wertauswahl, um eine Liste der Werte anzuzeigen, die Sie als Wert der Variablen auswählen können.

     Wenn Sie den Operator **Zwischen** oder **Nicht zwischen** wählen, müssen Sie einen Anfangs- und einen Endwert angeben. Wenn Sie die Operatoren **In einer Liste** oder **Nicht in einer Liste** ausgewählt haben, müssen Sie die Werteliste für den Vergleich angeben. Sie können Werte in das Feld eingeben oder einfügen. Klicken Sie auf **Wert hinzufügen**, um einen Wert aus der Datenquelle auszuwählen. Jeder Wert muss sich in einer neuen Zeile befinden.

     Für alle anderen Operatoren werden die ersten 25 Einzelwerte aus der Datenquelle in der Liste angezeigt. Wenn Sie mehr Werte anzeigen möchten, klicken Sie auf **Weitere Werte abrufen**. Bei jedem Klick auf die Schaltfläche **Weitere Werte abrufen** werden die nächsten 25 Einzelwerte für die Variable zusätzlich zu den bereits abgerufenen 25 Werten angezeigt.

Hinweis: Wenn sich die SAS-Datenquelle im Dateisystem Ihres Computers oder auf einem Server befindet, auf den Sie von Ihrem Computer aus Zugriff haben, wird die Werteliste in der Reihenfolge aufgefüllt, in der die Einzelwerte in den Daten vorkommen. Befindet sich die Datenquelle an einem anderen Speicherort, wird die Liste der Werte in aufsteigender Reihenfolge aufgefüllt.

6 (Optional) Wählen Sie aus der letzten Dropdown-Liste **UND** oder **ODER**, um den Daten zusätzliche Filter hinzuzufügen. Wiederholen Sie die Schritte 3 bis 5, um diese Filter zu erstellen.

Hinweis: Wenn Sie mehr als einen Filter angegeben haben, werden die Filter, die mit einer UND-Beziehung verbunden sind, zuerst ausgewertet. Die Filter werden nicht in der Reihenfolge ausgewertet, in der Sie sie erstellt haben. Angenommen, Sie haben den folgenden Filter erstellt: **`Geschlecht = 'M' oder Name = 'Carol' und Datum = '1980'`**.

In diesem Beispiel werden alle Beobachtungen, deren Name 'Carol' und deren Datum '1980' lautet herausgefiltert. Anschließend wird die ODER-Beziehung berechnet.

7 Klicken Sie auf **OK**, um die Änderungen zu speichern. Der Filter für die Eingabedatenquelle wird jetzt im Datenfenster der Anwendungsroutine angezeigt.

Hinzufügen von Benutzercode zu einer Anwendungsroutine

Mit der Option **Codevorschau** haben Sie die Möglichkeit, eigenen Code in den Code einzufügen, den SAS für die Anwendungsroutine generiert. Sie können SAS-Codeprozeduroptionen oder Anweisungen in den generierten Code einfügen und dabei alle Funktionen des SAS-Programmeditors nutzen. Sie können automatisch generierten Code im Fenster **Codevorschau für Anwendungsroutine** weder bearbeiten noch löschen, sondern nur neuen Code einfügen. Klicken Sie auf **Codevorschau**, um eigenen Code einzugeben.

1 Aktivieren Sie im Fenster **Codevorschau für Anwendungsroutine** das **Kontrollkästchen Einfügepunkte in benutzerdefiniertem Code anzeigen**. Die folgenden Zeilen werden an jeder Stelle, an der benutzerdefinierter Code eingefügt werden kann, zum Code der Anwendungsroutine hinzugefügt:

```
/* Start des benutzerdefinierten Codes (Framework_BeforeTaskCode) */
<Benutzerdefinierten Code hier einfügen>
/* Ende des benutzerdefinierten Codes (Framework_BeforeTaskCode) */
```

2 Klicken Sie auf die Position, an der der Code eingefügt werden soll, und geben Sie den Code ein.

3 Schließen Sie das Fenster **Codevorschau für Anwendungsroutine**, um Ihre Änderungen zu speichern.

Zuweisen von Spalteneigenschaften

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Klicken Sie mit der rechten Maustaste auf die Variable und wählen Sie **Eigenschaften**, um temporär ein beschreibendes Etikett und ein Format zuzuweisen. Alle zugewiesenen Eigenschaften wirken sich nur auf die aktuelle Anwendungsroutine aus.

Um ein Etikett zuzuweisen, geben Sie den Etiketttext im Feld **Etikett** ein. Das Etikett darf maximal 256 Zeichen enthalten. Die Dropdown-Liste enthält den Wert des Etiketts, das permanent mit der Variablen verknüpft ist.

In den meisten Anwendungsroutinen können Formate zugewiesen werden. So weisen Sie ein Format zu:

1 Klicken Sie auf **Ändern**. Das Fenster **Formate** wird angezeigt.

2 Wählen Sie die Kategorie des Formats im Feld **Kategorien** aus. Sie können ein SAS-Standardformat oder ein benutzerdefiniertes Format verwenden, das Sie mit der Anwendungsroutine **Format erstellen** erzeugt haben.

3 Wählen Sie den Namen des Formats im Feld **Formate** aus.

4 Geben Sie Änderungen an den Attributen im Feld **Attribute** ein und klicken Sie auf **OK**.

Hinweis: Im Fenster **Eigenschaften** ist die Eigenschaft **Sortiert** auf `Ja` gesetzt, wenn die Datei nach dieser Variablen sortiert wurde.

Auswählen des Grafikausgabeformats

In SAS Add-In for Microsoft Office können Sie folgende Grafikausgabeformate auswählen: ActiveX, ActiveX-Bild, GIF, JPEG, PNG und Java-Bilder. In SAS Enterprise Guide können Sie außerdem die Formate Java und SAS EMF auswählen.

Die Ausgabeformate ActiveX und Java sind interaktiv. Sie können mit der rechten Maustaste auf eine beliebige Grafik klicken, die in diesem Ausgabeformat generiert

wird, und dann bestimmte Optionen ändern. Die änderbaren Optionen unterscheiden sich für die Ausgabeformate ActiveX und Java. Alle Optionen, die Sie ändern, werden nur in der Ausgabe dargestellt, nicht aber in den Einstellungen, die Sie im Fenster für die Anwendungsroutine beim Erstellen der Grafik vorgenommen haben.

Die Ausgabeformate ActiveX image, GIF, JPEG, PNG, Java image und SAS EMF sind nicht interaktiv. Die Darstellung einer Grafik kann nach der Erstellung nicht mehr verändert werden.

Arbeiten mit Anwendungsroutinenvorlagen

Allgemeines über Anwendungsroutinenvorlagen

Wenn Sie eine Anwendungsroutine, wie z. B. die Anwendungsroutine **Balkendiagramm**, ausführen, möchten Sie eventuell die gleichen Einstellungen mehrmals verwenden. Sie brauchen die Anwendungsroutine nicht immer wieder zu öffnen, um diese Einstellungen für unterschiedliche Eingabedatenquellen anzugeben, denn Sie können Ihre Einstellungen für eine bestimmte Anwendungsroutine in einer Anwendungsroutinenvorlage speichern. Sie können dann diese Vorlage mit einer beliebigen Eingabedatenquelle ausführen.

Hinweis: Beachten Sie bei der Arbeit mit Anwendungsroutinenvorlagen die folgenden Einschränkungen:

- Werte oder Einstellungen, die von der Datenquelle abhängig sind, werden nicht in der Anwendungsroutinenvorlage gespeichert. Beispiel: Wenn Sie eine Variable einer bestimmten Rolle zuweisen, wird die Variablenzuweisung nicht in der Anwendungsroutinenvorlage gespeichert. Wenn eine Option von den Werten der Datenquelle abhängig ist, wird die Einstellung der Option ebenfalls nicht in der Anwendungsroutinenvorlage gespeichert. So werden beispielsweise für die Anwendungsroutine **Liniendiagramm** keine Symboldefinitionen in der Anwendungsroutinenvorlage gespeichert, die auf den Datenwerten basieren.
- Aus den folgenden Anwendungsroutinen heraus ist es nicht möglich, Anwendungsroutinenvorlagen zu erstellen: Tabelle anhängen, Daten vergleichen, Dateien und Formate löschen, Dateien auf PC herunterladen, Forecast Studio - Projekt erstellen, Forecast Studio - Projekt öffnen, Forecast Studio - Wertekorrektur-Projekt, Listenbericht, Modell-Scoring und Zusammenfassungstabellen.
- Wenn Sie eine Eingabeaufforderung in der Anwendungsroutine verwenden und dann eine Vorlage der Anwendungsroutine erstellen, wird die

Eingabeaufforderung in SAS Enterprise Guide nicht als Teil der Anwendungsroutinenvorlage gespeichert.

Hinweis: Darüber hinaus können Sie im SAS Enterprise Guide im Fenster **Abfrage erstellen** eine Vorlage erstellen.

Sie können Ihre Vorlagen in SAS Enterprise Guide mit dem Manager für Anwendungsroutinevorlagen verwalten.

Erstellen einer Anwendungsroutinenvorlage

1. Öffnen Sie die Anwendungsroutine und geben Sie die Optionen für die Vorlage an.
2. Klicken Sie neben **Ausführen** auf den Pfeil nach unten, und wählen Sie im Kontextmenü die Option **Vorlage erstellen**. Das Dialogfeld **Anwendungsroutinenvorlage erstellen** wird geöffnet.
3. Geben Sie einen Namen und eine Beschreibung für die Vorlage ein.
4. Geben Sie an, ob eine neue Anwendungsroutinenvorlage innnerhalb Gruppe erstellt oder eine vorhandene Vorlage ersetzt werden soll.
5. Klicken Sie auf **Erstellen**.

Bearbeiten einer Anwendungsroutinenvorlage

Sie können die von Ihnen erstellten Vorlagen ändern. Bevor Sie jedoch eine Vorlage ändern, die auf einer Abfrage beruht, müssen Sie zunächst prüfen, ob die Datenverweise noch gültig sind.

1. Öffnen Sie die Anwendungsroutinenvorlage, die Sie bearbeiten möchten. Das Anwendungsroutinenfenster wird mit den in der Vorlage festgelegten Optionen geöffnet.
2. Legen Sie die gewünschten Optionen für die Anwendungsroutine fest und wählen Sie **Ausführen** ⇨ **Vorlage erstellen**. Das Dialogfeld **Anwendungsroutinenvorlage erstellen** wird geöffnet.
3. Aktivieren Sie das Kontrollkästchen **Bestehende Anwendungsroutinenvorlage ersetzen** und wählen Sie die Vorlage aus, die Sie ersetzen möchten. Klicken Sie auf **Erstellen**. Die Änderungen werden in der Anwendungsroutinenvorlage gespeichert.

Ausführen von Anwendungsroutinenvorlagen

1. Öffnen Sie eine Anwendungsroutinenvorlage durch Auswahl von **Anwendungsroutinen** ⇨ **Anwendungsroutinevorlagen** ⇨ ***name-der-gruppe*** ⇨ ***name-der-vorlage***.
2. (Optional) Wählen Sie die Eingabedatenquelle. Standardmäßig ist die aktive Datenquelle im Projekt gleichzeitig die Eingabedatenquelle für die Anwendungsroutine. Bei den meisten Anwendungsroutinen können Sie die Eingabedatenquelle im Fenster **Daten** ändern.
3. Weisen Sie im Panel **Daten** des Anwendungsroutinefensters den Anwendungsroutinenrollen Variablen aus der Eingabedatenquelle zu.
4. (Optional) Ändern Sie weitere Optionen in der Anwendungsroutine.
5. Klicken Sie auf **Ausführen**.

Importieren einer Anwendungsroutinenvorlage

1. Klicken Sie im Fenster **Anwendungsroutinen** auf **Vorlage importieren**.
2. Klicken Sie auf **Hinzufügen**, um Vorlagen für den Import auszuwählen.
3. Wählen Sie in der Dropdown-Liste **Importieren in Gruppe** die Gruppe, in der Sie die importierten Vorlagen speichern möchten.
4. Klicken Sie auf **Importieren**. Die Vorlagen werden nun im Fensterabschnitt **Anwendungsroutinen** angezeigt und sind im Menü **Anwendungsroutinen** verfügbar.

Exportieren einer Anwendungsroutinenvorlage

1. Klicken Sie im Fenster **Anwendungsroutinen Vorlage exportieren**.
2. Wählen Sie die zu exportierenden Vorlagen aus und fügen Sie dieses zum Feld **Ausgewählte Anwendungsroutinenvorlagen** hinzu.
3. Um den Speicherort für diese Vorlagen auszuwählen, klicken Sie auf **Exportieren nach**.
4. Klicken Sie auf **Exportieren**. Eine Kopie der Vorlagen wird am ausgewählten Speicherort gespeichert.

2

Anwendungsroutine 'Tabelle anhängen'

Allgemeines über die Anwendungsroutine 'Tabelle anhängen' 11
Tabelle anhängen: Erstellen einer Liste der Eingabetabellen 12
Tabelle anhängen: Speichern von Ergebnissen 12

Allgemeines über die Anwendungsroutine 'Tabelle anhängen'

Die Anwendungsroutine **Tabelle anhängen** hängt zwei oder mehrere Eingabetabellen aneinander, um eine einzelne Ausgabedatei oder eine SQL-Ansicht zu erstellen. Sie können maximal 256 Tabellen miteinander verketten. Die entstandene Datei enthält jede Zeile und jede Variable aus allen Eingabetabellen. Eine Variable, die in einigen, aber nicht in allen Eingabetabellen vorkommt, weist in jeder Zeile der Ergebnisse, welche von einer Eingabetabelle stammen, in der diese Variable nicht enthalten ist, einen fehlenden Wert auf.

Beispiel: Sie haben vier Eingabedateien für die einzelne Quartale im Jahr 2002, in denen die Umsätze für ein bestimmtes Produkt aufgelistet sind. Um nun eine einzelne Ausgabedatei mit alle Umsätzen für 2002 zu erstellen, können Sie die Anwendungsroutine **Tabelle anhängen** verwenden.

Hinweis: Die Anwendungsroutine **Tabelle anhängen** ist nur im SAS Enterprise Guide verfügbar.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SQL
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	Ohne

Tabelle anhängen: Erstellen einer Liste der Eingabetabellen

Klicken Sie im Auswahlfenster auf **Tabellen**, um die zu verkettenden Tabellen auszuwählen und deren Reihenfolge zu bestimmen.

Der Bereich **Anzuhängende Tabellen** führt die zu verkettenden Tabellen in der Reihenfolge auf, in der Sie verkettet werden. Erstellen Sie in diesem Bereich die gewünschte Tabellenliste.

1. Klicken Sie auf **Tabelle hinzufügen**. Das Dialogfenster **Daten öffnen** wird angezeigt. Wählen Sie die Tabelle aus.
2. Wählen Sie im folgenden Dialogfeld den Namen der Tabelle aus, die sie hinzufügen möchten. Um mehr als eine Tabelle auszuwählen, wählen Sie bei gedrückt gehaltener Strg-Taste die Tabellen aus, die hinzugefügt werden sollen.
3. Klicken Sie auf **Öffnen** oder **OK**, um die Tabelle dem Bereich **Anzuhängende Tabellen** hinzuzufügen.

Hinweis: Sie können die Tabelle, mit der Sie die Anwendungsroutine **Tabelle anhängen** geöffnet haben, nicht löschen.

Tabelle anhängen: Speichern von Ergebnissen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um aus Ihren Daten eine SAS-Datei oder eine SQL-Ansicht zu erstellen.

Im Bereich **Ergebnisformat** können Sie wählen, ob Sie die Ergebnisse Ihrer angehängten Daten als SAS-Datei oder als SQL-Ansicht speichern möchten.

- Um die Ergebnisse als SAS-Datei zu speichern, klicken Sie auf **Datentabelle**. Hierbei handelt es sich um die Standardeinstellung.
- Um Ihre Ergebnisse als SQL-Ansicht zu speichern, klicken Sie auf **Datenansicht**.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

3

Flächendiagramm

Allgemeines über die Anwendungsroutine 'Flächendiagramm' ... 15
Flächendiagramm: Auswählen eines Diagrammtyps ... 16
Flächendiagramm: Zuweisen von Variablen zu Analyserollen ... 16
Flächendiagramm: Festlegen von Darstellungsoptionen ... 18
Anpassen einer Fläche ... 18
Weitere Darstellungsoptionen ... 18
Festlegen von Titeln und Fußnoten ... 19
Anzeigen von Eigenschaften ... 19

Allgemeines über die Anwendungsroutine 'Flächendiagramm'

Die Anwendungsroutine **Flächendiagramm** erstellt Flächen-, Spline-, Stufen- oder Überlagerungsdiagramme, die die mathematische Abhängigkeit zweier Variablen darstellen. Der Bereich unterhalb der Kurve wird mit einem Muster oder einer Farbkombination ausgefüllt, um die Lesbarkeit zu verbessern.

Weitere Informationen, siehe "Auswählen des Grafikausgabeformats" auf Seite 6.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GPLOT
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Flächendiagramm: Auswählen eines Diagrammtyps

Klicken Sie im Auswahlfenster auf **Flächendiagramm**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Flächen-Liniendiagramm	erstellt mithilfe der Werte der beiden Variablen ein Liniendiagramm. Die Fläche unter der Diagrammlinie ist schattiert.
Flächen-Spline-Plot	erstellt mithilfe der Spline-Routine ein Liniendiagramm. Die Fläche unter dem Spline-Plot ist schattiert.
Flächen-Stufendiagramm	erstellt mithilfe der Stufenfunktion ein Liniendiagramm. Die Fläche unter der Stufendiagrammlinie ist schattiert.
Mehrere vertikale sich überlagernde Flächendiagramme	erstellt ein separates Diagramm von jeder Y-Variablen vs. derselben X-Variablen. Aufgrund der Überlagerungsoption werden alle Diagramme in demselben Diagramm angezeigt. Die Fläche unter den Diagrammen wird jeweils in einer anderen Farbe schattiert.
Mehrere Flächendiagramme nach Spalte gruppiert	erstellt ein separates Diagramm für jeden eindeutigen Wert der Variablen, die der Rolle **Gruppe** zugewiesen ist. Alle Diagramme werden in derselben Grafik angezeigt. Die Fläche unter den Diagrammen wird jeweils in einer anderen Farbe schattiert.

Flächendiagramm: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Horizontal	Die Spalte, die Sie dieser Rolle zuweisen, ist die horizontale oder X-Achsenvariable für das Diagramm.
Vertikal	Die Spalte, die Sie dieser Rolle zuweisen, ist die vertikale oder Y-Achsenvariable für das Diagramm.
Vertikal (Rechts)	Die Spalte, die Sie dieser Rolle zuweisen, ist die vertikale bzw. Y-Achsenvariable für die rechte Seite der Grafik. Diese Variable wird gegen die X-Achsenvariable dargestellt, sodass ein anderer Plot auf der Grafik erstellt wird.
Gruppe	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen, wie viele Plotlinien im gleichen Diagramm gezogen werden.
Diagramme gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Diagramme, die erstellt werden. Für jeden eindeutigen Wert wird eine separate Grafik erzeugt. Aktivieren Sie das Kontrollkästchen **Gleiche Achsenskalierung für alle Diagramme**, wenn Sie möchten, dass für alle erzeugten Diagramme die gleiche Achsenskalierung verwendet wird. Standardmäßig basiert der Wertebereich für jede Achse auf den Minimal- und Maximalwerten der dargestellten Daten und könnte daher in den Diagrammen variieren. Wenn Sie diese Option auswählen haben alle Diagramme den gleichen Bereich für jede Achse.

Flächendiagramm: Festlegen von Darstellungsoptionen

Anpassen einer Fläche

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Fläche**, um auf diese Optionen zuzugreifen.

Wählen Sie einen Flächenfüllstil:

- **Leer**
- **Nicht transparent**
- **Linie**

Wenn Sie **Linie** auswählen, bestimmen Sie das Muster für den Füllstil und die Liniendichte.

Hinweis: Es werden nicht alle Flächenfüllstile von allen Grafikausgabeformaten unterstützt. Die Flächenfüllstile **Leer** oder **Linie** wirken sich nicht auf die Formate ActiveX, Java, ActiveX-Bild oder Java-Bild aus. Diese Grafikausgabeformate zeigen die Flächendiagramme genauso an wie der Flächenfüllstil **Ausgefüllt**.

Weitere Darstellungsoptionen

Einige Darstellungsoptionen gelten für mehrere Grafik-Anwendungsroutinen. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- "Festlegen von Achsenoptionen" auf Seite 745
- "Hinzufügen von Referenzlinien" auf Seite 748
- "Festlegen von Legendenoptionen" auf Seite 749
- "Anpassen des Diagrammbereichs" auf Seite 750

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

4

Anwendungsroutine 'ARIMA-Modellierung und -Prognosen'

Allgemeines über die Anwendungsroutine 'ARIMA-Modellierung und -Prognosen' 21

ARIMA-Modellierung und -Prognosen: Zuweisen von Variablen zu Analyserollen 22

ARIMA-Modellierung und -Prognosen: Festlegen von Identifizierungsoptionen 23
Differenzenbildung bei der Wirkungsreihe 23
Festlegen von Stationaritätstests 24
Erzeugen von Plots und Speichern von Ergebnissen 24

ARIMA-Modellierung und -Prognosen: Festlegen von Schätzoptionen 25
Aktivierung der Schätzungsschritte 25
Definieren des ARIMA-Modells 25
Auswählen von Schätzmethode und -optionen 26
Speichern von Ergebnissen 27

ARIMA-Modellierung und -Prognosen: Festlegen von Prognoseoptionen 28
Aktivieren der Prognoseschritte 28
Festlegen des Prognoseintervalls 28
Erzeugen von Plots und Speichern von Ergebnissen 29

Festlegen von Titeln und Fußnoten 30

Anzeigen von Eigenschaften 30

Allgemeines über die Anwendungsroutine 'ARIMA-Modellierung und -Prognosen'

ARIMA-Prognosenmodellierung analysiert und prognostiziert gleichmäßig verteilte, univariate Zeitreihendaten, Transferfunktionsdaten und Interventionsdaten mit Hilfe des Modells ARIMA (AutoRegressive Integrated Moving Average) oder ARMA

(Autoregressive Moving Average). Ein ARIMA-Modell sagt einen Wert in einer Zeitreihe als lineare Kombination der eigenen vergangenen Werte, vergangener Fehler (auch Shocks oder Innovationen genannt) und aktueller und vergangener Werte anderer Zeitreihen voraus.

Es gibt drei grundlegende Stufen bei der Durchführung einer ARIMA-Analyse:

- Stufe 1: Zeitreihen identifizieren
- Stufe 2: Zeitreihen schätzen
- Stufe 3: Künftige Werte der Zeitreihen prognostizieren.

Sie können die Anwendungsroutine **ARIMA-Modellierung und Prognose** verwenden, um folgende Analysen durchzuführen:

- Bestimmen der Auswirkung saisonaler Faktoren. Beispielsweise können Sie eine Zeitreihe mit Wettervariablen (Temperatur, Niederschlagsmenge, Feuchtigkeit usw.) als Eingaben verwenden.
- Vorhersagen eines Produktbedarfs auf der Grundlage historischer Muster.

Hinweis: Die Anwendungsroutine **ARIMA-Modellierung und -Prognosen** ist so strukturiert, dass sich jede der drei Stufen auf einer eigenen Seite befindet. Wenn Sie auf **Fertigstellen** klicken, werden nur die Stufe der jeweils ausgewählten Seite und die vorhergehenden Stufen ausgeführt. Wenn Sie also zu Seite **Stufe 3: Prognoseerstellung** gehen und Änderungen durchführen und dann zur Seite **Stufe 1: Identifikation** gehen und auf **Ausführen** klicken, wird nur die Identifikationsstufe ausgeführt.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	ARIMA
Erforderliche SAS-Produkte	Base SAS, SAS/ETS
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

ARIMA-Modellierung und -Prognosen: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Zeitreihenvariable	gibt die Variable für die Prognose an. Sie müssen dieser Rolle genau eine Variable zuweisen.
Zeit-ID-Variable	gibt die Variable an, mit der den Zeilen Daten zugeordnet werden. Sie müssen dieser Rolle genau eine Variable zuweisen. Die erste Variable, die ein SAS-Datums-, Uhrzeit- oder Datetime-Format enthält, wird dieser Rolle automatisch zugewiesen. Sie können auch eine Zeit-ID-Variable erstellen.
Prognosen gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits als abhängige Variable oder als unabhängige Variable ausgewählt ist.

ARIMA-Modellierung und -Prognosen: Festlegen von Identifizierungsoptionen

Differenzenbildung bei der Wirkungsreihe

Klicken Sie im Auswahlfenster unter der Überschrift Stufe 1: Identifikation auf **Differenzen**, um auf diese Optionen zuzugreifen.

Um für die Wirkungsreihe Differenzen zu bilden, markieren Sie das Kontrollkästchen **Differenzen der Wirkungsreihe**. Wenn Sie diese Kontrollkästchen markieren, werden standardmäßig erste Differenzen gebildet. Um den Grad der Differenzen zu ändern, geben Sie die neuen Werte in das Feld **Lags-Differenzen** ein. Geben Sie beispielsweise 1,1 an, um bei Lag Eins zweimal Differenzen zu bilden.

Festlegen von Stationaritätstests

Klicken Sie im Auswahlfenster unter der Überschrift **Stufe 1: Identifizierung** auf **Stationaritätstests**, um auf diese Optionen zuzugreifen.

Um einen Stationaritätstest anzugeben, wählen Sie die Optionsschaltfläche für den gewünschten Test aus. Standardmäßig werden keine Tests durchgeführt. Wenn Sie einen Stationaritätstest auswählen, können Sie die autoregressiven Ordnungen für den Test angeben. Standardmäßig werden autoregressive Ordnungen von 0, 1 und 2 angegeben. Geben Sie im Feld **Autoregressive Ordnungen** die gewünschte autoregressiven Ordnungen durch Kommata getrennt an, oder geben Sie einen Ordnungsbereich ein, der mit Auslassungspunkten getrennt ist. Geben Sie z.B. `1...3` an, um autoregressive Ordnungen von 1, 2 und 3 anzugeben.

Wenn Sie **Erweiterte Dickey-Fuller-Tests durchführen** auswählen, können Sie einen saisonalen Lag zur Durchführung eines saisonalen Dickey-Fuller-Tests angeben. Geben Sie einen ganzzahligen Wert in das Feld **Saisonaler Lag** ein. Der Standard für den saisonalen Lag ist `1`. Dies ist ein nichtsaisonaler Test.

Erzeugen von Plots und Speichern von Ergebnissen

Klicken Sie im Auswahlfenster unter der Überschrift **Stufe 1: Identifizierung** auf **Plots und Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Plot mit tatsächlichen Werten	zeigt in den Ergebnissen einen Plot der eingegebenen Zeitreihen an. Sie können die Anzahl der Lags angeben, die beim Berechnen der Autokorrelationen und Kreuzkorrelationen berücksichtigt werden sollen. Die Standardanzahl an Lags beträgt `24` oder ein Viertel der Anzahl der Beobachtungen, je nachdem, welche Zahl kleiner ist
Autokorrelationen und Kreuzkovarianzen speichern	speichert eine Ausgabetabelle mit Autokorrelationen und Kreuz-Kovarianzen. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist

Optionsname	Beschreibung
	und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Angezeigte Identifizierungsausgabe unterdrücken	unterdrückt die Identifizierungsausgabe aus den Ergebnissen.

ARIMA-Modellierung und -Prognosen: Festlegen von Schätzoptionen

Aktivierung der Schätzungsschritte

Klicken Sie im Auswahlfenster unter der Überschrift **Stufe 2: Schätzung** auf **Schätzungsschritte aktivieren**, um auf diese Optionen zuzugreifen.

Um die Optionen für die Schätzungsschritte zu aktivieren, wählen Sie das Kontrollkästchen **Schätzungsschritte ausführen**.

Hinweis: Die Anwendungsroutine **ARIMA-Modellierung und -Prognosen** ist in drei Stufen unterteilt. Die Optionen für **Stufe 1: Identifizierung** sind immer aktiviert. Sie müssen die entsprechenden Kontrollkästen auswählen, um die Optionen für **Stufe 2: Schätzung** und **Stufe 3: Prognose** zu aktivieren. Die Schätzschritte müssen aktiviert sein, damit Sie die Prognoseschritte aktivieren können.

Definieren des ARIMA-Modells

Klicken Sie im Auswahlfenster unter der Überschrift **Stufe 2: Schätzung** auf **Modelldefinition**, um auf diese Optionen zuzugreifen.

Sie können der Modellschätzung sowohl autoregressive als auch Moving-Average-Parameter hinzufügen.

So fügen Sie der Modellschätzung Parameter hinzu:

1. Geben Sie in die Felder **Faktoren für AR-Modell** oder **Faktoren für MA-Modell** einen ganzzahligen Wert, eine durch Kommata getrennte Liste mit Ganzzahlen oder einen Bereich mit Ganzzahlen an, der die Lags für die Parameter Autoregressiv oder Moving Average darstellt, die für jeden multiplikativen Faktor in der Liste geschätzt werden müssen. Geben Sie z.B. `1...3` , `1 TO 3` oder `1,2,3` ein, um (1, 2, 3) anzugeben; geben Sie `3` ein, um (3) anzugeben.
2. Klicken Sie auf **Hinzufügen**, oder drücken Sie die Eingabetaste, um der Liste den neuen Parameter hinzuzufügen.
3. Wiederholen Sie diesen Vorgang, um multiplikative Terme hinzuzufügen. Geben Sie beispielsweise `1` ein, klicken Sie auf **Hinzufügen**, geben Sie `12` ein und klicken Sie auf **Hinzufügen**, um (1)(12) anzugeben.

Nachdem der Parameter erstellt wurde, können Sie diesen bearbeiten oder aus der Liste in den Bereichen **Autoregressiv** bzw. **Moving Average** löschen.

Auswählen von Schätzmethode und -optionen

Klicken Sie im Auswahlfenster unter der Überschrift **Stufe 2:** Schätzung auf **Modelloptionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Schätzmethode	Wählen Sie eine Schätzmethode aus der Dropdown-Liste aus. Folgende Schätzmethoden stehen zur Verfügung: ■ Die **Maximum-Likelihood-Methode** erstellt Maximum-Likelihood-Schätzer. Die Likelihood-Funktion wird über nichtlineare kleinste Quadrate mithilfe der Marquardt-Methode maximiert. ■ Die **unbedingte Kleinste-Quadrate-Methode** erstellt Schätzungen für unbedingte kleinste Quadrate. Die ULS-Methode (Unconditonal Least Squares) wird auch ELS-Methode (Exact least Squares) genannt. ■ Die **bedingte Kleinste-Quadrate-Methode** erstellt Schätzungen für bedingte kleinste Quadrate. Diese Schätzung ist bedingt unter der Annahme, dass die letzten nichtbeobachteten Fehler kleiner 0 sind. Dies ist die Standardmethode.
Maximale Anzahl Iterationen	gibt die maximal zulässige Anzahl an Iterationen für die Stufe Schätzung an. Geben Sie einen ganzzahligen Wert im Feld **Maximale Anzahl an Iterationen** an. Der Standardwert beträgt `50`.

Optionsname	Beschreibung
Mittelwertsterm aus Modell entfernen	Schließt den Mittelwertsterm aus dem Modell aus.

Speichern von Ergebnissen

Klicken Sie im Auswahlfenster unter der Überschrift **Stufe 2: Schätzung** auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Ausgabedaten speichern	Sie können Ihre Ergebnisse teilweise oder vollständig in einer Ausgabetabelle speichern. Aktivieren Sie die entsprechenden Kontrollkästchen, um Ihre Ergebnisse zu speichern. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Angezeigte Schätzungsausgabe unterdrücken	Unterdrückt die Ausgabe der Schätzstufe.

ARIMA-Modellierung und -Prognosen: Festlegen von Prognoseoptionen

Aktivieren der Prognoseschritte

Klicken Sie im Auswahlfenster unter der Überschrift **Stufe 3: Prognose** auf **Prognoseschritte aktivieren**, um auf diese Optionen zuzugreifen.

Um die Optionen für die Prognoseschritte zu aktivieren, aktivieren Sie das Kontrollkästchen **Prognoseschritte ausführen**.

Hinweis: Die Anwendungsroutine **ARIMA-Modellierung und -Prognosen** ist in drei Stufen unterteilt. Die Optionen für **Stufe 1: Identifizierung** sind immer aktiviert. Sie müssen die entsprechenden Kontrollkästen auswählen, um die Optionen für **Stufe 2: Schätzung** und **Stufe 3: Prognose** zu aktivieren. Die Schätzschritte müssen aktiviert sein, damit Sie die Prognoseschritte aktivieren können.

Festlegen des Prognoseintervalls

Klicken Sie im Auswahlfenster unter der Überschrift **Stufe 3: Prognose** auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Zeitintervall zwischen Beobachtungen	Wählen Sie ein Zeitintervall aus der Dropdown-Liste aus, welches das Zeitintervall zwischen Beobachtungen in der Eingabedatentabelle darstellt. Das gewählte Intervall muss für die Eingabedaten geeignet sein, andernfalls schlägt die Analyse fehl oder erzeugt ungenaue Ergebnisse. Wenn Sie **Anzahl der Einheiten** als Zeitintervall auswählen, können Sie auch die Zeiteinheiten pro Intervall angeben. Der Standardwert ist 1. Bei Zeitintervallen, für die **Anzahl der Einheiten** nicht zutrifft, ist das Feld **Zeiteinheiten pro Intervall** nicht verfügbar.

Optionsname	Beschreibung
Zu prognostizierende Intervalle	Geben Sie die Anzahl der Zeitintervalle an, für die Sie Prognosen erzeugen möchten. Der Standardwert beträgt `24`. Wenn Sie z.B. Monat als Zeitintervall zwischen Beobachtungen eingeben, und 6 Zeitintervalle für die Prognose angeben, würden Sie Prognosen für 6 Monate nach Ende der Eingabedaten erhalten. Wählen Sie ein Konfidenzniveau aus der Dropdown-Liste **Konfidenzniveau**. Der Standardwert ist ein Konfidenzniveau von 95%.

Erzeugen von Plots und Speichern von Ergebnissen

Klicken Sie im Auswahlfenster unter der Überschrift **Stufe 3: Prognose** auf **Plots und Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Prognoseplots - Optionen	Erzeugt folgende Plots: ■ Aktivieren Sie das Kontrollkästchen **Prognosen**, um einen Plot für Istwerte, prognostizierte Werte (Ein-Schritt-voraus), mehrstufige Prognosen und Konfidengrenzen zu erstellen. ■ Wenn Sie Plots der Residuen erzeugen möchten, aktivieren Sie das Kontrollkästchen **Residuen**. Verwenden Sie die Felder **Intervalle der angezeigten Daten begrenzen** und **Intervalle zeigen**, um die Anzahl der Intervalle historischer Daten zu bestimmen, die vor den Prognosedaten dargestellt werden sollen.
Prognosen speichern	Um eine Ausgabetabelle für Prognosen zu erzeugen, markieren Sie das Kontrollkästchen **Prognosen speichern**. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Optionsname	Beschreibung
Angezeigte Prognoseausgabe unterdrücken	Unterdrückt die Prognoseausgabe aus den Ergebnissen.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine

- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

5

Zuweisen einer Projektbibliothek

Allgemeines über den Projektbibliotheks-Assistenten 33
Projektbibliothek zuweisen: Angeben eines Namens und eines Servers 34
Projektbibliothek zuweisen: Angeben einer Engine 34
Projektbibliothek zuweisen: Festlegen von Konfigurationsoptionen 37
Projektbibliothek zuweisen: Überprüfen der Zusammenfassungsinformationen 38

Allgemeines über den Projektbibliotheks-Assistenten

Mithilfe des **Projektbibliotheks-Assistenten** können Sie eine Bibliothek für das Projekt definieren. Der Assistent generiert die LIBNAME-Anweisung, die zum Definieren der Bibliothek erforderlich ist. Diese Bibliothek wird dann jedes Mal definiert, wenn Sie die Anwendungsroutine oder den Prozessfluss ausführen, der die Anwendungsroutine enthält.

Wenn Ihnen z.B. Daten vorliegen, die analysiert werden müssen. Sie befinden sich in Ihrem lokalen Verzeichnis und nicht in einer von Ihrem Administrator definierten SAS-Bibliothek. Sie müssen eine SAS-Bibliothek definieren, um auf die Daten zugreifen zu können. Sie können den **Projektbibliotheks-Assistenten** zur Erstellung dieser LIBNAME-Anweisung verwenden.

Hinweis: Diese Anwendungsroutine ist nur im SAS Enterprise Guide verfügbar.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	Ohne

Name der Anforderung	Prozedur- und Produktnamen
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	Ohne

Projektbibliothek zuweisen: Angeben eines Namens und eines Servers

1. Geben Sie den Namen der Bibliothek in das Feld **Name** ein. Der Name muss in Großbuchstaben geschrieben sein und darf maximal acht Zeichen enthalten. Der Name muss für jeden Server auf dem die Bibliothek erstellt wird eindeutig sein. Der Name darf die folgenden Zeichen nicht enthalten: \ / : * ? " < > |
2. Wählen Sie den Server in der Liste der verfügbaren Server aus.
3. Klicken Sie auf **Weiter**, um eine Engine für die Bibliothek auszuwählen.

Projektbibliothek zuweisen: Angeben einer Engine

1. Bevor Sie eine Engine auswählen können, müssen Sie für die Bibliothek einen Namen und einen Server angeben.
2. Wählen Sie den Engine-Typ aus. Folgende Typen sind verfügbar:

Engine-Typ	Beschreibung
Dateisystem	Geben Sie folgende Optionen an: ■ Damit SAS die Engine anhand des Typs der Dateien auswählen kann, die sich in dem für die Bibliothek angegebenen Pfad befinden, aktivieren Sie das Kontrollkästchen **SAS soll die Engine basierend auf den Inhalten des angegebenen Pfads auswählen**.

Engine-Typ	Beschreibung
	Hinweis: Aktivieren Sie dieses Kontrollkästchen nicht, wenn der Bibliothekspfad verschiedene Dateitypen enthält. ■ Wählen Sie die Engine für die Bibliothek aus. Wenn Sie **REMOTE-SAS/SHARE** als Engine verwenden, müssen Sie das Name-Wert-Paar slibref/SAMPSHR in den Konfigurationsoptionen für die Bibliothek festlegen. **Hinweis:** Sie können keine Engine auswählen, wenn Sie das Kontrollkästchen **SAS soll die Engine basierend auf den Inhalten des angegebenen Pfads auswählen** aktiviert haben. ■ Geben Sie in das Feld **Pfad** den physischen Pfad ein, unter dem die Bibliothek auf dem Server gespeichert ist. Klicken Sie auf **Durchsuchen**, um diesen Speicherort auszuwählen. Sie können im Pfad Platzhalter verwenden. <userid> ersetzt die Windows-Benutzer-ID <sample> ersetzt das Verzeichnis Sample im Enterprise Client <serveruser> ersetzt die Benutzer-ID für die Anmeldung am Server <serverpassword> ersetzt das Passwort für die Anmeldung am Server <libraryuser> ersetzt die Benutzer-ID für die Anmeldung an der Bibliothek <librarypassword> ersetzt das Passwort für die Anmeldung an der Bibliothek Wenn Benutzer-ID und Passwort für einen Zugriff auf den ausgewählten Server erforderlich sind, werden Sie aufgefordert, diese einzugeben. **Hinweis:** Wenn Sie **BASE**, **V9**, **V8**, **V7**, **V604** oder **V6** als Engine auswählen, brauchen Sie keinen Pfad anzugeben. Für alle anderen Engines ist ein Pfad erforderlich.
Datenbanksystem	Geben Sie folgende Optionen an: ■ Wählen Sie die Engine, die für den Zugriff auf die Dateien in der Bibliothek verwendet werden soll. ■ Wählen Sie den externen Datenbankserver für die Bibliothek. **Hinweis:** In der Dropdown-Liste werden die Datenbankserver angezeigt, die im aktiven SAS

Engine-Typ	Beschreibung
	Metadata Repository festgelegt wurden. Wenn der gewünschte Server in den Metadaten nicht festgelegt wurde, ist dieser Server auch nicht in der Dropdown-Liste aufgeführt. Ihr Administrator kann zum Definieren weiterer Server SAS Management Console verwenden. ■ Geben Sie das Datenbankschema an, das für den Zugriff auf die Daten auf dem Server verwendet werden soll. Der angegebene Schemaname muss mit dem Namen eines Schemas übereinstimmen, das Sie bereits auf dem Datenbankserver definiert haben. ■ Geben Sie einen Benutzernamen und ein Passwort für den Datenbankserver ein.
WebDAV	Geben Sie folgende Optionen an: ■ Wählen Sie den WebDAV-Server, der als Quelle für die Bibliothek verwendet werden soll. **Hinweis:** In der Dropdown-Liste werden die WebDAV-Server angezeigt, die im aktiven SAS Metadata Repository festgelegt wurden. Wenn der gewünschte Server in den Metadaten nicht festgelegt wurde, ist dieser Server auch nicht in der Dropdown-Liste aufgeführt. Ihr Administrator kann zum Definieren weiterer Server SAS Management Console verwenden. ■ Geben Sie in das Feld **Pfad** den physischen Pfad ein, unter dem die Bibliothek auf dem Server gespeichert ist. Sie können diese Platzhalterzeichen verwenden, wenn Sie den physischen Pfad, unter dem die Bibliothek auf dem Server gespeichert ist, oder die Bibliotheksoptionen angeben. <userid> ersetzt die Windows-Benutzer-ID <sample> ersetzt das Verzeichnis Sample im Enterprise Client <serveruser> ersetzt die Benutzer-ID für die Anmeldung am Server <serverpassword> ersetzt das Passwort für die Anmeldung am Server <libraryuser> ersetzt die Benutzer-ID für die Anmeldung an der Bibliothek <librarypassword> ersetzt das Passwort für die Anmeldung an der Bibliothek

Engine-Typ	Beschreibung
	■ Geben Sie einen Benutzernamen und ein Passwort für den WebDav-Server ein.

3 Klicken Sie auf **Weiter**, um für die Bibliothek die Konfigurationsoptionen festzulegen.

Projektbibliothek zuweisen: Festlegen von Konfigurationsoptionen

1 Bevor Sie die Konfigurationsoptionen festlegen können, müssen Sie für die Bibliothek eine Engine auswählen.

2 Legen Sie die Bibliotheksoptionen fest, die Name-Wert-Paare besitzen. Sie können nur Optionen mit verknüpften Werten in der Tabelle angeben. Beispiel für eine Option mit einem Name-Wert-Paar: ACCESS=READONLY. Diese Option kennzeichnet die Bibliothek als schreibgeschützt, so dass Sie Ausgabedaten nicht in diese Bibliothek schreiben können. Eine Liste der verfügbaren Optionen finden Sie in der Dokumentation zu Ihrer Betriebssystemumgebung.

Hinweis: Wenn Sie **REMOTE-SAS/SHARE** als Engine verwenden, müssen Sie das Name-Wert-Paar slibref/SAMPSHR in den Konfigurationsoptionen für die Bibliothek festlegen.

So geben Sie einen Optionsnamen oder -wert ein:

a Klicken Sie in der Tabelle auf eine leere Zelle. Wenn Sie einen Optionsnamen und -wert angeben, wird automatisch eine leere Zeile in der Tabelle angezeigt.

b Geben Sie den Namen der Option in die Spalte **Name** ein. Sie können jede beliebige gültige LIBNAME-Anweisungsoption angeben, Sie sollten jedoch keine Eingabeaufforderungsoptionen verwenden (wie z.B. DBPROMPT=YES).

c Geben Sie den der Option entsprechenden Wert in die Spalte **Wert** ein. Ein leerer Wert ist nicht zulässig.

Hinweis: Wenn die Bibliothek mit einer Sybase-Datenbank verbunden wird, müssen Sie READLOCK_TYPE=PAGE als Option einbeziehen.

3 Geben Sie alle Optionen, die keine Werte erfordern, in das Feld **Weitere Optionen** ein.

4 Klicken Sie auf **Weiter**, um die Zusammenfassungsinformationen zu prüfen.

Projektbibliothek zuweisen: Überprüfen der Zusammenfassungsinformationen

Bevor Sie die neue Bibliothek erstellen, können Sie die angegebenen Informationen prüfen.

1 Bevor Sie die Zusammenfassungsinformationen prüfen, müssen Sie die Konfigurationsoptionen für die Bibliothek angeben.

2 Prüfen Sie die Informationen im Textfeld. Um einen Wert zu ändern, klicken Sie auf **Zurück**.

3 Um die neue Bibliothek auf dem ausgewählten Server zu testen, klicken Sie auf **Bibliothek testen**. Wenn SAS Enterprise Guide die Verbindung zum Server herstellt und die Bibliothek zuweist, wird als Status dieses Servers **OK** angezeigt. Wenn Sie das Testlog anzeigen möchten, klicken Sie auf **Log anzeigen**.

4 Klicken Sie auf **Fertig stellen**, um die neue Bibliothek zu erstellen.

6

Balkendiagramm

Allgemeines über die Anwendungsroutine 'Balkendiagramm' ... **39**
Balkendiagramm: Auswählen des Diagrammtyps ... **40**
Balkendiagramm: Zuweisen von Variablen zu Analyserollen ... **42**
Balkendiagramm: Festlegen von Darstellungsoptionen ... **43**
Festlegen des Layouts ... 43
Weitere Darstellungsoptionen ... 43
Festlegen von Titeln und Fußnoten ... **44**
Anzeigen von Eigenschaften ... **44**

Allgemeines über die Anwendungsroutine 'Balkendiagramm'

Die Anwendungsroutine **Balkendiagramm** erstellt vertikale, horizontale oder dreidimensionale Balkendiagramme, die numerische Werte oder Statistiken zwischen unterschiedlichen Werten einer Diagrammvariablen vergleichen. Balkendiagramme zeigen die relative Größe von Daten durch Balken unterschiedlicher Höhe. Jeder Balken stellt eine Datenkategorie dar.

Sie können mit einem Balkendiagramm die Gesamtumsätze jeder Geschäftsfiliale vergleichen. Bei diesem Diagrammtyp stellt jeder Balken die Gesamtumsätze für jede Filiale dar.

Weitere Informationen, siehe “Auswählen des Grafikausgabeformats” auf Seite 6.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GCHART

Name der Anforderung	Prozedur- und Produktnamen
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Balkendiagramm: Auswählen des Diagrammtyps

Klicken Sie im Auswahlfenster auf **Balkendiagramm**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Einfach	zeigt den Betrag der Daten durch Balken an, wobei jeder Balken einer Datenkategorie entspricht. Die Länge (oder Höhe) der Balken stellt den Wert der Diagrammstatistik für den entsprechenden Midpoint dar.
Gruppiert	erstellt eine separate Balkengruppe für jeden eindeutigen Wert der Variablen, die der Rolle **Balken gruppieren nach** zugewiesen wurde.
Gestapelt	erstellt ein separates Segment innerhalb jedes Balkens für jeden eindeutigen Wert der Variablen, die der Rolle **Stapel** zugewiesen wurde. Jedes Segment innerhalb eines Balkens wird in einer anderen Farbe dargestellt.
Gruppiert/ Gestapelt	kombiniert die gruppierten und die gestapelten Balkendiagramme. Diese Option erstellt eine separate Balkengruppe für jeden eindeutigen Wert der Variablen, die der Rolle **Balken gruppieren nach** zugewiesen, sowie ein separates Segment innerhalb der einzelnen Balken für jeden eindeutigen Wert der Variablen, die der Rolle **Stapel** zugewiesen wurde.
Farbig	erstellt ein einfaches Balkendiagramm, in dem jeder Balken mit einer anderen Farbe dargestellt wird.
Gruppiert farbig	erstellt eine separate Balkengruppe für jeden eindeutigen Wert der Variablen, die der Rolle **Balken gruppieren nach** zugewiesen wurde. Jeder Balken innerhalb der Gruppe wird in einer anderen Farbe dargestellt.

Optionsname	Beschreibung
Gruppierte farbige Gruppen	erstellt eine separate Balkengruppe für jeden eindeutigen Wert der Variablen, die der Rolle **Balken gruppieren nach** zugewiesen wurde. Die Balken innerhalb einer Gruppe weisen dieselbe Farbe auf.
Verschachtelt	erstellt eine separate Balkengruppe für jeden eindeutigen Wert der Variablen, die der Rolle **Balken gruppieren nach** zugewiesen wurde. Die Balken innerhalb der einzelnen Gruppen werden angehängt, und jeder Balken wird in einer anderen Farbe dargestellt.
3D Gruppiert	erstellt eine separate Gruppe dreidimensionaler Balken für jeden eindeutigen Wert der Variablen, die der Rolle **Balken gruppieren nach** zugewiesen wurde.
3D Gruppierte/ Gestapelte	kombiniert die gruppierten und die gestapelten Balkendiagramme. Diese Option erstellt eine separate Gruppe dreidimensionaler Balken für jeden eindeutigen Wert der Variablen, die der Rolle **Balken gruppieren nach** zugewiesen, sowie ein separates Segment innerhalb der einzelnen Balken für jeden eindeutigen Wert der Variablen, die der Rolle **Stapel** zugewiesen wurde.
3D Gruppiert farbig	erstellt eine separate Gruppe dreidimensionaler Balken für jeden eindeutigen Wert der Variablen, die der Rolle **Balken gruppieren nach** zugewiesen wurde. Jeder Balken innerhalb der Gruppe wird in einer anderen Farbe dargestellt.
3D Gruppierte farbige Gruppen	erstellt eine separate Gruppe dreidimensionaler Balken für jeden eindeutigen Wert der Variablen, die der Rolle **Balken gruppieren nach** zugewiesen wurde. Die Balken innerhalb einer Gruppe weisen dieselbe Farbe auf.
Gruppen mit mehreren Maßen	ermöglicht die Zuweisung mehrerer Analysevariablen zur Rolle Summe von. Das daraus resultierende Diagramm enthält eine Gruppe von Balken für jeden eindeutigen Wert der Variablen, die der Rolle **Spalte für Diagramm** zugewiesen wurde. Die Balken innerhalb einer Gruppe weisen dieselbe Farbe auf.

Balkendiagramm: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Hinweis: Die jeweils verfügbaren Rollen hängen vom ausgewählten Diagrammtyp ab.

Optionsname	Beschreibung
Spalte für Diagramm	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Datenkategorien des Diagramms. Sie müssen dieser Rolle genau eine Variable zuweisen.
Balken gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Gruppen, in welche die Balken unterteilt werden. Sie müssen dieser Rolle genau eine Variable zuweisen.
Stapel	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Segmente in den einzelnen Balken.
Summe von	Die Spalte, die Sie dieser Rolle zuweisen, bestimmt die Länge der Balken. Um die bestimmte Statistik zur Bestimmung der Balkenlänge zu wählen, klicken Sie im Auswahlfenster auf **Erweitert**. Wenn Sie dieser Rolle keine Spalte zuweisen, wird die Länge der Balken oder die Größe der Abschnitte über die Häufigkeit jedes Werts der Spalte **Spalte für Diagramm** bestimmt.
Diagramme gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Diagramme, die erstellt werden. Für jeden eindeutigen Wert wird eine separate Grafik erzeugt.

Balkendiagramm: Festlegen von Darstellungsoptionen

Festlegen des Layouts

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Layout**, um auf diese Optionen zuzugreifen.

Standardmäßig sind die meisten Balkendiagramme zweidimensional und das Kontrollkästchen **2D** ist aktiviert. Um ein dreidimensionales Diagramm zu erstellen, deaktivieren Sie das Kontrollkästchen **2D** und wählen aus der Dropdown-Liste **Form** eine Balkenform.

Hinweis: Das Kontrollkästchen **2D** ist nur dann verfügbar, wenn Sie ein dreidimensionales Balkendiagramm als Diagrammtyp ausgewählt haben.

Mit Hilfe der Dropdown-Liste **Sortierreihenfolge** können Sie die Sortierreihenfolge für die Balken bestimmen. Standardmäßig wird die Sortierreihenfolge automatisch festgelegt. Sie können die Balken in aufsteigender oder absteigender Reihenfolge der Länge oder Höhe anordnen.

Aktivieren Sie das Kontrollkästchen **Umrissfarbe**, wenn Sie jeden Balken mit einem Umriss versehen möchten. Sie können eine Farbe für den Umriss aus der Dropdown-Liste auswählen.

Mit Hilfe der Dropdown-Liste **Balkengröße** können Sie die Balkengröße bestimmen. Standardmäßig wird die Balkengröße automatisch festgelegt. Sie können jedoch die Balkenbreite oder den Abstand zwischen den Balken als Prozentwert der gesamten Diagrammbreite im Textfeld angeben.

Weitere Darstellungsoptionen

Einige Darstellungsoptionen gelten für alle Grafik-Anwendungsroutinen. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Festlegen der Balkendarstellung” auf Seite 751
- “Festlegen von Achsenoptionen” auf Seite 745
- “Hinzufügen von Referenzlinien” auf Seite 748
- “Festlegen von Legendenoptionen” auf Seite 749

- “Anpassen des Diagrammbereichs” auf Seite 750
- “Festlegen erweiterter Optionen” auf Seite 756

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse

- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

7

Assistent für Balkendiagramme

Allgemeines über den Assistenten für Balkendiagramme 47
Assistent für Balkendiagramme: Auswählen von Daten 47
Assistent für Balkendiagramme: Zuweisen von Variablen zu Rollen 48
Assistent für Balkendiagramme: Festlegen der Darstellungsoptionen 49
Assistent für Balkendiagramme: Festlegen von Titeln und Fußnoten 50

Allgemeines über den Assistenten für Balkendiagramme

Der **Assistent für Balkendiagramme** unterstützt Sie bei der Erstellung von vertikalen oder horizontalen, zwei- oder dreidimensionalen Balkendiagrammen mit gruppierten oder gestapelten Balken.

Assistent für Balkendiagramme: Auswählen von Daten

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Klicken Sie auf **Weiter**, um Variablen Rollen zuzuweisen.

Assistent für Balkendiagramme: Zuweisen von Variablen zu Rollen

1. Bevor Sie Variablen zu Rollen zuweisen können, müssen Sie die Daten auswählen.
2. Weisen Sie den folgenden Rollen Variablen zu:
   - **Balken** - Die Werte der Variablen, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der verschiedenen Balken. Die Variable kann entweder alphanumerisch oder numerisch sein. Sie müssen dieser Rolle genau eine Variable zuweisen. Sie können auswählen, die Balken zu sortieren oder den Variablentyp anzugeben:
     - **Sortieren** – Legen sie die Sortierreihenfolge der Balken fest. Standardmäßig sortiert die Anwendungsroutine die Balken automatisch in aufsteigender Reihenfolge des Midpoint-Werts. Sie können die Balken nach aufsteigender oder nach absteigender Balkenhöhe sortieren.
     - **Variablentyp** – Standardmäßig wird die Anzahl der Balken automatisch bestimmt. Bei numerischen Spalten werden die Werte in Bereiche geteilt, mit einem Balken pro Bereich. Wenn die numerische Spalte diskrete Werte darstellt, können Sie die Option **Diskret** auswählen.
   - **Balkenhöhe** - Die Variable oder die Statistik, die Sie dieser Rolle zuweisen, bestimmt die Länge der Balken.
     - Wenn Sie dieser Rolle eine Variable zuweisen, können Sie auswählen, ob die Summe oder der Mittelwert jedes Balkens berechnet wird. Standardmäßig wird die Summe berechnet.
     - Wenn Sie dieser Rolle keine Variable zugewiesen haben, können Sie die Häufigkeit, die kumulative Häufigkeit, den Prozentwert oder den kumulativen Prozentwert für jeden Balken berechnen. Standardmäßig wird die Häufigkeit berechnet.

Hinweis: Wenn Sie ein horizontales Balkendiagramm erstellen heißt diese Option **Balkenlänge**.

3. (Optional) Weisen Sie folgenden Rollen Variablen zu:
   - **Gruppieren nach** - Die Werte der Variablen, die Sie dieser Rolle zuweisen, bestimmen die Gruppen, in welche die Balken unterteilt werden. Bei vertikalen Balkendiagrammen werden die Balkengruppen auf der X-Achse angezeigt. Bei horizontalen Balkendiagrammen werden die Balkengruppen auf der Y-Achse angezeigt.

Hinweis: Diese Option ist nicht verfügbar, wenn Sie der Rolle **Tiefe** eine Variable zuweisen.

- **Tiefe** - Die Werte der Variablen, die Sie dieser Rolle zuweisen, bestimmen die Gruppen, in welche die Balken unterteilt werden. Die Balkengruppen werden auf der Z-Achse angezeigt.

  Hinweis: Diese Option ist nicht verfügbar, wenn Sie ein horizontales Balkendiagramm erstellen, oder wenn Sie der Rolle **Gruppieren nach** eine Variable zuweisen.

- **Stapeln nach** - Die Werte der Variablen, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Segmente in jedem Balken innerhalb eines Balkendiagramms.
- **Gruppiert nach** - Für jede Gruppe werden separate Diagramme generiert. Die Gruppen werden von den Werten der Variablen bestimmt, die Sie dieser Rolle zuweisen.

4 Klicken Sie auf **Weiter**, um die Darstellungsoptionen festzulegen.

Assistent für Balkendiagramme: Festlegen der Darstellungsoptionen

1 Bevor Sie die Diagrammdarstellung festlegen können, müssen Sie die Zuweisung von Variablen zu Rollen vornehmen.

2 Aktivieren Sie das Kontrollkästchen **3D-Diagramm**, um ein dreidimensionales Balkendiagramm zu erstellen. Standardmäßig sind die Balken zweidimensional.

3 Wählen Sie in der Dropdown-Liste **Farbige Balken** die Farben für die Balken im Diagramm aus. Die Optionen **Alle Balken gleich** und **Balkenkategorie** sind standardmäßig verfügbar. Wenn Sie den Rollen **Gruppieren nach** oder **Tiefe** eine Variable zuweisen, sind ebenso die Optionen **Gruppierungsvariable** bzw. **Tiefenvariable** verfügbar.

Hinweis: Wenn Sie der Rolle **Stapeln nach** eine Variable zugewiesen haben, ist diese Option nicht verfügbar.

Wenn alle Balken die gleiche Farbe haben, wählen Sie aus der Dropdown-Liste **Balkenfarbe** eine Farbe aus. Wenn jede Balkenkategorie eine andere Farbe hat, wird das Farbschema vom Assistenten für Balkendiagramme automatisch ausgewählt.

4 Wenn eine Legende verfügbar ist, aktivieren Sie das Kontrollkästchen **Legende**, um sie in das Diagramm einzuschließen. Wählen Sie aus der Dropdown-Liste die Position für die Legende aus.

5 Wenn Sie die Statistik angeben möchten, die als Etikett für jeden Balken angezeigt werden soll, aktivieren Sie das Kontrollkästchen **Datenetiketten** und wählen anschließend die anzuzeigende Statistik aus der Dropdown-Liste. Wenn Sie der Rolle **Balkenhöhe** keine Variable zugewiesen haben, sind die Statistiken Häufigkeit, kumulative Häufigkeit, Prozentwert und kumulativer Prozentwert verfügbar. Wenn Sie der Rolle **Balkenhöhe** eine Variable zugewiesen haben, dann sind ebenso die Statistiken Summe und Mittelwert verfügbar.

6 Wenn Sie die Etiketten für jede zugewiesene Variable bearbeiten möchten, klicken Sie auf **Achsenetiketten**. Das Fenster **Achsenetiketten** wird geöffnet.

7 Wenn Sie Referenzlinien in das Diagramm einschließen möchten, aktivieren Sie das Kontrollkästchen **Referenzlinien verwenden**.

8 Wenn Sie Markierungen in das Diagramm einschließen möchten, aktivieren Sie das Kontrollkästchen **Markierungen**.

9 Klicken Sie auf **Weiter**, um Titel und Fußnoten anzugeben.

Assistent für Balkendiagramme: Festlegen von Titeln und Fußnoten

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

1 Vor dem Angeben der Titel und Fußnoten müssen Sie die Darstellungsoptionen festlegen.

2 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld. Sie können in Titeln und Fußnoten Makrovariablen verwenden.

3 Klicken Sie auf **Fertig stellen**, um die Anwendungsroutine auszuführen.

8

Balken-Linien-Diagramm

Allgemeines über die Anwendungsroutine 'Balken-Liniendiagramm' **51**

Balken-Linien-Diagramm: Zuweisen von Variablen zu Analyserollen **52**

Balken-Linien-Diagramm: Festlegen von Darstellungsoptionen **53**

Auswählen der Interpolationsmethode 53

Weitere Darstellungsoptionen 54

Festlegen von Titeln und Fußnoten **54**

Anzeigen von Eigenschaften **55**

Allgemeines über die Anwendungsroutine 'Balken-Liniendiagramm'

Die Anwendungsroutine **Balken-Liniendiagramm** erstellt ein vertikales Balkendiagramm mit einem überlagerten Liniendiagramm. Das Liniendiagramm stellt den Wert einer Statistik dar, die für eine der Variablen in der Eingabedatei berechnet wird.

Mit dieser Anwendungsroutine können folgende Aufgaben erledigt werden:

- Exakte und relative Beträge anzeigen und vergleichen
- Den Beitrag einzelner Teile an der Gesamtheit untersuchen
- Trends und Muster in den Daten ermitteln

Sie können mit dieser Anwendungsroutine beispielsweise die Beziehung zwischen Produktumsätzen und -kosten für jedes Land zeigen, in dem das Produkt verkauft wird. Die horizontale Achse ist die Midpoint-Achse und enthält einen Balken pro Land. Die vertikalen Achsen sind die abhängigen Achsen. Jeder Balken stellt die Umsätze für jedes Land dar. Auf der vertikalen rechten Achse wird ein Liniendiagramm über die Balken gelegt, das die Produktions- und Marketing-Kosten

des Produkts in jedem Land darstellt. Dieses Diagramm ermöglicht Ihnen, die Kosten und die Umsätze jedes Landes einzeln und länderübergreifend zu vergleichen.

Weitere Informationen, siehe “Auswählen des Grafikausgabeformats” auf Seite 6.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GBARLINE
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Balken-Linien-Diagramm: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Spalte für Diagramm	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Datenkategorien des Diagramms. Sie müssen dieser Rolle genau eine Variable zuweisen.
Balkensumme von	Die Spalte, die Sie dieser Rolle zuweisen, bestimmt die Länge der Balken. Wenn dieser Rolle keine Spalte zugewiesen wurde, wird standardmäßig die Gesamtzahl der Beobachtungen (die Häufigkeitsstatistik) verwendet, um die Länge der Balken zu bestimmen. Wenn dieser Rolle eine Spalte zugewiesen wurde, bestimmt die Summe der Werte für jede Kategorie die Länge der Balken. Um die

Optionsname	Beschreibung
	bestimmte Statistik zur Bestimmung der Balkenlänge zu wählen, klicken Sie im Auswahlfenster auf **Erweitert**. Weitere Informationen, siehe "Festlegen erweiterter Optionen".
Liniensumme von	Die Spalte, die Sie dieser Rolle zugewiesen haben, gibt die Variablen an, die für die Liniendiagramme verwendet werden. Sie können dieser Rolle mehr als eine Variable zuweisen. Um die Statistik auszuwählen, die zur Darstellung jeder Linie verwendet wird, wählen Sie **Plots** im Auswahlfenster. Standardmäßig wird die Summe der Werte für jede Kategorie für die Darstellung der einzelnen Linien verwendet. **Hinweis:** Wenn Sie dieser Rolle keine Spalte zuweisen, wird die Gesamtzahl der Beobachtungen für jede Kategorie (die Häufigkeitsstatistik) für die Darstellung der Linie verwendet. Um die Statistik auszuwählen, die zur Darstellung der Linie verwendet wird, wählen Sie **Erweitert** im Auswahlfenster.
Stapel	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Segmente in den einzelnen Balken.
Diagramme gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Diagramme, die erstellt werden. Für jeden eindeutigen Wert wird eine separate Grafik erzeugt.

Balken-Linien-Diagramm: Festlegen von Darstellungsoptionen

Auswählen der Interpolationsmethode

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Interpolationen**, um auf diese Optionen zuzugreifen.

1. Wählen Sie eine Variable aus der Liste aus. Hierbei handelt es sich um die Variablen, die Sie der Rolle **Liniensumme von** zugewiesen haben.
2. Wählen Sie eine Interpolationsmethode aus der Dropdown-Liste aus. Je nach ausgewählter Interpolationsmethode kann die Angabe weiterer Einstellungen erforderlich sein. Die folgenden Interpolationsmethoden sind verfügbar:

- **Linie** - Verbindet die Datenpunkte durch Geraden. Die Punkte werden in der Reihenfolge verbunden, in der sie in der Eingabedatei vorkommen.
- **Streuung** unterdrückt die Interpolation. Es wird eine Darstellung der Datenpunkte erstellt.
- **Stufe** - stellt die Daten mithilfe einer Stufenfunktion dar.

3 (Optional) Aktivieren Sie das Kontrollkästchen **Für alle übernehmen**, um die gleiche Interpolationsmethode und die gleichen Eigenschaftseinstellungen für alle Variablen der Liste zu verwenden. Diese Option ist nur dann verfügbar, wenn sich mehr als eine Variable in der Liste befinden.

Weitere Darstellungsoptionen

Einige Darstellungsoptionen gelten für alle Grafik-Anwendungsroutinen. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- "Festlegen der Balkendarstellung" auf Seite 751
- "Festlegen von Balkenoptionen" auf Seite 753
- "Festlegen der Liniendiagramm-Optionen" auf Seite 753
- "Festlegen von Achsenoptionen" auf Seite 745
- "Hinzufügen von Referenzlinien" auf Seite 748
- "Festlegen von Legendenoptionen" auf Seite 749
- "Anpassen des Diagrammbereichs" auf Seite 750
- "Festlegen erweiterter Optionen" auf Seite 756

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

9

Einfache Prognosen

Allgemeines über die Anwendungsroutine 'Einfache Prognosen' ... **57**

Einfache Prognosen: Zuweisen von Variablen zu Analyserollen ... **58**

Einfache Prognosen: Festlegen von Prognoseoptionen ... **59**
Auswählen von Prognosemethoden und Intervallen ... 59
Auswählen von Eingabedatenoptionen ... 60
Festlegen von Optionen für Prognosemethoden ... 61

Einfache Prognosen: Erzeugen von Plots ... **62**

Einfache Prognosen: Festlegen von Ergebnisoptionen ... **62**

Festlegen von Titeln und Fußnoten ... **63**

Anzeigen von Eigenschaften ... **63**

Allgemeines über die Anwendungsroutine 'Einfache Prognosen'

Die Anwendungsroutine **Einfache Prognosen** bietet eine schnelle und automatisierte Möglichkeit, in einem Schritt Prognosen für zahlreiche Zeitreihen zu erzeugen. Diese Anwendungsroutine verwendet extrapolative Prognosemethoden, wobei die Prognosen für eine Reihe nur Funktionen der Zeit- und Vergangenheitswerte der Reihen und nicht anderer Variablen sind.

Sie können diese Anwendungsroutine zum Vorhersagen der Umsatzzahlen für ein Produkt verwenden. Angenommen, Sie haben die monatlichen Umsatzzahlen für einige Produkte in einer Datei gespeichert und möchten die Umsatzzahlen für die nächsten 10 Monate vorhersagen. Mit dieser Anwendungsroutine können Sie die Daten für die Prognosewerte und Konfidenzgrenzen der vorhergesagten Umsatzzahlen berechnen. Sie können diese Ausgabe in einer SAS-Datei speichern. Außerdem können Sie die Parameter der verwendeten Prognosemodelle sowie

Statistiken speichern, die messen, wie gut die Prognosemodelle die Daten für eine SAS-Ausgabedatei anpassen. Schließlich können Sie die Prognosen, Konfidenzgrenzen, Residuen und tatsächlichen Werte darstellen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	FORECAST
Erforderliche SAS-Produkte	Base SAS, SAS/ETS
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Einfache Prognosen: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Prognosevariable	gibt die Variablen für die Prognose an. Sie müssen dieser Rolle genau eine Variable zuweisen.
Zeit-ID-Variable	gibt die Variable an, mit der den Zeilen Daten zugeordnet werden. Sie müssen dieser Rolle genau eine Variable zuweisen. Die erste Variable, die ein SAS-Datums-, Uhrzeit- oder Datetime-Format enthält, wird dieser Rolle automatisch zugewiesen. Sie können auch eine Zeit-ID-Variable erstellen.
Prognosen gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig

Optionsname	Beschreibung
	werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können nicht nach einer Variablen gruppieren, die bereits einer anderen Rolle zugewiesen ist.

Einfache Prognosen: Festlegen von Prognoseoptionen

Auswählen von Prognosemethoden und Intervallen

Klicken Sie im Auswahlfenster auf **Prognoseoptionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Prognosemethode	Um eine Prognosemethode anzugeben, wählen Sie die gewünschte Methode aus der Dropdown-Liste aus. Hinweis ■ **Schrittweise autoregressiv** (für autoregressive Prozesse) passt ein Zeittrendmodell an die Serien an und verwendet die Differenz zwischen den einzelnen Werten und dem geschätzten Trend. Anschließend wird die verbleibende Variation mithilfe eines autoregressiven Modells angepasst. Dies ist die Standardmethode. ■ **Exponentielles Glätten** (für Moving-Average-Prozesse) passt ein Trendmodell so an, dass die neuesten Daten eine schwerere Gewichtung als die Daten in den frühen Teilen der Serie erhalten. Die Gewichtung einer Beobachtung ist eine geometrische (exponentielle) Funktion der Periodenanzahl, welche die Beobachtung relativ zur aktuellen Periode in die Vergangenheit ausdehnt. ■ Die Methode **Winters (Multiplikative) Methode** verwendet Aktualisierungsgleichungen, die denen bei der exponentiellen Glättung gleichen, um die Parameter an das

Optionsname	Beschreibung
	Modell anzupassen. In diesem Modell haben Sie zwei Trendparameter und eine Funktion, die den der Zeit entsprechenden saisonalen Parameter auswählt. ■ Die **Winters Additive Methode** ähnelt der Winters (Multiplikative) Methode, addiert jedoch die saisonalen Parameter zum Trend, anstatt diese mit dem Trend zu multiplizieren. **Hinweis:** Das Auswählen einer Prognosemethode aktiviert bzw. deaktiviert andere Steuerelemente in diesem Fenster.
Zu prognostizierende Intervalle	Geben Sie die Anzahl der Zeitintervalle, für die Sie Prognosen erzeugen lassen möchten, im Feld **Zu prognostizierende Intervalle** ein. Der Standardwert ist `12`. Wenn Sie z.B. Monat als Zeitintervall zwischen Beobachtungen eingeben, und 6 Lead-Zeitintervalle für die Prognose angeben, würden Sie Prognosen für 6 Monate nach Ende der Eingabedaten erhalten.

Auswählen von Eingabedatenoptionen

Klicken Sie im Auswahlfenster auf **Prognoseoptionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Zeitintervall zwischen Beobachtungen	gibt das Zeitintervall zwischen den Zeilen in der Eingabedatentabelle an. Der Standardzeitintervall ist **Anzahl der Einheiten**. Wenn Sie **Anzahl der Einheiten** auswählen, können Sie auch die Zeiteinheiten pro Intervall auswählen. Bei anderen Zeitintervallen ist **Zeiteinheiten pro Intervall** deaktiviert. Das gewählte Intervall muss für die Eingabedaten geeignet sein, andernfalls schlägt die Analyse fehl oder erzeugt ungenaue Ergebnisse.
Saisonale Zykluslänge	gibt die saisonale Zykluslänge sowie die Intervalle pro saisonalem Zyklus für eine saisonale Prognosemethode an (Winters multiplikative Methode oder Winters additive Methode). Die verfügbaren saisonalen Zykluslängen hängen vom ausgewählten Zeitintervall ab. Der Standard für die saisonale Zykluslänge ist **Intervallanzahl**.

Festlegen von Optionen für Prognosemethoden

Klicken Sie im Auswahlfenster auf **Prognoseoptionen**, um auf diese Optionen zuzugreifen.

Plot-Name	Beschreibung
Grad des Zeittrendmodells	Um den Grad des Zeittrendmodells anzugeben, wählen Sie einen Wert aus der Dropdown-Liste aus. Es stehen folgende Modelle zur Verfügung: **Konstant**, **Linear** und **Quadratisch**. Standardmäßig verwenden alle Prognosemethoden ein lineares Zeittrendmodell. Ausnahme ist die exponentielle Glättungsmethode, die ein quadratisches Modell verwendet.
Konfidenzniveau	Wählen Sie ein Konfidenzniveau aus der Dropdown-Liste aus. Der Standardwert ist ein Konfidenzniveau von 95%.
Glättungsgewichtungen	Sie können glättende Gewichtungen für die Methoden **Exponentiell glättend**, **Winters Multiplikative** und **Winters Additive** angeben. Bei der Methode **Exponentiell glättend** können Sie nur für den Konstantenterm eine glättende Gewichtung angeben. Die Methoden **Winters Multiplikative** und **Winters Additive** lassen die Angabe einer glättenden Gewichtung für die Konstantenkomponente, die lineare und quadratische Komponente sowie die saisonale Komponente zu. Die lineare und quadratische Komponente wird nur verfügbar, wenn Sie ein konstantes Glättungsgewicht angeben. Analog können Sie eine saisonale glättende Gewichtung nur angeben, nachdem Sie eine für die lineare und quadratische Komponente angegeben haben. Der Standard für die glättenden Gewichtungen hängt vom angegebenen Grad des Zeittrendmodells ab sowie von den anderen angegebenen glättenden Gewichtungen. Um eine glättende Gewichtung zu ändern, markieren Sie das relevante Kontrollkästchen und geben einen neuen Wert in das Feld ein.

Einfache Prognosen: Erzeugen von Plots

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Markieren Sie die Kontrollkästchen im Bereich **Typen**, um Plots der tatsächlichen Werte, Prognosedaten und Residuen zu erzeugen.

Für Prognosedatenplots können Sie diese Daten einbeziehen

- **Tatsächliche Werte und Ein-Schritt-Prognosen** umfasst die tatsächlichen Datenwerte und die Ein-Schritt-Prognosen in den Plot. Eine Ein-Schritt-Prognose bezieht sich auf eine Prognose, die während der vorherigen Zeitperiode (*t*-1) für eine Zeitperiode in der Zukunft getroffen wurde.
- **Konfidenzgrenzen der Prognosen** berücksichtigt die Konfidenzgrenzen für jeden prognostizierten Wert in der Darstellung.

Verwenden Sie bei allen Plottypen das Kontrollkästchen **Anzeige tats. Daten begrenzen** und das Feld **Intervalle zeigen**, um die Anzahl der Intervalle historischer Daten zu bestimmen, die vor den Prognosedaten dargestellt werden sollen.

Hinweis: Das Kontrollkästchen **Anzeige tats. Daten begrenzen** ist nicht verfügbar, wenn Sie der Rolle **Prognosen gruppieren nach** eine Variable zuweisen.

Einfache Prognosen: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Die Datei mit den Prognosewerten wird immer gespeichert. Um eine permanent gespeicherte Tabelle mit Parameterschätzern zu erzeugen, markieren Sie das Kontrollkästchen **Parameterschätzer**. Wählen Sie weitere Statistiken aus, die aufgenommen werden sollen.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Sie können wählen, ob Sie Prognosen und Parameterschätzer in die Ergebnisse einschließen möchten. Erst wenn Sie das Kontrollkästchen **Parameterschätzer** im Bereich **Parameterschätzer** aktiviert haben können Sie das Kontrollkästchen **Parameterschätzer** im Bereich **Ausgabe anzeigen** aktivieren.

Hinweis: Wenn Sie keine Plots auswählen und keine Prognosen oder Schätzer anzeigen lassen möchten, wird keine Ausgabe erzeugt.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

   Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine

- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

10

Box-Regelkarte

Allgemeines über die Anwendungsroutine 'Box-Regelkarte' ... 65
Regelkarten: Zuweisen von Variablen zu Analyserollen ... 66
Regelkarten: Festlegen von Regelgrenzen ... 68
Regelkarte: Auswählen der durchzuführenden Tests ... 69
Regelkarten: Festlegen von Optionen für die Plotdarstellung ... 70
Festlegen der Optionen für Box-Regelkarten ... 70
Weitere Darstellungsoptionen ... 71
Regelkarten: Speichern von Analyseergebnissen ... 72
Festlegen von Titeln und Fußnoten ... 73
Anzeigen von Eigenschaften ... 73

Allgemeines über die Anwendungsroutine 'Box-Regelkarte'

Die Shewhart-Regelkarte ist ein Grafik- und Analysetool, mit dem bestimmt werden kann, ob ein Prozess innerhalb der statistischen Kontrolle liegt. Die Anwendungsroutine **Box-Regelkarte** erstellt ein Mittelwertdiagramm für die Mittelwerte der Untergruppen. Die Box-Regelkarten jeder Untergruppe werden über das Mittelwertdiagramm überlagert. Die Box-Regelkarte bietet eine detaillierte Analyse der Untergruppenverteilungen und wird für große Untergruppen-Stichprobengrößen (mindestens 10) empfohlen.

Mithilfe einer Box-Regelkarte können Sie die Verteilung der Ausgabe untersuchen und bestimmen, ob ein Prozess innerhalb der statistischen Kontrolle liegt. Angenommen, ein Erdölunternehmen verwendet eine Turbine zum Umwandeln von Wasser in Dampf, der in den Boden gepumpt wird, um das Öl dünnflüssiger zu machen, sodass es leichter extrahiert werden kann. Dieser Prozess ist täglich 20 Mal erforderlich, und die Leistungsmenge (in Kilowatt), die zum Erhitzen des Wassers bis zur gewünschten Temperatur erforderlich ist, wird aufgezeichnet. Sie

können mithilfe einer Box-Regelkarte die Verteilung der Leistungsausgabe für jeden Tag untersuchen und bestimmen, ob die Mittelwertebene des Erhitzungsprozesses innerhalb der statistischen Kontrolle liegt.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SHEWHART
Erforderliche SAS-Produkte	Base SAS, SAS/QC, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

Regelkarten: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Prozessmaß	gibt die zu analysierenden Prozessvariablen an. Diese Variablen enthalten die Einzelmaße. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Untergruppen-Identifier	gibt die Variable an, welche die Untergruppen in den Daten identifiziert. Die Werte der Variablen geben an, wie die Beobachtungen in der Eingabetabelle in rationale Untergruppen gegliedert werden. Dies sind die üblichen Werte: ■ die Indizes, welche die Reihenfolge angeben, in der die Untergruppenstichproben gesammelt wurden

Optionsname	Beschreibung
	■ die Daten oder Zeiten, zu denen die Untergruppenstichproben gesammelt wurden ■ die Etiketten, welche die Untergruppenstichproben eindeutig identifizieren. Ein Untergruppen-Identifier ist erforderlich. Sie können dieser Rolle nur eine Variable zuweisen. Um die Daten nach den identifizierten Untergruppen zu sortieren, aktivieren Sie das Kontrollkästchen **Nach Untergruppe sortieren**.
Stichprobengröße der Untergruppe	gibt die Stichprobengröße der Untergruppe als Werte einer Variablen an. Die Stichprobengröße der Untergruppe ist erforderlich, und Sie können dieser Rolle nur eine numerische Variable zuweisen. Aktivieren Sie das Kontrollkästchen **Numerischen Wert für Untergruppe verwenden (statt Variable)**, wenn die Stichprobengröße der Untergruppe fest ist und Sie einen konstanten Wert für die Größe angeben müssen. Geben Sie den gewünschten Wert im Feld **Stichprobengröße der Untergruppe** ein. Der Wert für die Stichprobengröße der Untergruppe muss zwischen dem maximalen Wert der Prozessvariablen und 99999 (einschließlich) liegen. **Hinweis:** Diese Rolle ist nicht erforderlich, wenn Sie das Kontrollkästchen **Numerischen Wert für Untergruppe verwenden (statt Variable)** im Fenster **Optionen** aktiviert haben.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.
Blockvariablen	gibt bis zu zwei Variablen an, welche die Daten in Blöcke aufeinanderfolgender Untergruppen gruppieren. Diese Blöcke werden in einer Legende beschriftet. Jede Blockvariable liefert eine Etikettenausprägung in der Legende.

Regelkarten: Festlegen von Regelgrenzen

Klicken Sie im Auswahlfenster auf **Regelgrenzen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Sigma-Grenzen	gibt die Breite der Regelgrenzen als Mehrfaches des Standardfehlers (Sigma) der beschreibenden Untergruppenstatistik an, die im Diagramm dargestellt wird. Die Breite muss positiv sein. Der Standardwert für das Mehrfache ist 3.
Berechnungsmethode auswählen	gibt die Berechnungsmethode für die Regelgrenzen an. Folgende Optionen können ausgewählt werden: ■ **Regelgrenzen aus den aktiven Daten berechnen** ■ **Regelgrenzen aus der ausgewählten Datei berechnen** Klicken Sie auf **Durchsuchen**, um die Datei anzugeben, die die Regelgrenzen bzw. die Parameter zu deren Berechnung enthält. Wählen Sie die gewünschte Datei aus und klicken Sie auf **Öffnen**. **Hinweis:** Bei der ausgewählten Datei muss es sich um eine LIMITS= Datei handeln. Weitere Informationen über die in einer LIMITS= Datei erforderlichen Variablen finden Sie in der Prozedurendokumentation zur erstellten Regelkarte. Wenn Sie die beschreibenden Statistiken für alle Untergruppen darstellen möchten, aktivieren Sie das Kontrollkästchen **Beschreibende Statistiken für alle Untergruppen darstellen, unabhängig davon, ob die Stichprobengröße der Untergruppe der nominalen Stichprobengröße für Regelgrenzen entspricht**. Sie können diese Statistiken beispielsweise darstellen, wenn fast alle Untergruppen eine gemeinsame Stichprobengröße haben und wenn Sie die festen (anstelle der variablen) Regelgrenzen, die der nominalen Stichprobengröße *n* entsprechen, anzeigen möchten. Wenn Sie diese Option auswählen und sich Ihre Stichprobengrößen für die Untergruppen deutlich unterscheiden, ist die Interpretation der Regelgrenzen nur für die Untergruppen sinnvoll, deren Stichprobengröße *n* entspricht.

Optionsname	Beschreibung
	Um spezielle Marker für die Datenpunkte hinzuzufügen, bei denen die Stichprobengröße ungleich `n` ist, aktivieren Sie das Kontrollkästchen **Spezielle Marker für Punkte hinzufügen, die Stichprobengrößen entsprechen, welche ungleich nominalen Stichprobengrößen für feste Regelgrenzen sind**. ■ **Regelgrenzen angeben** Geben Sie in den Feldern **Oben**, **Mitte** und **Unten** die obere Regelgrenze, die Mittellinie bzw. die untere Regelgrenze an. Wenn die Stichprobengrößen Ihrer Untergruppe variieren, müssen Sie eine nominale Stichprobengröße angeben. **Hinweis:** Wenn Sie der Rolle **Analyse gruppieren nach** Variablen zuweisen, können Sie keine Werte für Regelgrenzen angeben.

Regelkarte: Auswählen der durchzuführenden Tests

Klicken Sie im Auswahlfenster auf **Tests**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Tests wenden nicht die Regelgrenzen an, die keine Drei-Sigma-Grenzen sind oder die mit der Untergruppenstichprobengröße variieren.

Optionsname	Beschreibung
Test auswählen	Fordern Sie über die Kontrollkästchen einen oder mehrere Tests auf besondere Ursachen an, die auch als Ausführungstests, Mustertests und Western-Electric-Regeln bekannt sind. Diese Tests entdecken bestimmte nichtzufällige Muster in den Punkten, die in der primären Regelkarte dargestellt werden. Das Vorkommen eines nichtzufälligen Musters wird als Signal bezeichnet. Die Tests können auf das Vorhandensein einer besonderen Variationsursache hinweisen. Wenn Sie die Tests auswählen, erscheint eine Beschreibung des jeweiligen Tests im Feld **Beschreibung**.

Optionsname	Beschreibung
Etikett	Geben Sie in den Textfeldern die Etiketten für Punkte an, bei denen die Tests auf besondere Ursachen positiv sind. Die Etiketten dürfen maximal 16 Zeichen lang sein. Jedes Etikett erscheint bei allen Punkten im Test, an denen ein Signal vorliegt.
Text zum Identifizieren der Testsignale	■ Wählen Sie über die Dropdown-Liste eine Farbe für die angegebenen Etiketten aus. ■ Markieren Sie das Kontrollkästchen **Zonenlinien anzeigen**, um Linien einzuschließen, welche die Zonen A, B und C in einem primären Diagramm kennzeichnen. Diese Zonen werden gebildet, indem die Intervalle zwischen den Regelgrenzen in sechs gleichmäßig verteilte Unterintervalle unterteilt werden. ■ Verwenden Sie das Kontrollkästchen **3 Sigma-Grenzen überschreiben**, um nach speziellen Ursachen zu suchen, wenn Sie eine andere Sigma-Regelgrenze als das Standardvielfache von 3 angeben. Um die Sigma-Grenze zu ändern, klicken Sie im Auswahlfenster auf **Regelgrenzen**. ■ Über das Kontrollkästchen **Test auf überlappende Punktmuster anwenden** wenden Sie Tests auf besondere Ursachen auf überlappende Punktmuster an.

Regelkarten: Festlegen von Optionen für die Plotdarstellung

Festlegen der Optionen für Box-Regelkarten

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Innerhalb der Regelgrenzen,	legt die Farben dieser Diagrammelemente fest.

Optionsname	Beschreibung
Regelgrenzen und Rahmen	■ **Innerhalb der Regelgrenzen** gibt die Farbe innerhalb der unteren und oberen Regelgrenzen an. ■ **Regelgrenzen** gibt die Farbe der Regelgrenzen und der Mittellinie sowie der Etiketten für diese Linien an. Diese Option gibt außerdem die Farbe der Etiketten für diese Linien an. ■ **Rahmen** gibt die Farben zum Füllen des Rechtecks an, das durch die Achsen und den Rahmen begrenzt ist. **Hinweis:** Die Optionen **Innerhalb der Regelgrenzen** und **Regelgrenzen** sind für eine Box-Regelkarte nicht verfügbar, wenn Sie die Regelgrenzen unterdrücken. Um die Regelgrenzen zu aktivieren, entfernen Sie das Häkchen im Kontrollkästchen **Anzeige der Regelgrenzen unterdrücken**.
Box-Farben	legt die Umrissfarbe und Füllfarbe der Boxen fest.
Anzeige der Regelgrenzen unterdrücken	unterdrückt die Standardlegende für Untergruppenstichproben. Diese Option ist nützlich, wenn die Stichprobengrößen der Untergruppe konstant und mit der Stichprobengröße der Regelgrenze identisch sind, da die Stichprobengröße der Regelgrenze automatisch in der oberen rechten Ecke der Karte angezeigt wird.
Boxbreite nach Unterg.größe variieren	gibt an, dass die Box-Diagramme proportional zur Größe der Untergruppe geändert werden sollen.
Boxen einkerben	erstellt eingekerbte Boxplots. Die Endpunkte der Einkerbungen befinden sich am Median, plus und minus die Menge 1,58(IQR/sqrt(*n*)), wobei IQR der Interquartilsabstand ist und sqrt(*n*) die Quadratwurzel der Untergruppenstichprobengröße. Die Mediane (zentrale Linien) von zwei Box-and-Whisker-Plots unterscheiden sich signifikant ca. beim Niveau von 0,05, wenn die entsprechenden Einkerbungen nicht überlappen.

Weitere Darstellungsoptionen

Einige der Darstellungsoptionen sind für unterschiedliche Regelkarten-Anwendungsroutinen gleich. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

■ “Festlegen von Achsenoptionen” auf Seite 759
■ “Hinzufügen von Referenzlinien” auf Seite 760

■ “Festlegen von Blockvariablenoptionen” auf Seite 762

Regelkarten: Speichern von Analyseergebnissen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

Um die Analyseergebnisse zu speichern, können Sie Ausgabedateien mit diversen Statistiken speichern. Markieren Sie die Kontrollkästchen der Ausgabedaten, die Sie permanent speichern möchten.

Optionsname	Beschreibung
Datei mit Untergruppenstatistiken und Regelgrenzen	erstellt eine Ausgabedatei, welche die im Diagramm dargestellten Informationen enthält. Dies beinhaltet die Werte der Untergruppenvariablen und ihre entsprechenden beschreibenden Statistiken und Regelgrenzen. **Hinweis:** Diese Option erstellt eine Datei OUTTABLE=. Sie können diese Datei mit der Option TABLE= aus der SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Untergruppenstatistiken	erstellt eine Ausgabedatei mit den beschreibenden Statistiken der Untergruppe. **Hinweis:** Diese Option erstellt eine Datei mit OUTHISTORY=. Sie können diese Datei mit der Option HISTORY= aus der SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Regelgrenzen	erstellt eine Ausgabedatei, welche die Regelgrenzen speichert. Sie können diese Datei verwenden, wenn Sie die Berechnungsmethode für die Regelgrenzen später in dieser Anwendungsroutine auswählen.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen

- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

11

Box-Plot

Allgemeines über die Anwendungsroutine 'Box-Plot' ... **75**

Box-Plot: Auswählen eines Diagrammtyps ... **76**

Box-Plot: Zuweisen von Variablen zu Analyserollen ... **77**

Box-Plot: Festlegen von Darstellungsoptionen ... **78**
- Festlegen der Box-Plotoptionen ... 78
- Festlegen von Optionen für Hoch-Tief-Plots ... 79
- Weitere Darstellungsoptionen ... 80

Festlegen von Titeln und Fußnoten ... **80**

Anzeigen von Eigenschaften ... **80**

Allgemeines über die Anwendungsroutine 'Box-Plot'

Die Anwendungsroutine **Box-Plot** erstellt Box-Plots, Hoch-Tief-Plots oder Hoch-Tief-Schluss-Plots, die mehrere beschreibende Statistiken für einige numerische Variablen in verschiedenen Werten einer Diagrammvariablen anzeigen. Beispielsweise kann ein Box-Plot den Bereich, den Quartilsabstand und den Mittelwert einer Variablen anzeigen. Ein Hoch-Tief-Schluss-Plot kann die höchste, die niedrigste und die Schlussnotierung einer Aktie an verschiedenen Kalenderdaten anzeigen.

Weitere Informationen, siehe “Auswählen des Grafikausgabeformats” auf Seite 6.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GPLOT

Name der Anforderung	Prozedur- und Produktnamen
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Box-Plot: Auswählen eines Diagrammtyps

Klicken Sie im Auswahlfenster auf **Box-Plot**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Box-Plot	Gibt an, dass die oberen und unteren Ecken der Box die 25ten und 75ten Perzentile der Stichprobe darstellen. Die mittlere horizontale Linie wird in der 50. Perzentile (Median) gezeichnet. Vertikale Linien, oder Whisker, werden von der Box zum extremsten Punkt innerhalb des 1,5-fachen Quartilsabstands gezeichnet. Die Werte außerhalb dieses Bereichs werden mit einem Plotsymbol markiert.
Hoch-Tief-Plot	gibt an, dass eine durchgehende vertikale Linie die niedrigen (Minimum) und hohen (Maximum) Y-Werte für jeden X-Wert verbindet. Eine Achsenmarkierung weist auf den mittleren Y-Wert für jeden X-Wert hin. Dieser Plottyp eignet sich gut für Börsendaten.
Hoch-Tief-Schluss-Plot	gibt an, dass eine durchgehende vertikale Linie die niedrigen (Minimum) und hohen (Maximum) Y-Werte für jeden X-Wert verbindet. Eine Achsenmarkierung weist auf den Schlusswert hin. Dieser Plottyp eignet sich gut für Börsendaten.

Box-Plot: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Horizontal	Die Spalte, die Sie dieser Rolle zuweisen, ist die horizontale oder X-Achsenvariable für das Diagramm.
Vertikal	Die Spalte, die Sie dieser Rolle zuweisen, ist die vertikale oder Y-Achsenvariable für das Diagramm.
Vertikal (Rechts)	Die Spalte, die Sie dieser Rolle zuweisen, ist die vertikale bzw. Y-Achsenvariable für die rechte Seite der Grafik. Diese Variable wird gegen die X-Achsenvariable dargestellt, sodass ein anderer Plot auf der Grafik erstellt wird.
Diagramme gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Diagramme, die erstellt werden. Für jeden eindeutigen Wert wird eine separate Grafik erzeugt. Aktivieren Sie das Kontrollkästchen **Gleiche Achsenskalierung für alle Diagramme**, wenn Sie möchten, dass für alle erzeugten Diagramme die gleiche Achsenskalierung verwendet wird. Standardmäßig basiert der Wertebereich für jede Achse auf den Minimal- und Maximalwerten der dargestellten Daten und könnte daher in den Diagrammen variieren. Wenn Sie diese Option auswählen haben alle Diagramme den gleichen Bereich für jede Achse.

Box-Plot: Festlegen von Darstellungsoptionen

Festlegen der Box-Plotoptionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Box-Plot**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Optionen sind nur verfügbar, wenn Sie einen Box-Plot wählen.

Optionsname	Beschreibung
Linien zwischen Medianpunkten und Enden und Spitzen der Whiskers zeichnen	Sie können Linien zwischen den Medianpunkten sowie die Enden und Spitzen jedes einzelnen Whiskers zeichnen lassen.
Farben für die Box und die Ausreißer angeben	Um Ihre eigene Kombination von Füll- und Umrissfarben wählen zu können, aktivieren Sie das Kontrollkästchen **Farben angeben** und wählen die Farben aus den Dropdown-Listen **Boxfüllung und Ausreißer** und **Umriss** aus. Die Umrissfarbe gibt die Farbe für den Box-Umriss und die Whiskerlinien in Hoch-Tief-Diagrammen und Box-Plots an.
Einbeziehen von Werten außerhalb des Achsenbereichs	Um Werte, die außerhalb des Achsenbereichs liegen, in die Interpolationsberechnungen einzubeziehen, aktivieren Sie das Kontrollkästchen **Werte außerhalb des Achsenbereichs einschließen**. Wenn diese Option ausgewählt wurde, werden Werte, die außerhalb des Achsenbereichs liegen, zwar in die Interpolationsberechnungen einbezogen, jedoch nicht im Plot berücksichtigt. Standardmäßig ist dieses Kontrollkästchen deaktiviert, sodass Werte außerhalb des Achsenbereichs bei Interpolationsberechnungen nicht berücksichtigt werden.
Die Länge der Whiskers angeben	Sie können ein Längenperzentil für die Whiskers auswählen. Diese Optionen sind verfügbar: ■ **+- 1.5 mal den Quartilsabstand** zeichnet die vertikalen Linien, oder Whisker, von der Box zum extremsten Punkt innerhalb des 1,5-fachen Quartilsabstands.

Optionsname	Beschreibung
	■ **Hohe/niedrige Extreme** zieht die Whisker aus den extrem hohen und niedrigen Werten. **Hinweis:** Wenn Sie diese Option wählen, werden alle Werte außerhalb des Bereichs des Whisker-Längenperzentils (Ausreißer) nicht mit Plotsymbolen markiert. Wenn Sie für das Längenperzentil eine andere Option verwenden, weisen die Ausreißer dagegen eine entsprechende Markierung auf. ■ **1. Perzentil niedrig, 99. hoch** zieht die Whisker aus den Extremwerten im 1. Perzentil in das 99. Perzentil. ■ **5. Perzentil niedrig, 95. hoch** zieht die Whisker aus den Extremwerten im 5. Perzentil in das 95. Perzentil. ■ **10. Perzentil niedrig, 90. hoch** zieht die Whisker aus den Extremwerten im 10. Perzentil in das 90. Perzentil. ■ **Perzentil niedrig, 75. hoch** erstellt keine Whisker, da das Feld vom 25. bis zum 75. Perzentil reicht.
Ausreißer anpassen	Sie können Sie die Schriftart, das Symbol und die Höhe angeben.

Festlegen von Optionen für Hoch-Tief-Plots

Klicken Sie im Auswahlfenster unter der Überschrift "Darstellung" auf **Hoch-Tief**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Optionen sind nur verfügbar, wenn Sie ein Hoch-Tief- oder ein Hoch-Tief-Schluss-Plot auswählen.

Optionsname	Beschreibung
Linien zwischen Medianpunkten und Enden und Spitzen der Whiskers zeichnen	Sie können Linien zwischen den Medianpunkten sowie die Enden und Spitzen jedes einzelnen Whiskers zeichnen lassen.
Minimale und maximale Werte mit leeren Balken verbinden	Sie können statt der Linien auch Balken zwischen den minimalen und maximalen Y-Werten anzeigen lassen.
Achsenmarkierung bei Schlusswert zeichnen	Sie können eine Achsenmarkierung beim Schlusswert statt beim Mittelwert anzeigen lassen. Wenn Sie diese Option verwenden möchten, müssen exakt drei Y-

Optionsname	Beschreibung
	Werte (Hoch, Niedrig und Schließen) für jeden X-Wert vorliegen. Liegen mehr oder weniger Y-Werte für jeden X-Wert vor, wird die Achsenmarkierung am Mittelwert der Y-Werte platziert.

Weitere Darstellungsoptionen

Einige Darstellungsoptionen gelten für alle Grafik-Anwendungsroutinen. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Festlegen von Achsenoptionen” auf Seite 745
- “Hinzufügen von Referenzlinien” auf Seite 748
- “Anpassen des Diagrammbereichs” auf Seite 750

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

12

Blasendiagramm

Allgemeines über die Anwendungsroutine 'Blasendiagramm' ... **83**

Blasendiagramm: Zuweisen von Variablen zu Analyserollen ... **84**

Blasendiagramm: Festlegen von Darstellungsoptionen ... **85**

Festlegen der Optionen für das Blasendiagramm ... 85

Weitere Darstellungsoptionen ... 86

Festlegen von Titeln und Fußnoten ... **86**

Anzeigen von Eigenschaften ... **86**

Allgemeines über die Anwendungsroutine 'Blasendiagramm'

Die Anwendungsroutine **Blasendiagramm** erstellt einfache Blasendiagramme. Dabei werden Kreise unterschiedlicher Größe an Datenpunkten gezeichnet, die an den vertikalen und horizontalen Achsen dargestellt werden. Zwei der Variablen bestimmen die Position der Datenpunkte. Die ditte Variable bestimmt die Größe der Kreise.

Sie können ein Blasendiagramm zum Vergleichen von Gehaltsinformationen verwenden. Angenommen, Sie erstellen ein Diagramm, in dem jede Blase eine Ingenieurskategorie darstellt, wie Luftfahrt, Chemie, Bau, Elektrik und Maschinenbau. Die Position einer Blase wird durch das durchschnittliche Gehalt der Kategorie bestimmt, die diese darstellt. Die Größe einer Blase stellt die Anzahl der Ingenieure in der Kategorie relativ zur Gesamtanzahl der Ingenieure in den Daten dar.

Weitere Informationen, siehe “Auswählen des Grafikausgabeformats” auf Seite 6.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GPLOT
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Blasendiagramm: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Horizontal	Die Spalte, die Sie dieser Rolle zuweisen, ist die horizontale oder X-Achsenvariable für das Diagramm.
Vertikal	Die Spalte, die Sie dieser Rolle zuweisen, ist die vertikale oder Y-Achsenvariable für das Diagramm.
Vertikal (Rechts)	Die Spalte, die Sie dieser Rolle zuweisen, ist die vertikale bzw. Y-Achsenvariable für die rechte Seite der Grafik. Diese Variable wird gegen die X-Achsenvariable dargestellt, sodass ein anderer Plot auf der Grafik erstellt wird.
Blasengröße	Die Spalte, die Sie dieser Rolle zuweisen, bestimmt die relativen Größen der Blasen auf dem Diagramm der horizontalen und vertikalen Variablen. **Hinweis:** Wenn Sie der Rolle **Vertikal rechts** eine Variable zugewiesen haben, der Rolle **Bubble2-Größe** jedoch nicht, bestimmt diese Rolle auch die relativen Größen der Blasen auf dem Diagramm der horizontalen und vertikalen Variablen.

Optionsname	Beschreibung
Diagramme gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Diagramme, die erstellt werden. Für jeden eindeutigen Wert wird eine separate Grafik erzeugt. Aktivieren Sie das Kontrollkästchen **Gleiche Achsenskalierung für alle Diagramme**, wenn Sie möchten, dass für alle erzeugten Diagramme die gleiche Achsenskalierung verwendet wird. Standardmäßig basiert der Wertebereich für jede Achse auf den Minimal- und Maximalwerten der dargestellten Daten und könnte daher in den Diagrammen variieren. Wenn Sie diese Option auswählen haben alle Diagramme den gleichen Bereich für jede Achse.
Bubble2-Größe	Die Spalte, die Sie dieser Rolle zuweisen, bestimmt die relativen Größen der Blasen auf dem Plot der horizontalen und vertikalen rechten Variablen. **Hinweis:** Diese Option ist nur verfügbar, wenn Sie der Rolle **Vertikal rechts** eine Variable zuweisen.

Blasendiagramm: Festlegen von Darstellungsoptionen

Festlegen der Optionen für das Blasendiagramm

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Blasenoptionen**, um auf diese Optionen zuzugreifen.

Wählen Sie eine Blasenfarbe aus der Dropdown-Liste aus.

Sie können die Blasen beschriften. Das Etikett ist der Wert der Spalte, die der Rolle **Blasengröße** für die Kombination aus horizontalen und vertikalen Variablenwerten zugewiesen ist. Wenn diese Variable ein Format besitzt, wird der formatierte Wert verwendet. Standardmäßig sind Blasen nicht beschriftet.

Skalieren Sie über den Schieberegler alle Blasen auf eine größere oder kleinere Größe. Sie können die Ausgabe vor dem Einstellen dieses Werts ansehen.

Hinweis: Wenn Sie der Rolle **Vertikal rechts** eine Variable zugewiesen haben, können Sie auch die Bubble2-Optionen für die Blasen festlegen, die in dieser Kombination von horizontalen und vertikal rechts Variablen enthalten sind.

Weitere Darstellungsoptionen

Einige Darstellungsoptionen gelten für alle Grafik-Anwendungsroutinen. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Festlegen von Achsenoptionen” auf Seite 745
- “Hinzufügen von Referenzlinien” auf Seite 748
- “Anpassen des Diagrammbereichs” auf Seite 750

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine

- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

13

Kanonische Korrelation

Allgemeines über die Anwendungsroutine 'Kanonische Korrelation' ... 89
Kanonische Korrelation: Zuweisen von Variablen zu Analyserollen ... 90
Kanonische Korrelation: Festlegen von Statistikoptionen ... 91
Kanonische Korrelation: Erzeugen von Diagrammen ... 93
Kanonische Korrelation: Festlegen von Ergebnisoptionen ... 93
Festlegen von Titeln und Fußnoten ... 94
Anzeigen von Eigenschaften ... 95

Allgemeines über die Anwendungsroutine 'Kanonische Korrelation'

Bei der **kanonischen Korrelationsanalyse** untersuchen Sie die Beziehung zwischen den linearen Kombinationen eines Satzes mit X-Variablen und linearen Kombinationen eines Satzes mit Y-Variablen. Eine Reihe von Hypothesen prüfen, ob jede kanonische Korrelation und alle kleineren kanonischen Korrelationen in der Grundgesamtheit null sind. Die einfache und multiple Korrelation sind Sonderfälle der kanonischen Korrelation, bei denen eine oder beide Sätze eine einzelne Variable enthalten.

Sie können die Anwendungsroutine **Kanonische Korrelation** verwenden, um den Übereinstimmungsgrad zwischen einer Menge von Tätigkeitsmerkmalen und einer Menge an Messungen der Mitarbeiterzufriedenheit zu bestimmen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	CANCORR
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Kanonische Korrelation: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Satz-1-Variablen	führt Analysen auf Basis der Variablen durch, die Sie dieser Rolle zuweisen. Sie müssen dieser Rolle mindestens eine Variable zuweisen. Verwenden Sie den Bereich **Kanonische Variate aus Satz 1**, um ein Präfix und Etikett für die Satz-1-Variablen anzugeben.
Satz-2-Variablen	führt Analysen auf Basis der Variablen durch, die Sie dieser Rolle zuweisen. Sie müssen dieser Rolle mindestens eine Variable zuweisen. Verwenden Sie den Bereich **Kanonische Variate aus Satz 2**, um ein Präfix und Etikett für die Satz-2-Variablen anzugeben.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert.

Optionsname	Beschreibung
	Hinweis: Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits als **Satz-1-Variablen** bzw. **Satz-2-Variablen** zugewiesen haben.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Partielle Variablen	legt die Variablen fest, damit die kanonische Analyse auf Basis der partiellen Korrelationen erfolgt. Die Variablen in dieser Rolle stammen zum Teil aus den Satz-1- und zum Teil aus den Satz-2-Variablen.
Relative Gewichtung	gibt den Wert der Variable in jeder Zeile an, um die gewichteten Produkt-Moment-Korrelationskoeffizienten zu berechnen. Eine Zeile wird nur verwendet, wenn der Wert der Variablen in dieser Rolle größer als null ist. Sie können dieser Rolle maximal eine Variable zuweisen.

Kanonische Korrelation: Festlegen von Statistikoptionen

Klicken Sie im Auswahlfenster auf **Statistiken**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Durchzuführende Regressionsanalysen	legt die durchzuführende Regressionsanalyse fest. Folgende Optionen können ausgewählt werden: ■ **Keine Regressionsanalysen** führt keine Regressionsanalysen durch. ■ **Satz 1 sagt Satz 2 vorher, Satz 2 sagt Satz 1 vorher** fordert zwei multiple Regressionsanalysen an, sodass **Satz-1-Variablen** die abhängigen Variablen für die eine Analyse und die Regressoren für die andere Analyse sind.

Optionsname	Beschreibung
	■ **Satz 1 sagt Satz 2 vorher** fordert multiple Regressionsanalysen mit Variablen aus **Satz-2-Variablen** als abhängige Variablen und den Variablen aus **Satz-1-Variablen** als Regressoren an. ■ **Satz 2 sagt Satz 1 vorher** fordert multiple Regressionsanalysen mit Variablen aus **Satz-1-Variablen** als abhängige Variablen und den Variablen aus **Satz-2-Variablen** als Regressoren an.
Regressionsstatistiken	berechnet eine oder mehrere folgender Statistiken: ■ **Regressionskoeffizienten** erstellt die rohen Regressionskoeffizienten aus den Regressionsanalysen. ■ **Standardisierte Regressionskoeffizienten** erstellt standardisierte Regressionskoeffizienten. ■ **Standardfehler der Koeffizienten** erstellt Standardfehler der Regressionskoeffizienten. ■ **t-Statistik und Prob > |t|** erstellt *t*-Statistiken für die Regressionskoeffizienten und Wahrscheinlichkeitsniveaus für die *t*-Statistiken. ■ **Quadrierte multiple Korrelation** erstellt quadrierte multiple Korrelationen und *F*-Tests für die Regressionsanalysen.
Korrelationsstatistik	berechnet eine oder mehrere folgender Statistiken: ■ **Korrelationen der Regressionskoeffizienten** erstellt Korrelationen in den Regressionskoeffizientenschätzern. ■ **Partielle Korrelationen** erstellt partielle Korrelationen zwischen Regressoren und abhängigen Variablen, wobei aus allen abhängigen Variablen und Regressoren die Effekte aller anderen Regressoren entfernt werden. ■ **Quadrierte partielle Korrelationen** erstellt quadrierte partielle Korrelationen zwischen Regressoren und abhängigen Variablen, wobei aus allen abhängigen Variablen und Regressoren die Effekte aller anderen Regressoren entfernt werden. ■ **Semipartielle Korrelationen** erstellt semipartielle Korrelationen zwischen Regressoren und abhängigen Variablen, wobei die Effekte aller anderen Regressoren aus den einzelnen Regressoren entfernt werden. ■ **Quadrierte semipartielle Korrelationen** erstellt quadrierte semipartielle Korrelationen zwischen Regressoren und abhängigen Variablen, wobei die Effekte aller anderen Regressoren aus den einzelnen Regressoren entfernt werden.

Kanonische Korrelation: Erzeugen von Diagrammen

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Hinweis: Die Plotoptionen sind nicht verfügbar, wenn die Anzahl der kanonischen Variablen, für die Koeffizienten und Redundanzstatistiken gezeigt werden, gleich null ist. Weitere Informationen, siehe die Option Anzahl kanonischer Variablen in Ergebnissen.

Aktivieren Sie das Kontrollkästchen **Plots der kanonischen Variablen anzeigen**, um Plots anzuzeigen. Geben Sie in den Feldern **Startvariable** und **Endvariable** an, welche Variablen als Plot dargestellt werden sollen. Hier einige Beispiele:

- Um nur die erste kanonische Variable (V1 versus W1) darzustellen, geben Sie `1` in beide Eingabefelder ein.
- Um die erste, zweite und dritte kanonische Variable darzustellen (V1 versus W1, V2 versus W2 und V3 versus W3), geben Sie `1` in das erste Eingabefeld und `3` in das zweite Eingabefeld ein.

Kanonische Korrelation: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Ergebnisse in einer Ausgabedatei speichern	speichert das Ergebnisse in einer Ausgabedatei. Folgende Optionen können ausgewählt werden: ■ **Korrelationen speichern** erstellt eine Datei, die verschiedene Statistiken enthält, u. a. die kanonischen Korrelationen und Koeffizienten sowie die von Ihnen angeforderten Statistiken mit multiplen Regressionen. ■ **Scorewerte speichern** erstellt eine Datei, die alle Originaldaten und Scorewerte für die kanonischen Variablen enthält. Die Anzahl der neuen Variablen entspricht dem doppelten Wert, der im Feld **Anzahl der kanonischen Variablen** angegeben wurde.

Optionsname	Beschreibung
	Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Ergebnisse anzeigen	zeigt die Ausgabe an. Folgende Optionen können angegeben werden: ■ **Anzahl der kanonischen Variablen** gibt an, wie viele kanonische Variablen die volle Ausgabe in den Ergebnissen enthalten. ■ **Kanonische Redundanzanalyse einschließen** führt eine kanonische Redundanzanalyse, einschließlich roher (nicht standardisierter) und standardisierter Varianz sowie kumulativer Proportion der Varianz der einzelnen Variablenmengen durch. Diese werden durch ihre eigenen kanonischen Variablen sowie durch die entgegengesetzten kanonischen Variablen erläutert. ■ **Scorewerte einschließen** zeigt die Ausgabedaten an, die alle Originaldaten und die Scorewerte für die kanonischen Variablen enthalten. Wenn Sie eine permanente Datei zur späteren Verwendung erstellen möchten, wählen Sie das Kontrollkästchen **Scorewerte speichern** aus.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

14

c-Regelkarte

Allgemeines über die Anwendungsroutine 'c-Regelkarte' 97
c-Regelkarte: Zuweisen von Variablen zu Analyserollen 98
Regelkarten: Festlegen von Regelgrenzen 99
Regelkarte: Auswählen der durchzuführenden Tests 101
Regelkarten: Festlegen von Optionen für die Plotdarstellung 102
Regelkarten: Speichern von Analyseergebnissen 102
Festlegen von Titeln und Fußnoten 103
Anzeigen von Eigenschaften 104

Allgemeines über die Anwendungsroutine 'c-Regelkarte'

Die Shewhart-Regelkarte ist ein Grafik- und Analysetool, mit dem bestimmt werden kann, ob ein Prozess innerhalb der statistischen Kontrolle liegt. Die Anwendungsroutine **c-Regelkarte** erstellt c-Regelkarten für die Anzahl der Nichtkonformitäten (Defekte) in den Untergruppenstichproben.

Sie können eine c-Regelkarte zum Überwachen der Anzahl der Defekte verwenden, die bei einem neuen Produkt auftreten. Angenommen, ein Automobilunternehmen möchte die Anzahl der Lackfehler bei seinen neuen Lastkraftwagen überwachen. Es werden 20 Lkw desselben Modells untersucht, und die Anzahl der Lackfehler pro Lkw wird aufgezeichnet. Jeder Punkt in der c-Regelkarte stellt die Anzahl der Fehler für einen bestimmten Lkw dar. Die oberen und unteren Regelgrenzen werden markiert, sodass Sie sehen können, ob eventuell Punkte über oder unter diesen Grenzen liegen. Überschreitet keiner der Punkte die angegebenen Regelgrenzen, befindet sich der Lackierungsprozess innerhalb der statistischen Kontrolle.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SHEWHART
Erforderliche SAS-Produkte	Base SAS, SAS/QC, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

c-Regelkarte: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Anzahl der Nichtkonformitäten	gibt die Variablen an, welche die Anzahl der Nichtkonformitäten enthalten. Sie müssen dieser Rolle mindestens eine Variable zuweisen
Untergruppen-Identifier	gibt die Variable an, welche die Untergruppen in den Daten identifiziert. Die Werte der Variablen geben an, wie die Beobachtungen in der Eingabetabelle in rationale Untergruppen gegliedert werden. Dies sind die üblichen Werte: ■ die Indizes, welche die Reihenfolge angeben, in der die Untergruppenstichproben gesammelt wurden ■ die Daten oder Zeiten, zu denen die Untergruppenstichproben gesammelt wurden ■ die Etiketten, welche die Untergruppenstichproben eindeutig identifizieren. Ein Untergruppen-Identifier ist erforderlich. Sie können dieser Rolle nur eine Variable zuweisen.

Optionsname	Beschreibung
	Um die Daten nach den identifizierten Untergruppen zu sortieren, aktivieren Sie das Kontrollkästchen **Nach Untergruppe sortieren**.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.
Blockvariablen	gibt bis zu zwei Variablen an, welche die Daten in Blöcke aufeinanderfolgender Untergruppen gruppieren. Diese Blöcke werden in einer Legende beschriftet. Jede Blockvariable liefert eine Etikettenausprägung in der Legende.

Regelkarten: Festlegen von Regelgrenzen

Klicken Sie im Auswahlfenster auf **Regelgrenzen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Sigma-Grenzen	gibt die Breite der Regelgrenzen als Mehrfaches des Standardfehlers (Sigma) der beschreibenden Untergruppenstatistik an, die im Diagramm dargestellt wird. Die Breite muss positiv sein. Der Standardwert für das Mehrfache ist 3.
Berechnungsmethode auswählen	gibt die Berechnungsmethode für die Regelgrenzen an. Folgende Optionen können ausgewählt werden: ■ **Regelgrenzen aus den aktiven Daten berechnen** ■ **Regelgrenzen aus der ausgewählten Datei berechnen** Klicken Sie auf **Durchsuchen**, um die Datei anzugeben, die die Regelgrenzen bzw. die Parameter zu deren

Optionsname	Beschreibung
	Berechnung enthält. Wählen Sie die gewünschte Datei aus und klicken Sie auf **Öffnen**. **Hinweis:** Bei der ausgewählten Datei muss es sich um eine LIMITS= Datei handeln. Weitere Informationen über die in einer LIMITS= Datei erforderlichen Variablen finden Sie in der Prozedurendokumentation zur erstellten Regelkarte. Wenn Sie die beschreibenden Statistiken für alle Untergruppen darstellen möchten, aktivieren Sie das Kontrollkästchen **Beschreibende Statistiken für alle Untergruppen darstellen, unabhängig davon, ob die Stichprobengröße der Untergruppe der nominalen Stichprobengröße für Regelgrenzen entspricht**. Sie können diese Statistiken beispielsweise darstellen, wenn fast alle Untergruppen eine gemeinsame Stichprobengröße haben und wenn Sie die festen (anstelle der variablen) Regelgrenzen, die der nominalen Stichprobengröße *n* entsprechen, anzeigen möchten. Wenn Sie diese Option auswählen und sich Ihre Stichprobengrößen für die Untergruppen deutlich unterscheiden, ist die Interpretation der Regelgrenzen nur für die Untergruppen sinnvoll, deren Stichprobengröße *n* entspricht. Um spezielle Marker für die Datenpunkte hinzuzufügen, bei denen die Stichprobengröße ungleich `n` ist, aktivieren Sie das Kontrollkästchen **Spezielle Marker für Punkte hinzufügen, die Stichprobengrößen entsprechen, welche ungleich nominalen Stichprobengrößen für feste Regelgrenzen sind**. ■ **Regelgrenzen angeben** Geben Sie in den Feldern **Oben**, **Mitte** und **Unten** die obere Regelgrenze, die Mittellinie bzw. die untere Regelgrenze an. Wenn die Stichprobengrößen Ihrer Untergruppe variieren, müssen Sie eine nominale Stichprobengröße angeben. **Hinweis:** Wenn Sie der Rolle **Analyse gruppieren nach** Variablen zuweisen, können Sie keine Werte für Regelgrenzen angeben.

Regelkarte: Auswählen der durchzuführenden Tests

Klicken Sie im Auswahlfenster auf **Tests**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Tests wenden nicht die Regelgrenzen an, die keine Drei-Sigma-Grenzen sind oder die mit der Untergruppenstichprobengröße variieren.

Optionsname	Beschreibung
Test auswählen	Fordern Sie über die Kontrollkästchen einen oder mehrere Tests auf besondere Ursachen an, die auch als Ausführungstests, Mustertests und Western-Electric-Regeln bekannt sind. Diese Tests entdecken bestimmte nichtzufällige Muster in den Punkten, die in der primären Regelkarte dargestellt werden. Das Vorkommen eines nichtzufälligen Musters wird als Signal bezeichnet. Die Tests können auf das Vorhandensein einer besonderen Variationsursache hinweisen. Wenn Sie die Tests auswählen, erscheint eine Beschreibung des jeweiligen Tests im Feld **Beschreibung**.
Etikett	Geben Sie in den Textfeldern die Etiketten für Punkte an, bei denen die Tests auf besondere Ursachen positiv sind. Die Etiketten dürfen maximal 16 Zeichen lang sein. Jedes Etikett erscheint bei allen Punkten im Test, an denen ein Signal vorliegt.
Text zum Identifizieren der Testsignale	■ Wählen Sie über die Dropdown-Liste eine Farbe für die angegebenen Etiketten aus. ■ Markieren Sie das Kontrollkästchen **Zonenlinien anzeigen**, um Linien einzuschließen, welche die Zonen A, B und C in einem primären Diagramm kennzeichnen. Diese Zonen werden gebildet, indem die Intervalle zwischen den Regelgrenzen in sechs gleichmäßig verteilte Unterintervalle unterteilt werden. ■ Verwenden Sie das Kontrollkästchen **3 Sigma-Grenzen überschreiben**, um nach speziellen Ursachen zu suchen, wenn Sie eine andere Sigma-Regelgrenze als das Standardvielfache von 3 angeben. Um die Sigma-Grenze zu ändern, klicken Sie im Auswahlfenster auf **Regelgrenzen**.

Optionsname	Beschreibung
	■ Über das Kontrollkästchen **Test auf überlappende Punktmuster anwenden** wenden Sie Tests auf besondere Ursachen auf überlappende Punktmuster an.

Regelkarten: Festlegen von Optionen für die Plotdarstellung

Einige der Darstellungsoptionen sind für unterschiedliche Regelkarten-Anwendungsroutinen gleich. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Festlegen von Achsenoptionen” auf Seite 759
- “Hinzufügen von Referenzlinien” auf Seite 760
- “Festlegen weiterer Plotoptionen” auf Seite 760
- “Festlegen von Blockvariablenoptionen” auf Seite 762

Regelkarten: Speichern von Analyseergebnissen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

Um die Analyseergebnisse zu speichern, können Sie Ausgabedateien mit diversen Statistiken speichern. Markieren Sie die Kontrollkästchen der Ausgabedaten, die Sie permanent speichern möchten.

Optionsname	Beschreibung
Datei mit Untergruppenstatistiken und Regelgrenzen	erstellt eine Ausgabedatei, welche die im Diagramm dargestellten Informationen enthält. Dies beinhaltet die Werte der Untergruppenvariablen und ihre entsprechenden beschreibenden Statistiken und Regelgrenzen. **Hinweis:** Diese Option erstellt eine Datei OUTTABLE=. Sie können diese Datei mit der Option TABLE= aus der

Optionsname	Beschreibung
	SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Untergruppenstatistiken	erstellt eine Ausgabedatei mit den beschreibenden Statistiken der Untergruppe. **Hinweis:** Diese Option erstellt eine Datei mit OUTHISTORY=. Sie können diese Datei mit der Option HISTORY= aus der SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Regelgrenzen	erstellt eine Ausgabedatei, welche die Regelgrenzen speichert. Sie können diese Datei verwenden, wenn Sie die Berechnungsmethode für die Regelgrenzen später in dieser Anwendungsroutine auswählen.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

   Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

15

CDF-Plots

Allgemeines über die Anwendungsroutine 'CDF-Plots' ... **105**

Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen ... **106**

CDF-Plots: Auswählen der Verteilung ... **107**

CDF-Plots: Anpassen der Verteilung ... **108**

Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung ... **109**

Festlegen von Achsenoptionen ... 109
Hinzufügen von Referenzlinien ... 109
Einfügungen hinzufügen ... 110
Festlegen weiterer Plotoptionen ... 110
Festlegen von Optionen für Spezifikationsgrenzen ... 112

Prozessfähigkeitsanalyse: Festlegen der Analysetypen ... **113**

Festlegen von Titeln und Fußnoten ... **114**

Anzeigen von Eigenschaften ... **114**

Allgemeines über die Anwendungsroutine 'CDF-Plots'

Eine Prozessfähigkeitsanalyse vergleicht die Verteilung der Ausgabe eines Prozesses in statistischer Kontrolle mit den Spezifikationsgrenzen, um die Konsistenz zu bestimmen, mit der die Spezifikationen erfüllt werden. Ein CDF-Plot dient bei der Prozessfähigkeitsanalyse zur Darstellung der beobachteten kumulativen Verteilungsfunktion (Cumulative Distribution Function, CDF) einer Variablen. Die kumulative Verteilungsfunktion wird auch als empirische kumulative Verteilungsfunktion (Empirical Cumulative Distribution Function, ECDF) bezeichnet und ist eine Funktion mit ansteigenden Stufen.

Sie können mit einem CDF-Plot den Prozentwert der Beobachtungen verwenden, die unterhalb der unteren Spezifikationsgrenze oder über der oberen Spezifikationsgrenze liegen. Angenommen, ein Unternehmen, das Glasfaserkabel

herstellt, ist an der Bruchfestigkeit der Kabel interessiert. Sie erstellen eine Datei mit 50 Bruchfestigkeiten, gemessen in kg/mm². Die Anwendungsroutine **CDF-Plot** erstellt anhand dieser Daten einen CDF-Plot, der den Prozentwert der Beobachtungen in der Stichprobe anzeigt, die unterhalb der unteren Spezifikationsgrenze liegen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	CAPABILITY
Erforderliche SAS-Produkte	Base SAS, SAS/QC, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Analysevariablen	gibt die zu analysierenden numerischen Variablen an. Sie müssen dieser Rolle mindestens eine Variable zuweisen. Bei allen Plots, außer dem P-P-Plot, können Sie eine Obergrenze, einen Zielwert und eine Untergrenze im Bereich **Spezifikationsgrenzen für *Variablenname*** für jede Analysevariable angeben. Wenn Sie diese Grenzen angeben, enthält die Ausgabe automatisch eine Tabelle mit den Tests auf Normalverteilung, eine Tabelle mit den Spezifikationsgrenzen und eine Tabelle mit den Prozessfähigkeitsindizes.

Optionsname	Beschreibung
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt die Variable an, die als relative Gewichtungsvariable verwendet werden soll. Wenn Sie dieser Rolle eine Variable zuweisen, werden mit dem Wert der Variablen in jeder Beobachtung gewichtete Statistiken berechnet. Sie können dieser Rolle nur eine Variable zuweisen. **Hinweis:** Sie sollten der Rolle **Relative Gewichtung** keine Variable zuweisen, wenn Sie ein Histogramm, Wahrscheinlichkeitsplot, P-P-Plot, Q-Q-Plot oder CDF-Plot erstellen. Falls Sie dieser Rolle ein Variable zugewiesen haben, müssen Sie eine eventuelle Ausgabe mit Vorsicht interpretieren.
Klassifizierungsvariablen (nur verfügbar für die Anwendungsroutine Histogramm)	gibt eine oder zwei Variablen an, die als Klassifizierungsvariablen dienen. Wenn Sie dieser Rolle eine Variable zuweisen, erstellt die Anwendungsroutine ein Vergleichshistogramm, mit dem Sie die Verteilung einer Analysevariablen über Ausprägungen der Klassifizierungsvariablen hinweg vergleichen können. **Hinweis:** Um vergleichende Wahrscheinlichkeitsplots oder vergleichende Q-Q-Plots zu erstellen, verwenden Sie die Anwendungsroutine Verteilungsanalyse.

CDF-Plots: Auswählen der Verteilung

Klicken Sie im Auswahlfenster unter der Überschrift 'Verteilungen' auf **Zusammenfassung**, um auf diese Optionen zuzugreifen.

Sie können genau eine theoretische Verteilung auf einem CDF-Plot anpassen. Wenn Sie die Verteilung ausgewählt haben, wählen Sie den Namen der Verteilung im

Auswahlfenster aus, um die Parameterwerte und Linieneigenschaften für diese Verteilung anzugeben.

Hinweis: Wenn Sie mehrere Verteilungen auswählen, wird ein extra CDF-Plot für jede Verteilung erstellt.

Optionsname	Beschreibung
Verteilungen	Verwenden Sie die Kontrollkästchen, um eine oder mehrere der folgenden geschätzten Dichtekurven auszuwählen: ■ **Normal** passt eine normale Dichte mithilfe eines Mittelwerts (My) und einer Standardabweichung (Sigma) an. ■ **Lognormal** passt eine Lognormal-Dichte mithilfe eines Skalierungsparameters (Zeta), eines Schwellenwertparameters (Theta) und eines Formparameters (Sigma) an. ■ **Exponentiell** passt eine exponentielle Dichte mithilfe eines Skalierungsparameters (Sigma) und eines Schwellenwertparameters (Theta) an. ■ **Weibull** passt eine Weibull-Dichte mit drei Parametern mit einem Skalierungsparameter (sigma), einem Schwellenwertparameter (theta) und einem Formparameter (c) an. ■ **Beta** passt eine Beta-Dichte mit einem Skalierungsparameter (Sigma), einem Schwellenwertparameter (Theta) und Formparametern (Alpha und Beta) an. ■ **Gamma** passt eine Gamma-Dichte mit einem Skalierungsparameter (Sigma), einen Schwellenwertparameter (Theta) und einen Formparameter (Alpha) an.
Grafikstil	Wählen Sie den Grafikstil, der in der Ausgabe verwendet werden soll. Sie haben die Auswahl zwischen üblichen SAS-Grafiken und ODS Statistical Graphs (statistischen ODS-Grafiken). Weitere Informationen zu ODS-Grafiken finden Sie unter *SAS Output Delivery System: User's Guide*.

CDF-Plots: Anpassen der Verteilung

Klicken Sie im Auswahlfenster unter der Überschrift **Verteilungen** auf den Namen der Verteilung, die Sie anpassen möchten, um auf diese Optionen zuzugreifen.

Hinweis: Diese Verteilungsoptionen sind nur durch vorherige Auswahl der Verteilung auf der Seite **Zusammenfassung** verfügbar, oder durch Aktivieren des Kontrollkästchens ***Name der Verteilung*** auf der Seite dieser Verteilung.

Sie können die Verteilungsparameter für jede einzelne Analysevariable oder für alle Analysevariablen angeben. Standardmäßig schätzt die Anwendungsroutine die Parameterwerte für alle Analysevariablen.

So legen Sie für jede Analysevariable Parameter fest:

1 Deaktivieren Sie das Kontrollkästchen **Verteilung auf alle Variablen anwenden**.

2 Aktivieren Sie das Kontrollkästchen der Variablen im Feld **Analysevariablen**.

3 Geben Sie die Parameterwerte an.

Hinweis: Um zu den geschätzten Parameterwerten zurückzukehren, klicken Sie auf **Schätzer verwenden**.

Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung

Festlegen von Achsenoptionen

Klicken Sie im Auswahlfenster unter **Darstellung** ⇨ **Achsen** auf **Achsen**, um auf die Optionen zuzugreifen.

Folgende Optionen können angegeben werden:

- die Farbe der Achsen
- die Farbe der Achsenmarkierungen
- ein Etikett für die horizontalen und vertikalen Achsen

Hinzufügen von Referenzlinien

Klicken Sie im Auswahlfenster unter **Darstellung** ⇨ **Achsen** auf **Horizontal** bzw. **Vertikal**, um auf die Optionen zuzugreifen.

Um Referenzlinien zu verwenden, wählen Sie **Referenzlinien verwenden**. Sie können die Referenzlinie formatieren und die Position jeder Referenzlinie festlegen.

Einfügungen hinzufügen

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Einfügung**, um auf diese Optionen zuzugreifen.

Sie können einen Plot erweitern, indem Sie direkt in der Grafik ein Feld oder eine Tabelle (eine so genannte Einfügung) der beschreibenden Statistiken hinzufügen. Wenn Sie ein Vergleichshistogramm erstellen, erscheint in jedem Komponentenhistogramm eine Einfügung.

Wählen Sie **Einfügung einschließen**, um eine beschreibende Statistik zum Plot hinzuzufügen. Wählen Sie dann die Statistiken, die Sie in die Einfügung einschließen möchten.

Hinweis: Die Statistiken für die Berechnung der Prozessfähigkeitsindizes sind nur verfügbar, wenn Sie Spezifikationsgrenzen festgelegt und dann der Rolle **Analysevariablen** zugewiesen haben. Sie können Spezifikationsgrenzen für Histogramm, CDF-Plots, Wahrscheinlichkeitsplots und Q-Q-Plots erstellen. Weitere Informationen zur Berechnung der Prozessfähigkeitsindizes finden Sie in der Hilfe zur Prozedur CAPABILITY.

Festlegen weiterer Plotoptionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Optionen**, um auf diese Optionen zuzugreifen.

Die verfügbaren Plotdarstellungsoptionen hängen vom Typ des Plots ab, den die Anwendungsroutine erzeugt.

Optionsname	Beschreibung
Hintergrund	bestimmt die Hintergrundfarbe des Plots.
Balkenumriss und **Balkenfüllung**	legt die Farben der Balken in einem Histogramm fest. **Hinweis:** Diese Optionen sind nur für Vergleichshistogramme verfügbar.
Beide Achsen umkehren (Perzentile auf entgegengesetzte Achsen)	kehrt die Achsen auf dem Plot um. Wenn Sie diese Option auswählen, funktionieren die von Ihnen ausgewählten Achsenoptionen und Referenzlinien eventuell nicht mehr, da sich die Ausrichtung des Plots ändert. **Hinweis:** Diese Option ist nur für Wahrscheinlichkeitsplots und Q-Q-Plots verfügbar.

Optionsname	Beschreibung
Keine Standardlegenden	verhindert die Anzeige der Standardlegende mit dem Plot.
Midpoints verwenden	geben Sie Midpoints für das Histogramm an. Die in der Dropdown-Liste verfügbaren Optionen hängen davon ab, ob Sie ein Histogramm oder ein Vergleichshistogramm erstellen. Wenn Sie der Rolle **Klassifizierungsvariablen** eine Variable zuweisen, erstellen Sie ein Vergleichshistogramm. Für Vergleichshistogramme stehen folgende Optionen zur Verfügung: ■ **Midpoints angeben** - Ermöglicht das Angeben der Start- und End-Midpoints. Sie können auch das Schrittintervall festlegen, das für die Berechnung der anderen Midpoints verwendet werden soll. Der Midpoint ist der Mittelwert des Wertebereichs, der von den einzelnen Balken dargestellt wird. ■ **Schlüssel** - Bestimmt die Midpoints für die Daten in der Schlüsselzelle. Standardmäßig werden die Ausprägungen der Klassifizierungsvariablen von oben nach unten (links nach rechts) in aufsteigender Reihenfolge der internen (nicht formatierten) Werte der ersten Klassifizierungsvariablen angezeigt. Wenn Sie nur eine Klassifizierungsvariable angegeben haben, ist die Schlüsselzelle die Ausprägung, die als erstes in dieser Reihenfolge erscheint. Wenn Sie zwei Klassifizierungsvariablen angegeben haben, ist die Schlüsselzelle die Kombination der Ausprägungen von Variable 1 und Variable 2, die als erstes in dieser Reihenfolge erscheint. Die Auswahl der Schlüsselzelle bestimmt die einheitliche Horizontalachse, die für alle Zellen verwendet wird. Die Midpoint-Liste für die Schlüsselzelle wird dann wie erforderlich in beide Richtungen erweitert, bis sie die Daten in den übrigen Zellen abdeckt. ■ **Einheitlich** - Bestimmt die Anzahl der Midpoints abhängig von der Gesamt-Stichprobengröße. Wenn Sie der Rolle **Klassifizierungsvariablen** keine Variable zuweisen, erstellen Sie ein Histogramm. Für Histogramme sind die folgenden Optionen verfügbar: ■ **Midpoints angeben** - Ermöglicht das Angeben der Start- und End-Midpoints. Sie können auch das Schrittintervall festlegen, das für die Berechnung der anderen Midpoints verwendet werden soll. Der Midpoint ist der Mittelwert des Wertebereichs, der von den einzelnen Balken dargestellt wird. ■ **Anzahl der Intervalle angeben** - Ermöglicht, die Anzahl der Intervalle (auch Histogrammintervalle genannt) für die

Optionsname	Beschreibung
	Daten anzugeben. Standardmäßig basiert die Anzahl der angezeigten Intervalle auf dem Datenbereich. Sie können jedoch eine Maximalzahl anzuzeigender Intervalle angeben. **Hinweis:** Sie können nur dann Midpoints angeben, wenn Sie die Anwendungsroutine **Histogramme** verwenden.
Größe ben.def. Diagramm angeben (in Pixel)	legt die Bildgröße fest. Zum Vergleich wird die Gesamtbildschirmgröße angezeigt in den Feldern **Breite** und **Höhe** angezeigt.

Hinweis: Vergleichsoptionen sind für Vergleichshistogramme verfügbar. Sie können ein Vergleichshistogramm erstellen, indem Sie in der Anwendungsroutine **Histogramme** der Rolle **Klassifizierungsvariablen** eine Variable zuweisen.

Optionsname	Beschreibung
Zeilen und **Spalten**	Wenn Sie einer Analyserolle mindestens eine Klassifizierungsvariable zuweisen, können Sie über die Felder **Zeilen pro Seite** und **Spalten pro Seite** die Anordnung der Spalten und Zeilen im Vergleichshistogramm ändern. Wenn Sie eine Klassifizierungsvariable zuweisen, ist die Standardanordnung zwei Zeilen und eine Spalte pro Seite. Wenn Sie zwei Klassifizierungsvariablen zuweisen, ist die Standardanordnung zwei Spalten und zwei Zeilen pro Seite. Sie können auch über die Dropdown-Listen **Rahmenseite** und **Rahmenoberseite** Farben zum Füllen des Rahmenbereichs der Zeilenetiketten und Spaltenetiketten angeben.
Anzahl der Intervalle	Standardmäßig basiert die Anzahl der angezeigten Intervalle auf dem Datenbereich. Sie können jedoch eine Maximalzahl anzuzeigender Intervalle angeben. Sie können auch eine Standardabweichung angeben, so dass die Anzahl der angezeigten Intervalle von einem Bereich des Standardabweichungswerts über und unter dem Mittelwert der Daten abhängt.

Festlegen von Optionen für Spezifikationsgrenzen

Hinweis: Diese Optionen sind für Histogramme, Wahrscheinlichkeitsplot, Q-Q-Plots und CDF-Plots verfügbar.

Klicken Sie im Auswahlfenster unter **Darstellung** auf **Spezifikationsgrenzen**, um die untere Spezifikationsgrenze, das Ziel und die obere Spezifikationsgrenze anzugeben.

Hinweis: Die Optionen für **Spezifikationsgrenzen** sind nur verfügbar, wenn Sie eine oder mehrere Grenzen für mindestens eine Analysevariable angeben. Um eine Spezifikationsgrenze für eine Analysevariable anzugeben, klicken Sie im Auswahlfenster auf **Daten** und wählen dann die Analysevariable aus.

Prozessfähigkeitsanalyse: Festlegen der Analysetypen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

Standardmäßig enthält der Prozessfähigkeitsanalysebericht die folgenden Tabellen: eine mit Grundkonfidenzintervallen, eine mit Grundmaßen, eine mit Tests auf Lageparametern und eine mit Momenten.

Um eine Tabelle zum Bericht hinzuzufügen, wählen Sie die Tabelle im Feld **Tabellen** aus.

Um die Analyseoptionen festzulegen, wählen Sie die Tabelle im Feld **Tabellen** aus. Sind Optionen verfügbar, erscheint rechts ein Feld, in dem Sie die Optionen angeben können.

Um alle Tabellen mit beschreibenden Statistiken im Bericht zu unterdrücken, aktivieren Sie die Option **Beschreibende Statistiken und Tabellen für Prozessfähigkeitsindizes unterdrücken**. Folgende Tabellen werden durch Auswahl dieser Option nicht unterdrückt:

- Tabellen, die von der ausgewählten Verteilung erstellt wurden. Sie können diese Tabellen unterdrücken, indem Sie im Fenster für diese Verteilung das Kontrollkästchen **Verteilungstabellen unterdrücken** aktivieren.
- Tabellen, die mit der INTERVALS-Anweisung erstellt wurden. Um diese Tabellen zu unterdrücken, können Sie die Option NOPRINT angeben. Weitere Informationen zur Anweisung INTERVALS finden Sie in der Hilfe zur Prozedur CAPABILITY.

Um Ihre Ergebnisse in eine Ausgabedatei zu speichern, aktivieren Sie das Kontrollkästchen **Ausgabestatistiken in einer Datei speichern**. Die Datei enthält univariate Statistiken und Prozessfähigkeitsindizes.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen

- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

16

Charakterisieren von Daten

Allgemeines über den Assistenten zum Charakterisieren von Daten 117
Charakterisieren von Daten: Auswählen von Daten 118
Charakterisieren von Daten: Berichtoptionen auswählen 118
Charakterisieren von Daten: Angeben der Ausgabegrenzen 119

Allgemeines über den Assistenten zum Charakterisieren von Daten

Der Assistent zum Charakterisieren von Daten unterstützt Sie bei der Erstellung von Zusammenfassungsberichten, Grafiken, Häufigkeits- und univariaten SAS-Dateien, die die wesentlichen Eigenschaften der Daten enthalten.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	CONTENTS, FREQ, UNIVARIATE
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Charakterisieren von Daten: Auswählen von Daten

1 Klicken Sie auf **Hinzufügen**. Das Dialogfeld **Öffnen** wird geöffnet.

2 Wählen Sie die gewünschte Datei aus. Drücken Sie die STRG-Taste, um mehrere Dateien auszuwählen. Klicken Sie auf **Öffnen**. Die ausgewählten Dateien werden im Feld **Daten** angezeigt.

3 Klicken Sie auf **Weiter**, um die Berichtsoptionen auszuwählen.

Charakterisieren von Daten: Berichtoptionen auswählen

1 Damit Sie eine der Berichtoptionen auswählen können, müssen Sie zunächst die Daten auswählen.

2 Wählen Sie die Ausgabetypen aus. Standardmäßig werden ein Zusammenfassungsbericht, Grafiken und SAS-Dateien für Häufigkeitsdaten und für univariate Daten erstellt.

   Wenn Sie **SAS-Dateien** als Ausgabetyp wählen, durchsucht SAS Enterprise Guide die im Dialogfeld **Optionen** im Bereich **Anwendungsroutinen** > **Ausgabebibliothek** definierte Bibliotheksliste und speichert die Ausgabedaten in der ersten Bibliothek dieser Liste, in die geschrieben werden kann. Klicken Sie auf **Durchsuchen**, wenn Sie den Speicherort ändern möchten. Die entstandenen Dateien werden ebenfalls dem Projekt hinzugefügt.

3 Klicken Sie auf **Weiter**, um die Ausgabegrenzen anzugeben.

Charakterisieren von Daten: Angeben der Ausgabegrenzen

1. Damit Sie die Ausgabegrenzen angeben können, müssen Sie zunächst eine Berichtoption auswählen.
2. Aktivieren Sie das Kontrollkästchen **Kategorievariablen begrenzen**, um die maximale Anzahl eindeutiger kategorischer Werte für den Bericht pro Variable anzugeben. Standardmäßig werden `30` Werte in den Bericht einbezogen.
3. Klicken Sie auf **Fertig stellen**, um die Ausgabe zu generieren.

17

Clusteranalyse

Allgemeines über die Anwendungsroutine 'Clusteranalyse' 121
Clusteranalyse: Zuweisen von Variablen zu Analyserollen 122
Clusteranalyse: Festlegen von Clusteroptionen 123
Clusteranalyse: Erzeugen von Plots und Diagrammen 125
Clusteranalyse: Festlegen von Ergebnisoptionen 126
Festlegen von Titeln und Fußnoten 128
Anzeigen von Eigenschaften 129

Allgemeines über die Anwendungsroutine 'Clusteranalyse'

Die Anwendungsroutine **Clusteranalyse** erstellt hierarchische Cluster der Beobachtungen in einer SAS-Datei, die entweder Koordinaten- oder Abstandsdaten enthält. Enthält die Datei Koordinatendaten, berechnet die Anwendungsroutine Euklidsche Abstände bevor Clustermethoden verwendet werden. Mithilfe dieser Anwendungsroutine können Sie die Ergebnisse der hierarchischen Clusterbildung in einem Baumstrukturdiagramm grafisch darstellen (in einem sogenannten Dendrogramm).

Alternativ kann die Anwendungsroutine **Clusteranalyse** mittels der K-Mittelwert-Methode nicht hierarchische Cluster von Koordinatendaten bilden. Sie können die K-Mittelwert-Methode nicht für Abstandsdaten verwenden. Sie können diese Anwendungsroutine zum Analysieren von Bevölkerungsdaten verwenden.

Angenommen, Sie möchten herausfinden, ob die Zahlen eines Landes für Geburtenraten, Sterberaten und Kindersterblichkeitsraten verwendet werden können, um bestimmte Ländertypen oder -kategorien festzulegen. Sie können eine Clusteranalyse durchführen, um herauszufinden, ob die Beobachtungen in Gruppen zusammengefasst werden können, die sich aus den Daten ergeben.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	CLUSTER, FASTCLUS, TREE
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Clusteranalyse: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Analysevariablen	führt Analysen mit den Variablen durch, die Sie dieser Rolle zuweisen. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits als **Satz-1-Variablen** bzw. **Satz-2-Variablen** zugewiesen haben.

Optionsname	Beschreibung
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
ID-Etikett	gibt die Werte an, die zum Identifizieren von Zeilen im Clusterverlauf und in der Ausgabebaumtabelle verwendet werden sollen. Wird dieser Rolle keine Variable zugewiesen, werde die Zeilen mit OBS*n* gekennzeichnet, wobei *n* für die Zeilennummer steht.
Variablen kopieren	kopiert die Werte für diese Variable in den Ausgabebaumdatensatz. Beobachtungen in der Ausgabebaumdatei, die Cluster mit mehreren Beobachtungen aus der Eingabedatei darstellen, weisen fehlende Werte für die kopierten Variablen auf.

Clusteranalyse: Festlegen von Clusteroptionen

Klicken Sie im Auswahlfenster auf **Cluster**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Clustermethode	gibt die in der Analyse zu verwendenden Clustermethoden an. ■ **Average Linkage** verwendet den Gruppenmittelwert und die ungewichtete Paar-Gruppen-Methode mit arithmetischen Mittelwerten (UPGMA). ■ **Centroid-Methode** verwendet die Centroid-Methode (die ungewichtete Paar-Gruppen-Methode, welche Centroide, UPGMC, Centroid-Sortierung und die Methode gewichteter Gruppen einsetzt). ■ **K-Mittelwert-Algorithmus** verwendet Euklidsche Abstände, sodass die Clustermittelpunkte auf der Schätzung der kleinsten Quadrate basiert. Bei dieser Methode wird davon ausgegangen, dass die Clustermittelpunkte den Mittelwerten der Beobachtungen entsprechen, die jedem Cluster zugewiesen werden, wenn der Algorithmus zum Abschließen der Konvergenz ausgeführt wird. Mit jeder Iteration wird das

Optionsname	Beschreibung
	Kriterium der kleinsten Quadrate verringert, bis die Konvergenz erreicht wurde. ■ **Minimum-Varianz-Methode nach Ward** verwendet die Minimum-Varianz-Methode nach Wert, die eine Fehlerquadratsumme, Ablaufverfolgung W berechnet.
K-Mittelwert-Clusteroptionen	gibt die zusätzlichen Optionen für den K-Mittelwert-Algorithmus an. ■ **Maximalanzahl an Clustern** gibt die maximale Anzahl der zulässigen Cluster an. ■ **Maximale Anzahl an Iterationen** gibt die maximale Anzahl an Iterationen zum Berechnen der Clusterausgangswerte an. ■ **Ausgangswertersetz.** gibt an, wie die Ausgangswertersetzung durchgeführt wird. □ **Voll** ersetzt den Ausgangswert, wenn der Abstand zwischen der Beobachtung und dem nächsten Ausgangswert größer ist als der Mindestabstand zwischen den Ausgangswerten, oder wenn der kleinste Abstand der Beobachtung zu allen Ausgangswerten außer dem nächsten Ausgangswert größer ist als der kleinste Abstand vom nächsten Ausgangswert zu allen anderen Ausgangswerten. □ **Partiell** ersetzt den Ausgangswert nur dann, wenn der Abstand zwischen der Beobachtung und dem nächsten Ausgangswert größer ist als der Mindestabstand zwischen den Ausgangswerten. □ **Keine** unterdrückt die Ausgangswertersetzung. □ **Zufällig** verwendet eine einfache Pseudo-Zufallsstichprobe einer kompletten Beobachtungsreihe als anfänglichen Clusterausgangswert. ■ **Zufälliger Ausgangswert** gibt eine positive Ganzzahl als Anfangswert für den Zufallszahlengenerator an. **Hinweis:** Diese Option ist nur verfügbar, wenn Sie Zufällig als Option für **Ausgangswertersetzung** wählten.

Clusteranalyse: Erzeugen von Plots und Diagrammen

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Hinweis: Plots und Diagramme sind nicht mit der Clustermethode K-Mittelwert-Algorithmus verfügbar.

Eine grafische Darstellung des Clustering-Prozesses kann bei der Interpretation der Cluster nützlich sein. Plots werden mit einer Teilmenge der Eingabedatenquelle erstellt. Wie die Daten in Teilmengen untergliedert werden, hängt davon ab, ob das kubisches Cluster-Kriterium (CCC) berechnet werden kann. Wenn die CCC-Statistik berechnet werden kann, werden die Daten anhand der nichtfehlenden Werte der CCC-Statistik in Teilmengen untergliedert. Wenn die CCC-Statistik nicht berechnet werden kann, werden die Daten anhand der nichtfehlenden Werte der Pseudo *F*-Statistik in Teilmengen untergliedert.

Optionsname	Beschreibung
Pseudo-F-Statistiken	zeigt ein Streuungsdiagramm für die Pseudo-*F*-Statistik gegen die Anzahl der Cluster an. **Hinweis:** Wenn Sie die Option **Pseudo-F- und t-Quadrat-Statistiken** im Fenster **Ergebnisse** auswählen, wird immer ein Graph der Pseudo-*F*-Statistik erstellt.
Pseudo t-Quadrat-Statistiken	zeigt ein Streuungsdiagramm für die Pseudo-t^2-Statistik gegen die Anzahl der Cluster an. **Hinweis:** Wenn Sie die Option **Pseudo-F- und t-Quadrat-Statistiken** im Fenster **Ergebnisse** auswählen, wird immer ein Graph der Pseudo-*F*-Statistik erstellt.
CCC-Statistiken	zeigt ein Streuungsdiagramm für die Statistik mit kubischem Clusterkriterium (CCC) gegen die Anzahl der Cluster an. **Hinweis:** Diese Option steht nicht zur Verfügung, wenn die Eingabedatei vom Typ DISTANCE ist.
Ausrichtung Baumdiagramm	zeigt ein Baumdiagramm der Cluster an. Die Objekte in den Clustern werden als Blätter dargestellt. Das Cluster, das alle Objekte enthält, ist die Wurzel. Ein Cluster, das mindestens zwei, nicht jedoch alle Objekte enthält, ist ein Zweig. Im Allgemeinen werden Blätter, Zweige und Wurzeln auch als Knoten bezeichnet. Wenn das Cluster A die Vereinigung der

Optionsname	Beschreibung
	Cluster B und C ist, ist A entsprechend das übergeordnete Objekt von B und C und B und C sind untergeordnete Objekte von A. Ein Blatt ist also ein Knoten ohne untergeordnete Objekte und eine Wurzel ein Knoten ohne übergeordnetes Objekt. Wenn jedes Cluster über mindestens zwei untergeordnete Objekte verfügt, ist das Baumdiagramm ein binärer Baum. Wählen Sie Horizontal, um das Baumdiagramm so auszurichten, dass die Höhenachse horizontal verläuft und die Wurzel am linken Rand liegt. Wählen Sie Vertikal, um das Baumdiagramm so auszurichten, dass die Höhenachse vertikal verläuft und die Wurzel am oberen Rand liegt. Dies ist die Standardeinstellung.

Clusteranalyse: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Ausgabedaten speichern	erzeugt die Ausgabedatei. Folgende Optionen können ausgewählt werden: ■ Clusterbaumdaten erstellt eine Ausgabedatei, die verwendet werden kann, um ein Baumdiagramm der Clusterhierarchie zu erstellen. Die Datei enthält eine Beobachtung für jede Beobachtung in der Eingabedatei sowie eine Beobachtung für jeden Cluster, der aus zwei oder mehreren Beobachtungen besteht (d.h. eine Beobachtung für jeden Knoten des Clusterbaums). Die Gesamtanzahl der Ausgabebeobachtungen beträgt normalerweise 2n-1, wobei n die Anzahl der Eingabebeobachtungen ist. ■ **Baumausgabe** erstellt eine Ausgabedatei für die Average-Linkage-, Centroid- und Ward-Methoden (minimale Varianz). Die Datei enthält eine Beobachtung für jedes Objekt im Baum sowie die Variablen CLUSTER und CLUSNAME, welche die Clusterzugehörigkeit auf einer beliebigen Stufe im Baum anzeigen. Sie können auch die Anzahl der Cluster angeben, die in der Tabelle berücksichtigt werden sollen.

Optionsname	Beschreibung
	◻ Um die Anzahl der Cluster für die Ausgabedatei festzulegen, aktivieren Sie das Kontrollkästchen **Cluster**. Standardmäßig beträgt die Anzahl der Cluster 10. Die tatsächliche Anzahl der Cluster kann jedoch von der von Ihnen angegeben Zahl abweichen, wenn eine der folgenden Bedingungen erfüllt ist: ▪ Die Anzahl der Blätter im Baum ist kleiner der Anzahl der gewünschten Cluster. ▪ Die Anzahl der nicht miteinander verbundenen Bäume in der Datei ist größer der Anzahl der gewünschten Cluster. ▪ Ein mehrfach verzweigter Baum enthält keine Ebene mit der angegebenen Anzahl an Clustern. ◻ Um die Anzahl der Ebenen für die Ausgabedatei festzulegen, aktivieren Sie das Kontrollkästchen **Entfernen** und geben Sie die Ebene des Baums an, mit der die getrennten Cluster in der Ausgabedatei definiert werden. Es werden nur die Cluster zwischen der Root und der angegebenen Ebene angezeigt. ▪ **K-Mittelwert-Cluster** erstellt eine Ausgabedatei für die K-Mittelwert-Algorithmus-Methode. Die Ausgabedatei enthält alle ursprünglichen Daten sowie die CLUSTER- und DISTANCE-Variablen, die durch die Clusteranalyse generiert werden. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Ausgabe anzeigen	zeigt die Ausgabe an. Folgende Optionen können angegeben werden: ▪ **Ausgabe anzeigen** zeigt Informationen zum Verlauf des Clusterbildungsprozesses an. Hierzu gehören die Anzahl der Cluster, die Namen der verbundenen Cluster und die Anzahl der Beobachtungen im neuen Cluster. Geben Sie im Feld Clustergenerationen an, wie viele Generationen des Clusterverlaufs angezeigt werden sollen. ▪ **Einfache beschreibende Statistiken** zeigt Mittelwerte, Standardabweichungen, Schiefe, Kurtosis und einen Koeffizienten der Bimodalität an. ▪ **Kubisches Cluster-Kriterium (CCC)** zeigt das kubische Clusterkriterium und das angenäherte, erwartete R-Quadrat unter der einheitlichen Nullhypothese an. Diese Option gilt nur für Koordinatendaten. Die Statistiken eignen sich zum Bestimmen der Anzahl an Clustern in den Daten. CCC-Werte über 2 oder 3 weisen auf gute Cluster hin. Werte zwischen 0

Optionsname	Beschreibung
	und 2 weisen auf potenzielle Cluster hin, die jedoch mit Vorsicht betrachtet werden sollten. Große negative Werte können auf Ausreißer hinweisen. **Hinweis:** Diese Option steht nicht zur Verfügung, wenn die Eingabedatei vom Typ DISTANCE ist. ■ **Pseudo-F- und t-Quadrat-Statistiken** zeigt Pseudo-*F*- und *t*-Quadrat-Statistiken an. Diese Statistiken können nützliche Hinweise zur Anzahl der Cluster enthalten. Relativ große Werte der Pseudo-*F*- Statistik weisen auf einen Haltepunkt hin. Gemäß der allgemeinen Regel für die Interpretation der Werte der Pseudo-*t*-Quadrat-Statistik sollte in der Spalte so weit nach unten gegangen werden, bis der erste Wert gefunden wird, der erheblich größer ist als der vorherige Wert. Anschließend muss in der Spalte wieder ein Cluster nach oben gegangen werden. **Hinweis:** Wenn Sie diese Option aktivieren, werden Plots der Pseudo-*F*- und *t*-Quadrat-Statistiken erzeugt, wenn Sie die Optionen Pseudo-*F*-Statistiken und Pseudo- *t*-Quadrat-Statistiken im Fenster **Plots** wählen. **Hinweis:** **Ausgabe anzeigen** ist die einzige verfügbare Anzeigeoption für die Clustermethode K-Mittelwert-Algorithmus.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

18

Vergleichen von Daten

Allgemeines zum Assistenten zum Vergleichen von Daten 131
Daten vergleichen: Auswählen von Daten 132
Daten vergleichen: Auswählen der zu vergleichenden Variablen 132
Daten vergleichen: Angeben der Vergleichskriterien 133
Daten vergleichen: Festlegen der Ausgabeoptionen 133

Allgemeines zum Assistenten zum Vergleichen von Daten

Der Assistent zum Vergleichen von Daten unterstützt Sie bei der Erstellung eines Berichts, der zwei Dateien oder zwei Variablen in einer oder in verschiedenen Dateien vergleicht.

Der Assistent zum Vergleichen von Daten ermöglicht Ihnen, die an einer SAS-Datei vorgenommenen Änderungen zu vergleichen. Beispiel: Sie haben eine Datei kopiert, und jetzt möchten Sie wissen, ob es zwischen der Kopie und der ursprünglichen Datei Unterschiede gibt. Mit dem Assistenten zum Vergleichen von Daten können Sie eine Zusammenfassung der an der ursprünglichen Datei vorgenommenen Änderungen anzeigen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	COMPARE
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	Ohne

Daten vergleichen: Auswählen von Daten

1. Geben Sie an, ob Sie Variablen aus zwei Dateien oder innerhalb derselben Datei vergleichen möchten.
2. Wählen Sie die Base- und Vergleichsdatei aus, die Sie verwenden möchten.
3. Klicken Sie auf **Weiter**, um die zu vergleichenden Variablen auszuwählen.

Daten vergleichen: Auswählen der zu vergleichenden Variablen

Standardmäßig werden alle Variablen mit dem gleichen Namen und vom gleichen Typ in der Tabelle verglichen. Sie können zusätzliche Variablen einschließen, die zu vergleichenden Variablen bearbeiten und Variablen löschen.

So geben Sie die zu vergleichenden Variablen an:

1. Um die zu vergleichenden Variablen auswählen zu können, müssen Sie zunächst die Daten auswählen.
2. Wählen Sie in der Dropdown-Liste **Daten vergleichen nach** das Kriterium für den Variablenvergleich aus. Sie können die Daten entweder nach Beobachtungen oder nach ID-Variablen vergleichen. Wenn Sie die Daten nach ID-Variablen vergleichen möchten, muss die Variable sowohl in der Base- als auch in der Vergleichsdatei vorhanden sein. Wenn Sie Variablen innerhalb derselben Datei vergleichen, können Sie die Daten nur nach Beobachtungen vergleichen.
3. Klicken Sie auf **Neu**, um der Tabelle ein neues Variablenpaar hinzuzufügen. Der Tabelle wird eine neue Zeile hinzugefügt. Wählen Sie mit Hilfe der Dropdown-Listen in jeder Zelle die zu vergleichenden Variablen aus.

   Wenn ein Variablenpaar falsch ist, wählen Sie in der Dropdown-Liste den richtigen Variablennamen aus. Sie können ebenso auf **Bearbeiten** klicken.

   Markieren Sie die Zeile eines Variablenpaars in der Tabelle und klicken Sie auf **Löschen**, wenn Sie ein Variablenpaar löschen möchten.
4. Klicken Sie auf **Weiter**, um die Vergleichskriterien anzugeben.

Daten vergleichen: Angeben der Vergleichskriterien

1. Um die Vergleichskriterien angeben zu können, müssen Sie zunächst die zu vergleichenden Variablen auswählen.
2. Geben Sie im Feld **Gleichheitskriterium** den Wert für das Gleichheitskriterium an.
3. Wählen Sie aus der Dropdown-Liste **Methode zur Gleichheitseinschätzung** die Methode zur Einschätzung der Gleichheit aus. Numerische Werte sind ungleich, wenn der durch die ausgewählte Methode festgelegte Betrag ihrer Differenz größer ist als der Wert des Gleichheitskriteriums. Die folgenden Methoden sind verfügbar:
   - **Absolut** - vergleicht die absolute Differenz der Werte mit dem Wert des Gleichheitskriteriums. Wenn der absolute Wert von y minus x größer als das Gleichheitskriterium ist, dann gelten die Werte als ungleich.
   - **Exakt** - testet die exakte Gleichheit. Wenn der Wert von y nicht gleich dem Wert von x ist, dann gelten die Werte als ungleich.
   - **Prozent** - vergleicht die absolute prozentuale Differenz mit dem Wert des Gleichheitskriteriums.
   - **Relativ** - vergleicht die absolute relative Differenz mit dem Wert des Gleichheitskriteriums.
4. Geben Sie an, wie fehlende Werte behandelt werden sollen. Folgende Optionen können ausgewählt werden: **Fehlenden Wert in BASE-Datei mit einem Wert desselben Typs gleichsetzen** oder **Fehlenden Wert in VERGLEICHS-Datei mit einem Wert desselben Typs gleichsetzen**.
5. Klicken Sie auf **Weiter**, um die Ausgabeoptionen festzulegen.

Daten vergleichen: Festlegen der Ausgabeoptionen

Standardmäßig erstellt die Anwendungsroutine "Daten vergleichen" einen HTML-Bericht mit einer Zusammenfassung der Datei, der Variablen, der Beobachtung und des Wertvergleichs sowie die Ergebnisse des Wertvergleichs. Sie können ebenso

eine Ausgabedatei erstellen, die die Ausgabe von Beobachtungen enthält und eine Ausgabedatei, die die Zusammenfassungsstatistiken enthält.

1 Um die Ausgabeoptionen angeben zu können, müssen Sie zunächst die Vergleichskriterien angeben.

2 Wählen Sie die Option **Ausgabedatei einschließen**, um eine Ausgabedatei zu erstellen, die eine Zeile für jede übereinstimmende Beobachtung enthält. Diese Datei enthält eine Spalte für jede Variable in der Beobachtung, eine Spalte für den Beobachtungstyp (_TYPE_) und eine Spalte für die Beobachtungsnummer (_OBS_). Mögliche Wertetypen der Spalte _TYPE_:

   - BASE - Die Werte in dieser Beobachtung stammen von einer Beobachtung in der Base-Datei.
   - COMPARE - Die Werte in dieser Beobachtung stammen von einer Beobachtung in der Vergleichsdatei.
   - DIF - Die Werte in dieser Beobachtung entsprechen der Differenz zwischen den Werten in der Base- und der Vergleichsdatei.
   - PERCENT - Die Werte in dieser Beobachtung entsprechen der prozentualen Differenz zwischen den Werten in der Base- und der Vergleichsdatei.

   Sie können entscheiden, dass folgende Beobachtungen in der Ausgabedatei enthalten sein sollen:

   - **Für jede Beobachtung in der Base-Datei eine Beobachtung schreiben** - schreibt die Beobachtungen der Base-Datei in die Ausgabedatei. Der Wert in der Spalte _TYPE_ der Ausgabedatei wird auf BASE gesetzt.
   - **Eine Beobachtung für jede Beobachtung in der Vergleichsdatei schreiben** - schreibt die Beobachtungen der Vergleichsdatei in die Ausgabedatei. Der Wert in der Spalte _TYPE_ der Ausgabedatei wird auf COMP gesetzt.
   - **Unterschiedlichen Wert einbeziehen** - schreibt die Differenz zwischen den Werten der Base- und der Vergleichsdatei in die Ausgabedatei. Der Wert in der Spalte _TYPE_ der Ausgabedatei wird auf DIF gesetzt.
   - **Werte zu prozentualen Unterschieden einbeziehen** - schreibt die prozentuale Differenz zwischen den Werten der Base- und der Vergleichsdatei in die Ausgabedatei. Der Wert in der Spalte _TYPE_ der Ausgabedatei wird auf PERCENT gesetzt.
   - **Beobachtungen unterdrücken, wenn alle Werte gleich sind** - schließt Beobachtungen, bei denen die Werte der Variablen gleich sind, nicht in die Ausgabedatei ein.

3 Aktivieren Sie die Option **Ausgabedatei für Zusammenfassungsstatistiken einschließen**, um eine Ausgabedatei zu erstellen, die für jede Zusammenfassungsstatistik für jedes Variablenpaar eine Zeile enthält. Die Ausgabedatei enthält folgende Spalten:

   - _VAR_ - enthält den Namen der Variablen aus der Base-Datei.
   - _WITH_ - enthält den Namen der Variablen aus der Vergleichsdatei.
   - _TYPE_ - enthält den Namen der Statistik in der Beobachtung.

- _BASE_ - enthält den Wert der Statistik, der sich aus den Werten der Variablen in der Base-Datei mit übereinstimmenden Beobachtungen in der Vergleichsdatei berechnet.
- _COMP_ - enthält den Wert der Statistik, der sich aus den Werten der Variablen in der Vergleichsdatei mit übereinstimmenden Beobachtungen in der Base-Datei berechnet.
- _DIF_ - enthält den Wert der Statistik, der sich aus den Differenzen der Werte der Variablen in der Base-Datei und der übereinstimmenden Variablen in der Vergleichsdatei berechnet.
- _PCTIF_ - enthält den Wert der Statistik, der sich aus den prozentualen Differenzen der Werte der Variablen in der Base-Datei und der übereinstimmenden Variablen in der Vergleichsdatei berechnet.

4 Geben Sie die maximale Anzahl von Differenzen für die Ausgabe an.

5 Klicken Sie auf **Fertig stellen**, um die Ausgabe zu generieren.

19

Konturdiagramm

Allgemeines über die Anwendungsroutine 'Konturdiagramm' ... **137**

Konturdiagramm: Auswählen eines Diagrammtyps ... **138**

Konturdiagramm: Zuweisen von Variablen zu Analyserollen ... **139**

Konturdiagramm: Festlegen von Darstellungsoptionen ... **140**
Festlegen der Optionen für Konturdiagramme ... 140
Weitere Darstellungsoptionen ... 141

Festlegen von Titeln und Fußnoten ... **141**

Anzeigen von Eigenschaften ... **142**

Allgemeines über die Anwendungsroutine 'Konturdiagramm'

Die Anwendungsroutine **Konturdiagramm** erstellt Linien-, gefüllte, Muster- oder Glättungsdiagramme, welche die mathematischen Beziehungen zwischen drei numerischen Variablen darstellen. Die Diagramme stellen dreidimensionale Zusammenhänge in zwei Dimensionen dar. Linien oder Flächen in einem Konturdiagramm stellen Größenebenen dar (z), die einer Position (x, y) in einer Ebene entsprechen.

Sie können diese Anwendungsroutine auch zur Illustration eines Vorkommens in der Natur verwenden. Beispielsweise können Sie ein Konturdiagramm erstellen, das den Prozentsatz von Klee illustriert, der in Erdstichproben aus verschiedenen Standorten einer Testumgebung gefunden wurde. Die X- und Y-Achsen des Diagramms entsprechen der grafischen Darstellung der Oberflächenhöhe an verschiedenen x-y-Positionen. Die Konturlinien innerhalb des Diagramms stellen die Standorte auf der Ebene dar, deren Prozentwerte in der Legende angegeben werden.

Weitere Informationen, siehe “Auswählen des Grafikausgabeformats” auf Seite 6.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GCONTOUR, G3GRID
Erforderliche SAS-Produkte	Base SAS, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	Ohne

Konturdiagramm: Auswählen eines Diagrammtyps

Klicken Sie im Auswahlfenster auf **Konturdiagramm**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Konturdiagramm	verwendet Linien zum Darstellen der Konturvariablen (Z).
Gefülltes Konturdiagramm	füllt jede Konturlinie mit einer anderen Farbe. Dabei entspricht eine Farbe jeweils einer Konturebene.
Gefülltes Umrissdiagramm	füllt jede Konturlinie mit einer anderen Farbe. Anschließend wird jede Fläche mit einer Umrisslinie versehen, um eindeutig identifizieren zu können, wo sich die Konturlinien befinden.
Musterkonturdiagramm	stellt die Konturebenen durch mit Mustern gefüllte Rechtecke dar. Das Muster für die einzelnen Rechtecke wird durch Berechnung des Mittelwerts der Werte der Z-Variablen für die vier Ecken des Rechtecks bestimmt. Anschließend wird das Muster der Ausprägung zugewiesen, die diesem Mittel am nächsten ist.
Glättungskonturdiagramm	verbindet nebeneinander liegende Gitterzellen mit demselben Muster, so dass sie einen einzigen Musterbereich bilden.

Konturdiagramm: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Horizontal	Die Spalte, die Sie dieser Rolle zuweisen, ist die horizontale oder X-Achsenvariable für das Diagramm.
Vertikal	Die Spalte, die Sie dieser Rolle zuweisen, ist die vertikale oder Y-Achsenvariable für das Diagramm.
Tiefe	Die Spalte, die Sie dieser Rolle zuweisen, ist die Variable der dritten Dimension bzw. der Z-Achse für das Plot.
Diagramme gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Diagramme, die erstellt werden. Für jeden eindeutigen Wert wird eine separate Grafik erzeugt.

Konturdiagramm: Festlegen von Darstellungsoptionen

Festlegen der Optionen für Konturdiagramme

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Kontur**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Anzahl der Konturebenen angeben	Geben Sie die Anzahl der Konturebenen an, d. h. die Anzahl der Bereiche, durch die Sie die Werte der Variablen dividieren, die der Rolle **Tiefe** (oder Z-Achsenvariable) zugewiesen wurde.
Automatisch Etiketten für Konturebenen erzeugen	Aktivieren Sie das Kontrollkästchen, um die Konturebenen automatisch zu beschriften. Das Etikett für jede Konturlinie ist der entsprechende Wert für die **Tiefe** der Konturebene. Die Legende erscheint immer noch. **Hinweis:** Diese Option wird von den Grafikformaten ActiveX, ActiveX-Bild, Java und Java-Bild nicht unterstützt.
Datenvorverarbeitung mit Interpolationsoptionen	Wenn Sie die Anwendungsroutine Konturdiagramm ausführen und im Log den Fehler 'Less than half of the grid cells have data values' erhalten, aktivieren Sie dieses Kontrollkästchen, und führen Sie die Anwendungsroutine erneut aus. Die Software erstellt eine Datei, deren horizontale (X und Y) Variablenwerte ein vollständiges Gitter bilden und interpoliert den Wert der vertikalen (z) Variablen für jeden Punkt auf der X-Y-Ebene. Sie können die Interpolationsmethode steuern, indem Sie eine dieser Interpolationsoptionen wählen: ■ **Joinen** Lineare Interpolation in einem Satz dreieckiger Bereiche. Diese Interpolationsmethode erstellt Werte im Bereich der Anfangswerte der vertikalen Variablen. Die resultierende interpolierte Oberfläche ist in der Regel jedoch nicht geglättet. ■ **Ohne Skalierung** verhindert vor der Interpolation die Skalierung der X- und Y-Variablen auf diesen Bereich. ■ **Partiell** Verwendet einen Spline, um die Ableitungen für die Interpolation zu schätzen. Ein bivariater Spline wird

Optionsname	Beschreibung
	an die nächsten Nachbarn angepasst. Er wird verwendet, um die benötigten Ableitungen zu schätzen. Die Konturen sind nicht genauso geglättet wie mit der Option Spline. Diese Option benötigt jedoch weniger Computerressourcen. ■ **Spline** verwendet einen bivariaten Spline für die Interpolation. Diese Methode kann sehr zeitaufwändig sein, besonders bei der Einbeziehung von mehr als 100 Datenpunkten.

Weitere Darstellungsoptionen

Einige Darstellungsoptionen gelten für alle Grafik-Anwendungsroutinen. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Festlegen von Achsenoptionen” auf Seite 745
- “Hinzufügen von Referenzlinien” auf Seite 748
- “Festlegen von Legendenoptionen” auf Seite 749
- “Anpassen des Diagrammbereichs” auf Seite 750
- “Festlegen erweiterter Optionen” auf Seite 756

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

20

Dateien kopieren

Allgemeines über die Anwendungsroutine 'Dateien kopieren' 143
Kopieren von Dateien von ihrem lokalen Computer auf einen SAS Server 144

Allgemeines über die Anwendungsroutine 'Dateien kopieren'

Mit der Anwendungsroutine **Dateien kopieren** ist es möglich, Dateien von Ihrem lokalen Computer auf einen SAS-Server zu übertragen. Ebenso können Dateien in die andere Richtung von einem SAS-Server auf einen lokalen Rechner transferriert werden. Die Anwendungsroutine **Dateien kopieren** ist mit einer FTP-Anwendung vergleichbar. Jedoch beruht diese Anwendungsroutine auf SAS-Protokollen für den Dateitransfer. Ein FTP-Server ist hierfür nicht erforderlich.

Die Anwendungsroutine verwendet SAS Integration Technologies, um die Dateien innerhalb des Netzwerks zu kopieren. Es besteht daher keine Verwandschaft zu PROC DOWNLOAD oder PROC UPLOAD in SAS/CONNECT.

Hinweis: Diese Anwendungsroutine ist nur im SAS Enterprise Guide verfügbar. Sie können den von dieser Anwendungsroutine erstellten Code nicht außerhalb von SAS Enterprise Guide ausführen. Wenn Sie ein SAS Enterprise Guide-Projekt oder einen Prozessfluss als SAS-Programm exportieren, werden keine Schritte aus der Anwendungsroutine "Dateien kopieren" einbezogen.

Kopieren von Dateien von ihrem lokalen Computer auf einen SAS Server

1. Wählen Sie den Server, der als Quell- oder Zielort für die zu kopierenden Dateien verwendet werden soll.
2. Wählen Sie den Transfertyp aus. Sie können Dateien von Ihrem lokalen Computer auf den SAS-Server hochladen oder vom SAS-Server auf den lokalen Rechner herunterladen.
3. Geben Sie den vollständigen Pfad für die zu kopierenden Dateien an. Im Feld **Zu kopierende Quelldateien** können Sie dazu Standardplatzhalter verwenden. Ein Sternchen (*) steht dabei für eine unbegrenzte Anzahl beliebiger Zeichen bis zum nächsten Zeichen, das kein Platzhalter ist. Ein Fragezeichen (?) in einer Zeichenkette ersetzt ein beliebiges einzelnes Zeichen, das sich an derselben Stelle im Dateinamen befindet.

   Nachfolgend einige Anwendungsbeispiele:

   `/u/dept/finance/*.xls`
   Entspricht allen XLS-Dateien im Verzeichnis /u/dept/finance.

   `C:\Data\??Sep2012.csv`
   Entspricht allen Dateien, deren Name mit zwei Zeichen beginnt und auf "Sep2012.csv" endet. Beispielsweise den Dateien "01Sep2012.csv" und "15Sep2012.csv". Die Datei mit dem Namen "199Sep2011.csv" entspricht nicht diesem Muster und würde nicht gefunden, da ihr Präfix aus drei anstelle von zwei Zeichen besteht.

   `C:\Data\15???2012.*`
   Stimmt mit allen Dateien überein, deren Name mit "15" beginnt und anschließend drei Zeichen sowie die Angabe "2012" enthält. Es kann sich dabei um eine Datei mit beliebiger Dateierweiterung handeln. Beispiele für Dateinamen nach diesem Muster: "15Sep2012.csv" und "15Oct2012.xls".

   Hinweis: Paltzhalter können nur in Dateinamen verwendet werden. Der Einsatz in Ordnernamen bzw. im übrigen Verzeichnispfad ist nicht zulässig. Beispiel: **`c:\data\Sept*\report.xls`** ist keine gültige Quelldateiangabe für diese Anwendungsroutine.

4. Geben Sie an, wohin die Dateien auf Ihrem lokalen Computer oder dem SAS-Server kopiert werden sollen. Beachten Sie beim Festlegen eines Zielordners folgende Regeln:

   - Der Ordner muss bereits auf dem Zielsystem vorhanden sein. Mit der Anwendungsroutine **Dateien kopieren** können keine neuen Ordner erstellt werden.

- Sie müssen über Schreibzugriff für den entsprechenden Ordner verfügen.
- Für den Zielordner muss ein absoluter Pfad angegeben werden (auch dann, wenn Sie Dateien hochladen und der SAS Workspace Server mit SASUSER als Start-Stammpfad oder mit einem anderen Nicht-System-Startpfad konfiguriert wurde).

5 Geben Sie an, ob bei der Datei- bzw. Ordnerangabe Makrovariablen verwendet werden sollen. Aktivieren Sie das Kontrollkästchen **SAS-Makrovariablen in Quell- und Zielpfaden auflösen**, wenn Sie in der Anwendungsroutine **Dateien kopieren** SAS-Makrovariablen und -Ausdrücke im Dateinamen oder dem Zielordnerpfad einsetzen möchten. Wenn Sie diese Option aktivieren, werden die Makrovariablen und Makrofunktionen evaluiert, bevor der Kopiervorgang gestartet wird.

Nachfolgend sehen Sie einige Beispiele für Werte mit Makroverweisen sowie deren Auflösung.

`/u/&SYSUSERID/Data`
Ergibt aufgelöst einen Ordner, der die Benutzer-ID des SAS-Benutzers auf dem SAS Workspace Server enthält.

`C:\Data\%sysfunc(substr(&SYSUSERID,,3))`
Ergibt aufgelöst einen Ordner, der einen Teil der SAS Benutzer-ID auf dem SAS Workspace Server enthält.

`/u/Data/&outputFile..xls`
Ergibt aufgelöst den Namen einer XLS-Datei, bei dem der Stamm des Dateinamens in der Makrovariable &OUTPUTFILE gespeichert ist. Diese Makrovariable nimmt an, dass ein zuvor gelaufener Prozess (etwa ein SAS-Programm) die Makrovariable so festgelegt hat, dass sie mit dem Namen einer vorhandenen Datei übereinstimmt.

WARNUNG
Die Makrovariablen müssen in Ihrer SAS-Sitzung vorhanden sein (entweder als integrierte Standard-SAS-Makrovariable oder sie muss zuvor definiert worden sein). Werden die Makrovariablen nicht korrekt angegeben oder es wird eine fehlerhafte Syntax für eine Makrofunktion verwendet, so kann dies bei der Ausführung der Anwendungsroutine "Dateien kopieren"zu Fehlern oder unerwarteten Ergebnissen führen.

Klicken Sie auf **Makrowerte testen**, um zu überprüfen, ob Makrovariablen korrekt aufgelöst werden. Ein kleines SAS-Programm überprüft daraufhin die Auflösung der Makrovariablen und zeigt die Ergebnisse in einem eigenen Fenster an. Beachten Sie beim Testen der Werte, dass die Werte basierend auf dem aktuellen Stand der Makrovariablen oder der Ergebnisse von Makrofunktionen ermitttelt werden. Diese Werte können sich bis Sie die Anwendungsroutine später ausführen verändern.

6 Geben Sie an, ob vorhandene Dateien im Zielordner überschrieben werden sollen. Standardmäßig werden vorhandene Dateien nicht durch die Anwendungsroutine "Dateien kopieren" überschrieben. Wenn der Name einer Quelldatei mit dem einer vorhandenen Datei übereinstimmt, wird diese Datei nicht von der Anwendungsroutine kopiert. Darüber hinaus wird ein entsprechender Vermerk ins Log geschrieben.

7 Geben Sie an, ob zu transferierende Textdateien mit der Anwendungsroutine automatisch erkannt und ggf. neu geschrieben werden sollen (mit der entsprechenden Formatierung am Zeilenende). In UNIX und Windows werden unterschiedliche Konventionen für die Zeichen am Zeilenende einer Textdatei verwendet. In UNIX-Umgebungen wird ein Zeilenvorschub-Zeichen (LF, Line-Feed), in Windows-Umgebungen hingegegen ein Zeilenvorschub (LF, Line-Feed) kombiniert mit einem nachfolgenden Wagenrücklauf-Zeichen (CR, Carriage-Return) eingesetzt. Durch das Fehlen des Wagenrücklaufs sind UNIX-Textdateien in Standardtextanwendungen, wie z.B. Notepad, ggf. schwieriger zu lesen.

21

Korrelationen

Allgemeines über die Anwendungsroutine 'Korrelationen' ... 147
Korrelation: Zuweisen von Variablen zu Analyserollen ... 148
Korrelation: Festlegen von Korrelationsoptionen ... 149
Korrelation: Festlegen von Ergebnisoptionen ... 151
Korrelationen: Erstellen der Ausgabedaten ... 152
Festlegen von Titeln und Fußnoten ... 153
Anzeigen von Eigenschaften ... 153

Allgemeines über die Anwendungsroutine 'Korrelationen'

Datenkorrelation ist eine statistische Prozedur für das Beschreiben des Zusammenhangs zwischen numerischen Variablen. Die Beziehung wird durch Berechnen von Korrelationskoeffizienten für die Variablen beschrieben.

Sie können die Anwendungsroutine **Korrelationen** verwenden, um detailliertere Informationen zu den Produktumsätzen zu erhalten. Beispielsweise können Sie mithilfe dieser Anwendungsroutine bestimmen, ob eine Korrelation zwischen der Größe eines Kunden (am Umsatz gemessen) und der Menge der Produkte, die dieser Kunde von Ihrem Unternehmen bezieht, vorliegt.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	CORR
Erforderliche SAS-Produkte	Base SAS, SAS/STAT

Name der Anforderung	Prozedur- und Produktnamen
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Korrelation: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Analysevariablen	führt Analysen mit den Variablen durch, die Sie dieser Rolle zuweisen. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Korrelieren mit	korreliert jede dieser Rolle zugewiesene Variable mit den Analysevariablen.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können die Analyse nicht nach einer Variable gruppieren, die Sie bereits als **Analysevariablen** oder **Korrelieren mit** zugewiesen haben.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n*

Optionsname	Beschreibung
	Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Partielle Variablen	gibt die Variablen an, die in der Pearson-Teilkorrelation, der Spearman-Teilkorrelation nach Rangfolge oder der Kendall-Teiltau-b-Korrelation verwendet werden sollen, je nachdem, welche Art von Korrelation Sie auswählen.
Relative Gewichtung	gibt den Wert der Variablen in jeder Zeile an, um die Produkt-Moment-Korrelationskoeffizienten zu berechnen. Sie können dieser Rolle maximal eine Spalte zuweisen.

Korrelation: Festlegen von Korrelationsoptionen

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Korrelationstypen	gibt die Korrelationsanalyse an. ■ **Pearson** berechnet die Pearsonsche Produkt-Moment-Korrelation. Hierbei handelt es sich um ein parametrisches Assoziationsmaß für zwei kontinuierliche Zufallsvariablen. Die Korrelationen liegen zwischen -1 und 1. ■ Mit der Option **Hoeffding** wird das Hoeffdingsche Abhängigkeitsmaß D berechnet. Hierbei handelt es sich um ein nichtparametrisches Assoziationsmaß, das allgemeinere Unabhängigkeitsabweichungen erkennt. Diese D-Statistik ist 30 Mal größer als die normale Definition und skaliert den Bereich zwischen -0,5 und 1, sodass nur große positive Werte auf eine Abhängigkeit hinweisen. ■ **Kendall** berechnet Kendalls tau-b-Koeffizient. Hierbei handelt es sich um ein nichtparametrisches Assoziationsmaß, das auf der Anzahl der Konkordanzen und Diskordanzen in Beobachtungspaaren basiert. Eine Konkordanz liegt vor, wenn die Beobachtungspaare zusammen variieren, während eine Diskordanz vorliegt, wenn Beobachtungspaare unterschiedlich variieren. Die Werte von Kendalls tau-b-Koeffizient liegen zwischen -1 und 1.

Optionsname	Beschreibung
	■ **Spearman** berechnet den Spearmanschen Rang-Korrelationskoeffizienten. Hierbei handelt es sich um ein nichtparametrisches Assoziationsmaß, das auf dem Rang der Datenwerte basiert. Die Korrelationen liegen zwischen -1 und 1. **Hinweis:** Die Korrelationstypen Hoeffding, Kendall und Spearman sind nicht verfügbar, wenn Sie der Rolle **Relative Gewichtung** eine Variable zuweisen. Der Korrelationstyp Hoeffding ist auch nicht verfügbar, wenn Sie der Rolle **Partielle Variablen** eine Variable zuweisen.
Optionen für Pearson-Korrelation	legt die Optionen für die Pearson-Korrelation fest: ■ **Cronbachs Alpha-Koeffizient** berechnet Cronbachs Alpha-Koeffizienten. Es werden separate Koeffizienten berechnet, die rohe und standardisierte Werte verwenden (durch die Skalierung der Variablen auf eine Einheitenvarianz von 1). Für jede Korrelationsvariable wird die Korrelation zwischen der Variablen und der Summe der verbleibenden Variablen berechnet. Cronbachs Alpha-Koeffizient wird nur mithilfe der verbleibenden Variablen berechnet. **Hinweis:** Diese Option ist nicht verfügbar, wenn Sie der Rolle **Korrelieren mit** eine Variable zuweisen. ■ **Kovarianzen** berechnet eine Kovarianzmatrix. Wenn Sie eine partielle Variable zuweisen, wird eine partielle Kovarianzmatrix berechnet. ■ **Quadratsumme und Kreuzprodukte** berechnet die Quadratsumme und Kreuzprodukte. Wenn Sie eine partielle Variable zuweisen, wird die nichtpartielle SSCP-Matrix berechnet. ■ **Korrigierte Quadratsumme und Kreuzprodukte** berechnet die korrigierten Quadratsummen und Kreuzprodukte. Wenn Sie eine partielle Variable zuweisen, wird eine nichtpartielle und eine partielle CSSCP-Matrix berechnet. ■ **Pearsonsche Korrelationen in den Ergebnissen unterdrücken** schließt Pearson-Korrelationen aus den Ergebnissen aus.
Fisher-Optionen	verwendet die Fisher-Optionen zur Anforderung von Konfidenzgrenzen und *p*-Werten unter einer angegebenen Nullhypothese für Korrelationskoeffizienten unter Einsatz der Fisher *z*-Transformation. **Hinweis:** Diese Optionen sind verfügbar, wenn Sie **Pearson** oder **Spearman** als Korrelationstyp wählen. ■ **Wert von alpha angeben** gibt das Niveau der Konfidenzgrenzen für die Korrelation an. Der Wert dieser

Optionsname	Beschreibung
	Option muss zwischen `0` und `1` liegen. Der Standardwert ist `0.05`. ■ **Typ der Konfidenzgrenzen** gibt den Typ der Konfidenzgrenzen an. Standard ist eine zweiseitige Konfidenzgrenze der zweiseitigen Alternative. ■ **Wert rho0 in der Nullhypothese angeben** gibt den Wert von rho0 in der Nullhypothese an. Der Standardwert ist `0`. ■ **Korrektur systematischer Fehler beim Erstellen von Konfidenzniveaus verwenden** gibt an, ob die Korrektur systematischer Fehler beim Erstellen von Konfidenzniveaus verwendet werden soll. Standardmäßig ist dieses Kontrollkästchen aktiviert und eine neue Korrelationsschätzung wird unter Verwendung der Korrektur systematischer Fehler vorgenommen.
Teiler für Varianz	gibt den Teiler an, der beim Berechnen von Varianzen, Standardabweichungen und Kovarianzen verwendet wird. Standardmäßig entspricht die Berechnung des Teilers für Varianz den Freiheitsgraden.
Zeilen mit fehlenden Werten für die korrelierten Variablen weglassen	verwendet nur die Zeilen, in denen keine der Analysevariablen fehlende Werte enthält. Standardmäßig verwendet die Anwendungsroutine **Korrelationen** alle nicht fehlenden Wertepaare für jedes Paar von Analysevariablen. Dies bedeutet, dass einige Korrelationen mit mehr Beobachtungen berechnet werden als andere.

Korrelation: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Plots	erzeugt für jedes Korrelationspaar ein Streuungsdiagramm.
Anzuzeigende Ergebnisse	legt die Optionen für die Korrelationsanalyse fest. Folgende Werte können in die Ergebnisse einbezogen werden:

Optionsname	Beschreibung
	■ Anzeige beschreibender Statistiken für die einzelnen Analysevariablen, einschließlich Mittelwert, Standardabweichung, Summe, Minimum und Maximum. ■ Anzeige der Signifikanzwahrscheinlichkeiten, die den einzelnen Korrelationen zugeordnet sind. ■ Zuordnen der Korrelationskoeffizienten zu Rangfolgen vom höchsten zum niedrigsten Rang. Für jede Analysevariable werden die **Korrelieren mit**-Variablen in absteigender Reihenfolge bezüglich ihres absoluten Werts des Korrelationskoeffizienten aufgelistet. Standardmäßig werden Korrelationskoeffizienten nicht vom höchsten zum niedrigsten Rang bewertet. Wenn Sie die Korrelationskoeffizienten vom höchsten zum niedrigsten Rang bewerten möchten, können Sie auch die Anzahl der Korrelationen angeben, die für jede Variable angegeben werden soll. Wenn Sie weniger Korrelationen angeben als maximal zugelassen sind, führt die Anwendungsroutine Korrelationen die Korrelationen vom höchsten zum niedrigsten absoluten Wert an und meldet die n besten Korrelationen. Dabei entspricht n der Anzahl der von Ihnen angegebenen Korrelationen.

Korrelationen: Erstellen der Ausgabedaten

Klicken Sie im Auswahlfenster auf **Ausgabedaten**, um auf diese Optionen zuzugreifen.

Aktivieren Sie das Kontrollkästchen **Ausgabedaten speichern**. Die Korrelationsstatistiken, die in der endgültigen Ausgabe berücksichtigt werden, hängen von den Korrelationstypen ab, die Sie im Bereich **Korrelationstypen** ausgewählt haben.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

22

Erstellen eines Formats

Allgemeines über die Anwendungsroutine 'Format erstellen' ... **155**

Erstellen eines Formats: Festlegen der Formatoptionen ... **156**

Erstellen eines Formats: Definieren des neuen Formats ... **157**
So definieren Sie ein neues Format ... 157
Beispiel 1: Alphanumerisches Format ... 158
Beispiel 2: Numerisches Format ... 159

Anzeigen von Eigenschaften ... **160**

Allgemeines über die Anwendungsroutine 'Format erstellen'

Die Anwendungsroutine **Format erstellen** erstellt ein numerisches oder alphanumerisches Format. Ein numerisches Format gibt den Text an, der angezeigt wird, wenn die gespeicherten Werte einer numerischen Variablen angezeigt oder gedruckt werden. Ein alphanumerisches Format führt dieselbe Funktion für eine alphanumerische Variable aus. Sie können diese Anwendungsroutine zum Formatieren Ihrer Ausgabe verwenden.

Mit dieser Anwendungsroutine können folgende Aufgaben erledigt werden:

- numerische Variablen als alphanumerische Werte drucken (z.B. der Wert "1" wird als "männlich", der Wert "2" als "weiblich" ausgegeben)
- eine alphanumerische Zeichenkette mit anderen Werten darstellen (z.B. der Wert "JA" wird als "YES" ausgegeben)
- numerische Werte mit einer Vorlage ausdrucken (z.B. der Wert "062214150" wird als "06221/4150" ausgegeben).

Hinweis: Diese Anwendungsroutine ist nur im SAS Enterprise Guide verfügbar.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	FORMAT
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	Ohne

Erstellen eines Formats: Festlegen der Formatoptionen

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Formatname	legt den Namen für das erstellte Format fest. Dieser Name wird künftig immer verwendet, wenn Sie dieses Format anwenden. Formatnamen müssen den folgenden Regeln entsprechen: ■ Sie dürfen nur Buchstaben, Ziffern und Unterstriche enthalten. Alle anderen Zeichen sind nicht zulässig. GESCHLECHT und M_ODER_W sind gültige Formatnamen, M-ODER-W und M#W jedoch nicht. ■ Sie dürfen nicht mit einer Zahl beginnen oder enden. HEX2DEC ist ein gültiger Formatname, 4MAT jedoch nicht. ■ Sie haben eine maximale Länge, die vom Formattypen abhängt. Der Name eines alphanumerischen Formats darf nicht länger als 31 Zeichen sein. Namen numerischer Formate dürfen maximal 32 Zeichen enthalten. **Hinweis:** Wenn Sie eine ältere Version als SAS 9 verwenden, darf der Name des alphanumerischen Formats nicht länger als 7 Zeichen, und der Name des numerischen Formats nicht länger als 8 Zeichen sein. Die Verwendung eines längeren Formatnamens kann unerwartete Ergebnisse hervorbringen. MAENNLICHODERWEIBLICH ist als Formatname zu lang und für keinen Formattyp zulässig. STATCODE ist ein gültiger Formatname für ein numerisches Format, als Name für ein alphanumerisches Format jedoch zu lang.

Optionsname	Beschreibung
Formatbreite angeben	gibt die Breite eines Formats an. Standardmäßig entspricht die Formatbreite der Länge des längsten formatierten Werts.
Speicherort zum Speichern des Formats	gibt die Bibliothek an, in der das neue Format gespeichert werden soll. Standardmäßig wird das Format in der Bibliothek EGTASK auf dem ausgewählten Server gespeichert, wenn vorhanden. Wenn keine Bibliothek EGTASK vorhanden ist, wird das Format in der SASUSER-Bibliothek gespeichert. In jedem Fall bleibt das Format weiterhin bestehen, wenn Sie die Anwendung beenden, und steht für die weitere Verwendung zur Verfügung. **Hinweis:** Wenn Sie das Format in der Bibliothek WORK speichern möchten, wird das Format beim Beenden gelöscht. Ein Format kann existieren und dennoch nicht zur Verfügung stehen. Ein Format ist nur dann verfügbar, wenn die Bibliothek, in der es gespeichert ist, im Suchpfad für Formate enthalten ist.
Formattyp	gibt an, ob ein alphanumerisches oder ein numerisches Format erstellt wird. Zum Schutz vor der inhärenten Ungenauigkeit eines Gleitkommaspeichers, wird ein numerisches Format mit einem Fuzzyfaktor angewendet. Standardmäßig lautet der Fuzzyfaktor `1E-12`. Um den Fuzzyfaktor zu ändern, aktivieren Sie das Kontrollkästchen **Fuzzyfaktor angeben** und geben einen neuen Wert in das Feld ein.

Erstellen eines Formats: Definieren des neuen Formats

So definieren Sie ein neues Format

Klicken Sie im Auswahlfenster auf **Formate definieren**, um auf diese Seite zuzugreifen.

Auf dieser Seite können Sie die gespeicherten Werte einer Variablen mit dem Text verknüpfen, der beim Anzeigen oder Drucken dieser Werte angezeigt wird. Die Formatdefinition teilt sich in zwei Schritte auf. In der Spalte **Etikett** wird der Text angegeben, der angezeigt oder ausgedruckt wird. Die Spalte **Bereiche** gibt den bzw. die Bereiche der gespeicherten Werte an, die in diesen Text umgewandelt werden. Eine rote Umrandung um einen Bereich bedeutet, dass es zwischen diesem Bereich und einem anderen eine Überschneidung gibt. Beispielsweise kommt es zu einer

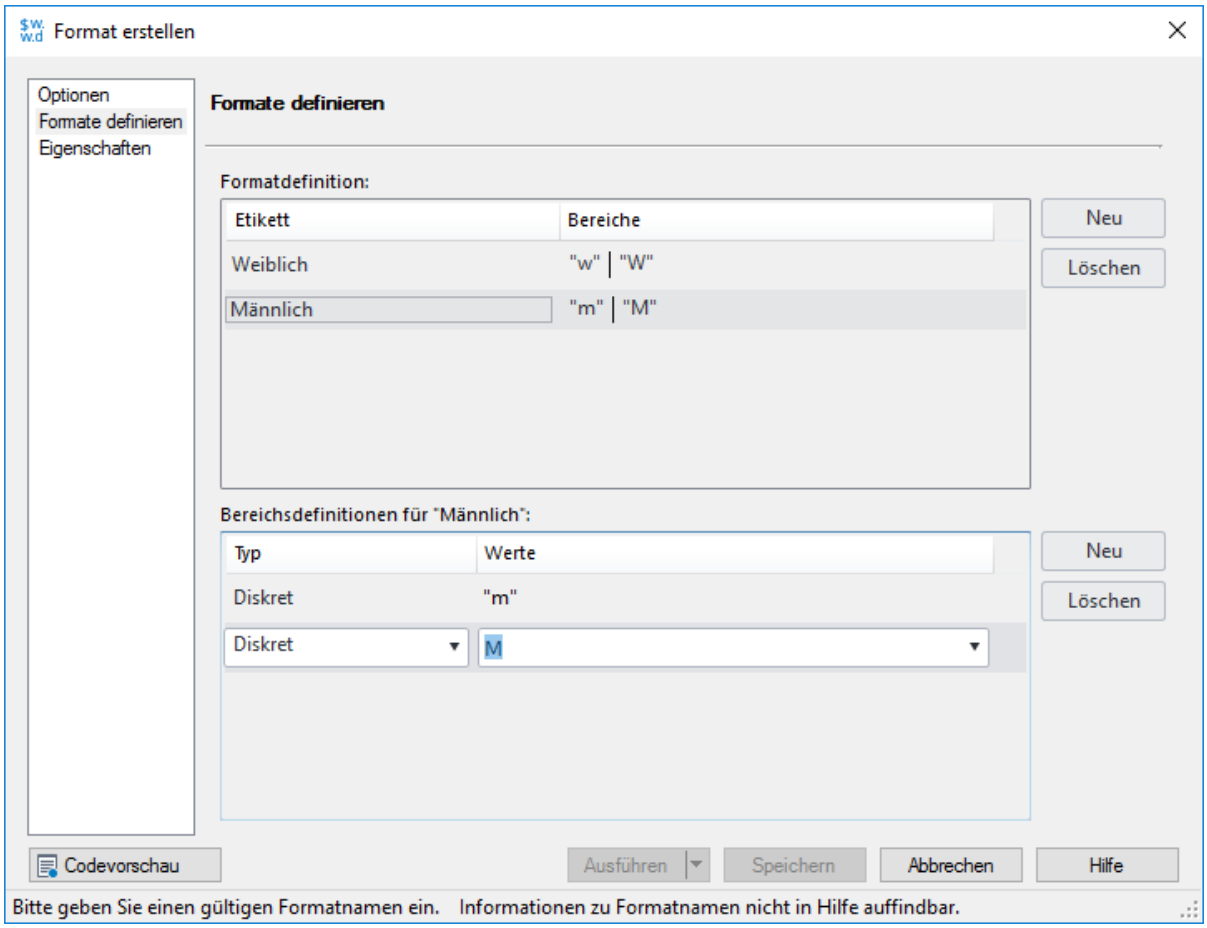
Überschneidung, wenn Sie alle Zahlen zwischen 0 und 5 dem Etikett NIEDRIG und alle Zahlen zwischen 3 und 7 dem Etikett MITTEL zuweisen.

So definieren Sie ein neues Format

1. Klicken Sie neben dem Feld **Formatdefinition** auf **Neu**. Den Feldern **Formatdefinition** und **Bereichsdefinitionen** wird eine neue Zeile hinzugefügt.
2. Geben Sie in der Spalte **Etikett** den Text ein, der angezeigt bzw. gedruckt werden soll.
3. Geben Sie im Feld **Bereichsdefinitionen** die gespeicherten Werte ein, denen Sie den Text zuordnen möchten. So legen Sie eine Bereichsdefinition fest:
   a. Wählen Sie einen Typ für die Bereichsdefinition aus (**Diskret** oder **Bereich**).
   b. Geben Sie die Werte ein, die dem Etikett in der Formatdefinition zugeordnet werden sollen. Jeder diskrete Werte bzw. Bereich muss einzeln eingegeben werden.

Beispiel 1: Alphanumerisches Format

Angenommen, Sie definieren ein alphanumerisches Format für eine Variable namens GESCHLECHT, in der die folgenden vier Werte gespeichert sind: w, W, m, M. Für die Werte w und W soll "Weiblich", für die Werte m und M soll "Männlich" angezeigt werden. Die Seite **Formate definieren** sollte dann wie folgt aussehen:

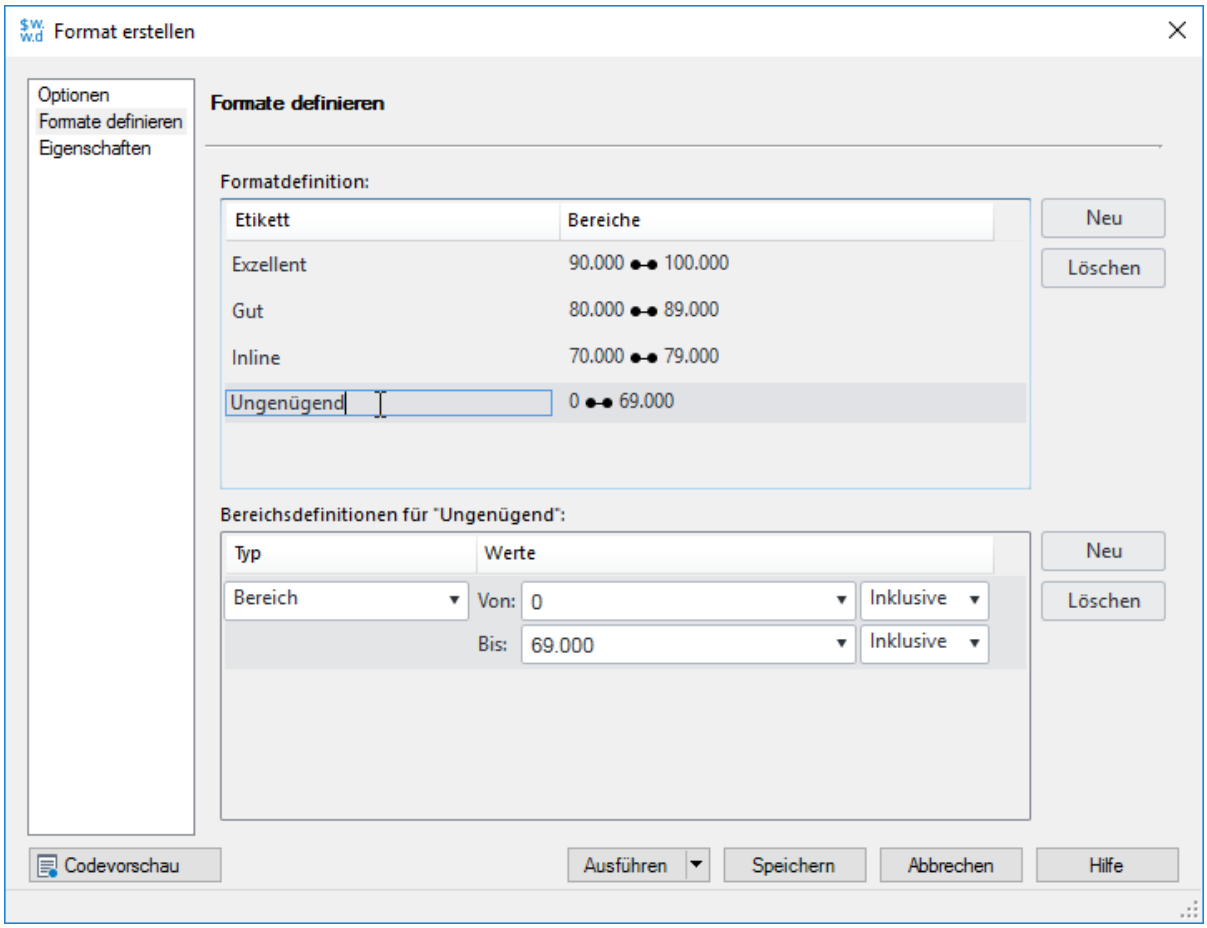


Beispiel 2: Numerisches Format

Angenommen Sie definieren ein numerisches Format für die Variable UMSATZ. Für festgelegte Umsatzbereiche soll ein Attribut in Textform angezeigt werden.

- Exzellent = 90.000 - 100.000 (inklusive)
- Gut = 80.000 - 89.000 (inklusive)
- Ausreichend = 70.000 - 79.000 (inklusive)
- Ungenügend = 0–69.000 (inklusive)

Die Seite **Formate definieren** sollte wie folgt aussehen:



Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten

- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

23

Format aus Datei erstellen

Allgemeines über die Anwendungsroutine 'Format aus Datei erstellen' 163
Erstellen eines Formats aus einer Datei 165

Allgemeines über die Anwendungsroutine 'Format aus Datei erstellen'

Die Anwendungsroutine **Format aus Datei erstellen** ermöglicht es Ihnen, unter Verwendung von Daten, die in einer SAS-Datei gespeichert sind, ein neues SAS-Format zu erstellen. Mit einem SAS-Format können Sie Rohdatenwerten formatierte Werte zuweisen. Sie können SAS-Formate für alphanumerische und numerische Daten erstellen.

Anwendungsbeispiele für die Anwendungsroutine **Format aus Datei erstellen**:

- Ihr Unternehmen plant für das nächste Jahr Kundenveranstaltungen in den einzelnen Vertriebsregionen (North, South, East, West). Die Umsatzdaten für 2009 enthalten die Adressinformationen der einzelnen Kunden nach Bundesstaat. Da Sie die Veranstaltungen jedoch in den Vertriebsregionen planen, benötigen Sie die Umsatzdaten nun nach Regionen kategorisiert. Dazu haben Sie bereits eine Zuordnung der einzelnen Bundesstaaten zu den Regionen in der Datei REGIONEN durchgeführt:

State	Region
Arizona	West
California	West
Florida	South

State	Region
Vermont	North
Virginia	East

Wenn Sie die Datei REGIONEN als Eingabedatei für die Anwendungsroutine **Format aus Datei erstellen** verwenden, können Sie nun ein SAS-Format erstellen, das Sie später auf die Werte für die Staaten in Ihrer Umsatzdatei anwenden können.

- Ihr Unternehmen hat in der jüngsten Vergangenheit eine Reihe medizinischer Tests durchgeführt. In den Patientendaten ist das Alter der Testpersonen enthalten. Die Datei ALTER wurde bereits in bestimmte Altersbereiche unterteilt:

Age	Low	High
Infant	0	2
Toddler	3	4
Child	5	9
Preteen	10	12
Teenage	13	17
Adult	18	100

Hinweis: Diese Anwendungsroutine ist nur im SAS Enterprise Guide verfügbar.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	FORMAT
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	Ohne

Erstellen eines Formats aus einer Datei

1 Wählen Sie die Eingabedatei aus. Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Quelldatei** angezeigt.

2 Geben Sie einen Namen für das Format an.

Formatnamen müssen den folgenden Regeln entsprechen:

- Sie dürfen nur Buchstaben, Ziffern und Unterstriche enthalten. Alle anderen Zeichen sind nicht zulässig. GESCHLECHT und M_ODER_W sind gültige Formatnamen, M-ODER-W und M#W jedoch nicht.
- Sie dürfen nicht mit einer Zahl beginnen oder enden. HEX2DEC ist ein gültiger Formatname, 4MAT jedoch nicht.
- Sie haben eine maximale Länge, die vom Formattypen abhängt. Der Name eines alphanumerischen Formats darf nicht länger als 31 Zeichen sein. Namen numerischer Formate dürfen maximal 32 Zeichen enthalten.

Hinweis: Wenn Sie eine ältere Version als SAS 9 verwenden, darf der Name des alphanumerischen Formats nicht länger als 7 Zeichen, und der Name des numerischen Formats nicht länger als 8 Zeichen sein. Die Verwendung eines längeren Formatnamens kann unerwartete Ergebnisse hervorbringen. MAENNLICHODERWEIBLICH ist als Formatname zu lang und für keinen Formattyp zulässig. STATCODE ist ein gültiger Formatname für ein numerisches Format, als Name für ein alphanumerisches Format jedoch zu lang.

3 Geben Sie an, ob Sie ein alphanumerisches oder numerisches Format erstellen. Zum Schutz vor der inhärenten Ungenauigkeit eines Gleitkommaspeichers, wird ein numerisches Format mit einem Fuzzyfaktor angewendet. Standardmäßig lautet der Fuzzyfaktor `1E-12`.

4 Wählen Sie die Bibliothek aus, in der das neue Format gespeichert werden soll. Standardmäßig wird das Format in der Bibliothek WORK auf dem gewählten Server gespeichert. Beachten Sie, dass es sich bei der Bibliothek um einen Speicher für temporäre Daten handelt. Ihr Format wird daher gelöscht, wenn Sie den SAS Enterprise Guide beenden. Wählen Sie daher eine andere Ausgabebibliothek, wenn Sie Ihr Format auch künftig verwenden möchten.

Hinweis: Ein Format kann existieren und dennoch nicht zur Verfügung stehen. Ein Format ist nur dann verfügbar, wenn die Bibliothek, in der es gespeichert ist, im Suchpfad für Formate enthalten ist.

5 Geben Sie an, ob die Werte in der Eingabedatei diskrete Werte sind oder ob es sich um einen Wertebereich handelt.

So erstellen Sie ein Format aus diskreten Werten:

a Wählen Sie im Feld **Wertetyp** die Option **Discrete/Lookup**.

b Wählen Sie im Bereich **Variablen** aus der Dropdown-Liste **Diskrete Werte** die Variable in der Datei aus, in der die zu formatierenden Werte enthalten sind. Die Werte für die ausgewählte Variable werden in der Vorschautabelle in der Spalte **Lookup-Wert** angezeigt.

c Wählen Sie aus der Dropdown-Liste **Etiketten** die Variable in der Datei aus, in der die Etiketten für die Werte enthalten sind. Die Werte für die ausgewählte Variable werden in der Vorschautabelle in der Spalte **Etikett** angezeigt.

So erstellen Sie ein format aus einem Wertebereich:

a Wählen Sie im Feld **Wertetypen** die Option **Bereiche** aus.

b Wählen Sie im Bereich **Variablen** aus der Dropdown-Liste **Niedrige Werte** die Variable in der Datei aus, in der Anfangswerte für den Bereich enthalten sind. Die Werte für die ausgewählte Variable werden in der Vorschautabelle in der Spalte **Start** angezeigt.

c Wählen Sie aus der Dropdown-Liste **Hohe Werte** die Variable in der Datei aus, in der die Endwerte für den Wertebereich enthalten sind. Die Werte für die ausgewählte Variable werden in der Vorschautabelle in der Spalte **Ende** angezeigt.

d Wählen Sie aus der Dropdown-Liste **Etiketten** die Variable in der Datei aus, in der die Etiketten für die Werte enthalten sind. Die Werte für die ausgewählte Variable werden in der Vorschautabelle in der Spalte **Etikett** angezeigt.

6 Legen Sie die maximale Länge des Etiketts fest. Standardmäßig kann ein Etikett bis zu acht Zeichen lang sein.

7 Geben Sie ein Etikett für sonstige Werte an. Wenn Sie dieses Format auf eine Datenquelle anwenden, kann es vorkommen, dass einige Werte nicht entsprechend formatiert werden. Sie können für Werte, denen kein formatierter Wert zugewiesen werden kann ein eigenes Format erstellen. Standardmäßig erhalten diese Werte das Format **Sonstige**.

8 Geben Sie an, ob in den Ergebnissen ein Zusammenfassungbericht enthalten sein soll. Standardmäßig wird kein Zusammenfassungsbericht erstellt.

24

Erstellen einer Feature-Tabelle für Karten

Allgemeines über die Anwendungsroutine 'Feature-Tabelle für Karten erstellen' **167**

Erstellen einer Feature-Tabelle für Karten: Zuweisen von Variablen zu Analyserollen **168**

Erstellen einer Feature-Tabelle für Karten: Festlegen von Optionen **169**
Festlegen von Optionen für die Kartendatenprojektion 169
Festlegen von Optionen für die Feature-Tabelle 170

Erstellen einer Feature-Tabelle für Karten: Festlegen der Ergebnisse **171**

Festlegen von Titeln und Fußnoten **171**

Anzeigen von Eigenschaften **171**

Allgemeines über die Anwendungsroutine 'Feature-Tabelle für Karten erstellen'

Mit der Anwendungsroutine **Feature-Tabelle für Karten** erstellen wird eine Feature-Tabelle für benutzerdefinierte oder andere Kartendateien erstellt, die Sie mit der Anwendungsroutine Kartengrafik verwenden. Sie ermöglicht Ihnen auch die Projektion einer Kartendatei - d. h. die sphärischen Koordinaten in einer Kartendatei werden in kartesische Koordinaten konvertiert. Diese Anwendungsroutine ermöglicht Ihnen die Verwendung einer von mehreren Kartenprojektionstechniken zur Projektion der Koordinaten in einer traditionellen Kartendatei auf eine zweidimensionale Ebene, während gleichzeitig versucht wird, Verzerrungsbereich, Abstand, Richtung und Formeigenschaften der ursprünglichen Kugel zu minimieren.

Hinweis: Diese Anwendungsroutine erfordert das mit SAS Release 8.2 und späteren Releases ausgelieferte Makro %ADDFEAT. Ansonsten wenden Sie sich

bitte an die Technische Hotline von SAS, um zu erfahren, wie Sie dieses Makros bekommen können.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GPROJECT
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	KARTEN

Erstellen einer Feature-Tabelle für Karten: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Hinweis: Die Eingabedatenquelle für diese Anwendungsroutine muss eine X- und eine Y-Spalte enthalten, die wiederum die räumlichen Datenkoordinaten für die Karte enthalten. Die Bibliothek "Maps"enthält Kartendateien, die bereits einer Feature-Tabelle zugewiesen wurden. Die Datei "US2" ist beispielsweise die Feature-Tabelle für die Karte der Vereinigten Staaten.

Die Bibliothek "Maps" ist geschützt, sodass Sie die Dateien in dieser Bibliothek nicht als Eingabedatenquelle für die Anwendungsroutine "Feature-Tabelle für Karten erstellen" verwenden können. Sie können jedoch Dateien aus der Bibliothek "Maps" an einen ungeschützten Speicherort kopieren und sie dann in dieser Anwendungsroutine verwenden.

Optionsname	Beschreibung
Identifikationsvariablen	Die Variablen, die Sie dieser Rolle zuweisen, identifizieren Abschnitte in der Karte. Mindestens eine Identifikationsvariable ist erforderlich, um die Daten zu projizieren und/oder eine Feature-Tabelle zu erstellen.

Optionsname	Beschreibung
Beizubehaltende Attributvariablen	Dies sind Variablen in der Eingabekartendatei (Geometrie), die ebenfalls in der Feature-Tabelle enthalten sein sollen. **Hinweis:** Diese Datenrolle wird nur angezeigt, wenn die Option **Feature-Tabelle für Karten erstellen** ausgewählt ist.
Zu projizierende Längenvariable	Dies ist die Variable in der Eingabedatei, welche den unprojizierten (sphärischen) Datenpunkt für die Länge enthält. **Hinweis:** Diese Datenrolle wird nur angezeigt, wenn die Option **Kartendaten projizieren (aus sphärischen kartesischen Koordinaten)** ausgewählt ist.
Zu projizierende Breitenvariable	Dies ist die Variable in der Eingabedatei, welche den unprojizierten (sphärischen) Datenpunkt für die Breite enthält. **Hinweis:** Diese Datenrolle wird nur angezeigt, wenn die Option **Kartendaten projizieren (aus sphärischen kartesischen Koordinaten)** ausgewählt ist.

Erstellen einer Feature-Tabelle für Karten: Festlegen von Optionen

Festlegen von Optionen für die Kartendatenprojektion

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

1 Markieren Sie das Kontrollkästchen **Kartendaten projizieren (aus sphärischen kartesischen Koordinaten)**.

2 Wählen Sie die geeignete Projektionsmethode aus der Dropdown-Liste **Projektionsmethode**.

- **Flächentreue Schnittkegelprojektion nach Albers** ist eine konische Projektion von der Oberfläche der Kugel zu einem Kegelsekans zur Kugel. Dabei wird diese an zwei Standardbreitenparallelen geschnitten. Die Achse

des Kegels fällt mit einer Verlängerung der Polarachse der Kugel zusammen. Jeder Abschnitt der resultierenden Karte weist ein konstantes Verhältnis zur Kugelfläche auf. Im Allgemeinen vergrößert sich die Formverzerrung zu den Breitenpolen hin, die außerhalb der beiden Standardparallelen liegen. Die Albers-Projektion eignet sich hervorragend zur Darstellung von Bereichen mit großer und kleiner Ost-West-Ausdehnung. Allerdings müssen beide Standardparallelen auf derselben Seite des Äquators liegen.

- **Lamberts konforme Projektion** ist eine Projektion, bei der die Längenmeridiane von der Spitze eines Sekanskegels ausstrahlen, während es sich bei den Breitenparallelen um konzentrische Kreise handelt. Diese Projektion eignet sich optimal für Navigationsdiagramme und -karten mit relativ kleiner Ost-nach-West-Ausdehnung. Allerdings müssen beide Standardparallelen auf derselben Seite des Äquators liegen.
- **Gnomonische Projektion (azimutal)** ist eine direkte planare Projektion von der Oberfläche der Kugel auf eine imaginäre Ebene, die tangential zur Kugel am Kartenprojektionspol liegt. Standardmäßig befindet sich der Projektionspol in der Mitte der zu projizierenden Kartendatei. Die Verzerrung wird mit zunehmendem Abstand vom Kartenpol erhöht. Diese Projektion eignet sich am besten zum Zuordnen der Flächen mit kleiner Ost-nach-West-Ausdehnung.

3 Wählen Sie die **Koordinateneinheiten** für die unprojizierten Längen- und Breitenwerte in der Eingabekartendatei. Der Standardwert ist Radiane.

4 Wählen Sie die **Richtung zunehmender Länge**. Der Standardwert ist West.

5 Wenn Sie Beobachtungen beibehalten möchten, bei denen die projizierten Werte der X- und Y-Variablen mit denen in der vorherigen Beobachtung identisch sind, aktivieren Sie das Kontrollkästchen **Nachfolgende identische Beobachtungen beibehalten**. Standardmäßig werden aufeinanderfolgende identische Beobachtungen gelöscht.

Festlegen von Optionen für die Feature-Tabelle

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

So erstellen Sie eine neue Feature-Tabelle, die mit der Eingabekartendatei verknüpft ist:

1 Markieren Sie das Kontrollkästchen **Feature-Tabelle für Karten erstellen**.

2 Wenn Sie den Namen der Kartengeometrievariablen ändern möchten, geben Sie den neuen Namen im Feld **Name der neuen Kartengeometrievariablen** an. Der Name darf nicht mit einer Zahl beginnen und keine Leerzeichen, Satzzeichen oder Sonderzeichen enthalten.

Erstellen einer Feature-Tabelle für Karten: Festlegen der Ergebnisse

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Standardmäßig speichert diese Anwendungsroutine die Dateien der Projektkarte und Feature-Tabelle in der SASUSER-Bibliothek. Um Ihren Ausgabedaten einen anderen Speicherort zuzuweisen, verwenden Sie das Dialogfeld **Speichern unter**.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

25

Assistent zum Erstellen von Zeitreihendaten

Allgemeines über den Assistenten zum Erstellen von Zeitreihendaten 173
Assistent zum Erstellen von Zeitreihendaten: Auswählen der Daten 174
Assistent zum Erstellen von Zeitreihendaten: Zuweisen von Variablen zu Rollen 174
Assistent zum Erstellen von Zeitreihendaten: Festlegen der Optionen für Akkumulation, fehlende Werte und Null-Werte 175
Assistent zum Erstellen von Zeitreihendaten: Zusätzliche Ausgaben auswählen 176
Assistent zum Erstellen von Zeitreihendaten: Festlegen von Titeln und Fußnoten 176

Allgemeines über den Assistenten zum Erstellen von Zeitreihendaten

Der **Assistent zum Erstellen von Zeitreihendaten** hilft Ihnen bei der Umwandlung transaktionaler Daten in Zeitreihen mit festem Intervall. Transaktionale Daten sind Daten, die unregelmäßig erfasst wurden (breite Streuung der Zeitstempel).

Als Beispiel könnten dies Umsatzdaten von Ihrer Unternehmenswebseite sein. Diese Umsatzdaten weisen einen Zeitstempel auf, wurden aber nicht in einem bestimmten zeitlichen Intervall erfasst. Mit dem Assistenten können Sie solche Daten in Zeitreihendaten umwandeln.

Assistent zum Erstellen von Zeitreihendaten: Auswählen der Daten

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Klicken Sie auf **Weiter**, um Variablen Rollen zuzuweisen.

Assistent zum Erstellen von Zeitreihendaten: Zuweisen von Variablen zu Rollen

1 Bevor Sie Variablen Rollen zuweisen können, müssen Sie die Daten auswählen.

2 Weisen Sie diesen erforderlich Rollen Variablen zu:

- **Zeit-ID** - Die Spalte, die Sie dieser Rolle zuweisen, ist die Zeit-ID für die Daten. Die Variable Zeit-ID ist eine Variable in der Eingabedatei, die den SAS-Datums-, -Zeit- oder -Datetime-Wert jeder Beobachtung enthält. Diese Variable wird verwendet, um die Häufigkeit und die Reihenfolge der Daten zu bestimmen, und um die Zeit-ID-Werte für die Prognosen zu extrapolieren. Sie können dieser Rolle nur eine Variable zuweisen, die eine Datumsvariable, eine Datetime-Variable oder eine numerische Variable sein muss, die Datums- oder Datetime-Werte enthält.
- **Zeitreihen für** - Die Spalte, die Sie dieser Rolle zuweisen, ist die Variable für die Modellbildung und für die Prognose. Sie müssen dieser Rolle mindestens eine numerische Variable zuweisen. Legen Sie mit dem Nach-oben- bzw. Nach-unten-Pfeil die Reihenfolge der Variablen fest.

3 (Optional) Weisen Sie der Rolle **Eigene Zeitreihen für Werte von (Gruppierungsvariable)** eine Variable zu. Diese Variable gruppiert Beobachtungen mit dem gleichen Wert für die Gruppierungsvariable. Durch die Zuweisung einer Gruppierungsvariablen erhalten Sie separate Analysen für

gruppierte Beobachtungen. Sie können dieser Rolle alphanumerische und numerische Variablen zuweisen.

Die Reihenfolge der BY-Variablen beschreibt die Struktur der Hierarchie. Legen Sie mit dem Nach-oben- bzw. Nach-unten-Pfeil die Reihenfolge der Variablen fest.

4 Klicken Sie auf **Weiter**, um die Optionen für Akkumulation, fehlende und Null-Werte festzulegen.

Assistent zum Erstellen von Zeitreihendaten: Festlegen der Optionen für Akkumulation, fehlende Werte und Null-Werte

1 Bevor Sie diese Optionen festlegen können, müssen Sie den Rollen Variablen zuweisen.

2 Wenn die Anwendungsroutine das Intervall aus der Eingabedatei erkannt hat, wird der Standardintervall in der Dropdown-Liste **Intervall** angezeigt. Sie können das Intervall ändern. Die Dropdown-Liste enthält jedoch nur die Intervalle, die für die Variable geeignet sind, die Sie der Rolle **Zeit-ID** zugewiesen haben.

Bei großen Datenquellen braucht die Anwendungsroutine eventuell länger, um ein Intervall zu erkennen. Wurde das Intervall nicht erkannt, wenn Sie bei diesem Schritt angelangt sind, können Sie aus der Dropdown-Liste **Intervall** ein Intervall auswählen. Bis ein Intervall erkannt wird, werden alle möglichen Werte für die Intervalloptionen aufgeführt. Nach Erkennung des Intervalls wird eine Warnmeldung angezeigt, wenn der Variablentyp, den Sie der Rolle Zeit-ID zugewiesen haben und der Intervallwert miteinander in Widerspruch stehen. Die Liste der Intervallwerte wird geändert, so dass nur die für den Variablentyp geeigneten Werte angezeigt werden. Weitere Informationen zu den einzelnen Statistiken finden Sie im Glossar für statistische Begriffe.

3 Wählen Sie aus dem Bereich **Zeitreihen** die Zeitreihenvariable aus. Diese Variablen haben Sie der Rolle **Zeitreihen für** zugewiesen.

4 Legen Sie die Statistik zum Aggregieren von Daten in verschiedenen Hierarchieebenen in der Dropdown-Liste **Akkumulationsstatistiken** fest.

5 Legen Sie in der Dropdown-Liste **Fehlende Werte wie** den Umgang mit fehlenden Werten fest.

6 Wiederholen Sie die Schritte 3-5 für jede neue Zeitreihenvariable.

7 Klicken Sie auf **Weiter**, um die Ausgabeoptionen festzulegen.

Assistent zum Erstellen von Zeitreihendaten: Zusätzliche Ausgaben auswählen

1 Vor dem Festlegen der Ausgabeoptionen müssen Sie die Optionen für Akkumulation, fehlende Werte und Null-Werte festlegen.

2 Aktivieren Sie das Kontrollkästchen **Zeitreihendiagramm**, wenn Sie ein Diagramm der Zeitreihendaten in die Ausgabe einbeziehen möchten.

3 Aktivieren Sie das Kontrollkästchen **Zeitreihen in Datei speichern**.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

4 Klicken Sie auf **Weiter**, um die Titel und Fußnoten anzugeben.

Assistent zum Erstellen von Zeitreihendaten: Festlegen von Titeln und Fußnoten

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

1 Vor der Angabe des Titels und der Fußnoten müssen Sie die Ausgabeoptionen festlegen.

2 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld. Sie können in Titeln und Fußnoten Makrovariablen verwenden.

3 Klicken Sie auf **Fertig stellen**, um die Anwendungsroutine auszuführen.

26

Dateiattribute

Allgemeines über den Assistenten für Dateiattribute 179
Dateiattribute: Auswählen von Daten 180
Dateiattribute: Festlegen der Ausgabe 180

Allgemeines über den Assistenten für Dateiattribute

Der Assistent für Dateiattribute unterstützt Sie bei der Erstellung eines Berichts, der sowohl das Datum und die Uhrzeit der Erstellung, den Speicherort der Datei, die Anzahl der Beobachtungen als auch die Variablennamen, Etiketten, Typen und Formate enthält.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	DATASETS
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	Ohne

Dateiattribute: Auswählen von Daten

1. Klicken Sie auf **Hinzufügen**. Das Dialogfeld **Öffnen** wird geöffnet.
2. Markieren Sie die Datei(en), die Sie verwenden möchten. Drücken Sie die Strg-Taste, um mehrere Dateien auszuwählen. Klicken Sie auf **Öffnen**. Die ausgewählten Dateien werden im Feld **Daten** angezeigt.
3. Klicken Sie auf **Weiter**, um die Ausgabe anzugeben.

Dateiattribute: Festlegen der Ausgabe

1. Damit Sie die Ausgabeoptionen angeben können, müssen Sie zunächst die Daten auswählen.
2. Geben Sie den Speicherort für die Ausgabedatei an. Standardmäßig wird die Ausgabedatei in der Bibliothek SASUSER gespeichert. Wenn Sie den Speicherort oder den Namen der Ausgabedatei ändern möchten, klicken Sie auf **Durchsuchen**. Die entstandene Datei wird ebenfalls dem Projekt hinzugefügt.
3. Wählen Sie den Berichtstyp aus. Folgende Optionen können ausgewählt werden:
   - **Standardbericht** - enthält die Ausgabe aus der Prozedur DATASETS. Dieser Bericht enthält die folgenden Datenattribute: das Datum, an dem die Datei erstellt und zuletzt geändert wurde, die Anzahl der Beobachtungen, die Kodierung, alle Engine/Host-abhängigen Informationen und eine alphabetische Liste der Variablen und ihrer Attribute.
   - **Erweiterter Bericht** - enthält die Ausgabe aus der Prozedur REPORT. Dieser Bericht zeigt die Tabellen- und Variablenattribute an. In diesem Bericht können Sie den Tabellentyp, das Erstellungs- und Änderungsdatum der Tabelle, die Anzahl der Beobachtungen, die Variablenetiketten und die Variablentypen festlegen. Mit folgenden Optionen können Sie die Zeilen in der Variablentabelle in aufsteigender bzw. in absteigender Reihenfolge sortieren:
     - **Variablenname**
     - **Variablenreihenfolge in Tabelle**
     - **Variablentyp**
     - **Variablenformat**
     - **Variablenetikett**

4 Klicken Sie auf **Fertig stellen**, um die Ausgabe zu generieren.

27

Löschen von Dateien und Formaten

Allgemeines über die Anwendungsroutine 'Dateien und Formate löschen' 183
Löschen von Dateien, Ansichten und benutzerdefinierten Formaten 184
Anzeigen von Eigenschaften 184

Allgemeines über die Anwendungsroutine 'Dateien und Formate löschen'

Mit der Anwendungsroutine **Dateien und Formate löschen** können Sie Dateien, Ansichten und benutzerdefinierte Formate endgültig löschen. Sie können dabei gleichzeitig mehrere Elemente aus verschiedenen Bibliotheken und Katalogen löschen.

Hinweis: Diese Anwendungsroutine ist nur im SAS Enterprise Guide verfügbar.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	CATALOG, SQL
Erforderliche SAS-Produkte	Base SAS

Name der Anforderung	Prozedur- und Produktnamen
Empfohlene zusätzliche SAS-Produkte	Ohne

Löschen von Dateien, Ansichten und benutzerdefinierten Formaten

1 Wählen Sie in der Dropdown-Liste **Server** den Server aus, auf dem die Dateien, Ansichten und benutzerdefinierten Formate gespeichert sind.

2 Markieren Sie im Bereich **Bibliotheken und Kataloge** die Bibliothek bzw. den Katalogl. Die Dateien, Ansichten oder benutzerdefinierten Formate in der gewählten Bibliothek bzw. im gewählten Katalog werden in der Listenansicht angezeigt. Standardmäßig sind diese Elemente nach Name sortiert. Klicken Sie auf die Spaltenüberschrift, um die Elemente nach Änderungsdatum oder Typ zu sortieren.

3 Aktivieren Sie ein oder mehrere Kontrollkästchen für die Dateien, Ansichten oder benutzerdefinierten Formate, die Sie löschen möchten. Hinweis

Hinweis: Unterhalb des Bereichs **Bibliotheken und Kataloge** wird die Gesamtzahl der ausgewählten Dateien, Ansichten und benutzerdefinierten Formate und die Anzahl der ausgewählten Elemente im aktuellen Ordner angezeigt.

4 Wiederholen Sie die Schritte 2 und 3, bis alle Elemente, die Sie löschen möchten, ausgewählt sind.

5 Klicken Sie auf **Ausführen**, um die ausgewählten Elemente endgültig zu löschen.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

28

Diskriminanzanalyse

Allgemeines über die Anwendungsroutine 'Diskriminanzanalyse' ... **187**

Diskriminanzanalyse: Zuweisen von Variablen zu Analyserollen ... **188**

Diskriminanzanalyse: Festlegen der Analyseoptionen ... **189**

Diskriminanzanalyse: Klassifizieren neuer Daten ... **191**
Festlegen der Dateneingabetabelle ... 191
Angeben der Testvariable ... 192

Diskriminanzanalyse: Festlegen von Ergebnisoptionen ... **193**

Festlegen von Titeln und Fußnoten ... **194**

Anzeigen von Eigenschaften ... **194**

Allgemeines über die Anwendungsroutine 'Diskriminanzanalyse'

Für eine Menge an Beobachtungen, die mehrere quantitative Variablen und eine Klassifizierungsvariable enthält, die Beobachtungsgruppen definiert, entwickelt die Anwendungsroutine Diskriminanzanalyse ein Diskriminanzkriterium zur Klassifizierung der einzelnen Beobachtungen in einer der Gruppen. Das aus dieser Datei abgeleitete Diskriminanzkriterium kann während derselben Ausführung der Diskriminanzfunktion auf eine zweite Datei angewendet werden.

Sie können diese Anwendungsroutine z. B. zum Klassifizieren von Beobachtungen in der Natur verwenden. So könnten Sie beispielsweise eine SAS-Datei erstellen, die die Arten, Gewichte und drei verschiedene Längen, Höhen und Breiten von 159 Fischen enthält. Anhand dieser Daten könnten Sie eine Diskriminanzfunktion basierend auf diesen sechs Variablen finden, die jeden Fisch einer der sieben Arten zuweist.

Hinweis: Diese Anwendungsroutine ist nur im SAS Enterprise Guide verfügbar.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	DISCRIM
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Diskriminanzanalyse: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Analysevariablen	führt Analysen für jede Variable in der Liste durch. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Klassifizierungsvariable	definiert die Gruppen für die Analyse. Klassifizierungsausprägungen werden durch die formatierten Werte dieser Variablen bestimmt. Sie müssen dieser Rolle genau eine Variable zuweisen.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert.

Optionsname	Beschreibung
	Hinweis: Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variable gruppieren, die bereits für die Rollen **Analysevariablen** oder **Klassifizierungsvariable** ausgewählt haben.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
ID-Etikett	zeigt die Werte der Klassifizierungsvariablen anstelle der Beobachtungsnummer an. Die Zuweisung einer Variablen zu dieser Rolle ist nur wirksam, wenn Sie das Kontrollkästchen **Ergebnisse der Rückersetzungsklassifizierung** aktiviert haben. Sie können dieser Rolle maximal eine Variable zuweisen.
Relative Gewichtung	verwendet die Werte dieser Variablen, um die entsprechenden Werte der Analysevariablen zu gewichten. Sie können dieser Rolle maximal eine Variable zuweisen.

Diskriminanzanalyse: Festlegen der Analyseoptionen

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Ergebnisse der Kreuzvalidierungsklassifizierung	Klassifiziert jede Beobachtung in der Eingabedatei mithilfe einer Diskriminanzfunktion, die sich aus den anderen Beobachtungen in der Datei, mit Ausnahme der klassifizierten Beobachtung, berechnet. Diese Option wird nur angewendet, wenn eine parametrische Methode (Standard) verwendet wird.
Ergebnisse der Kreuzvalidierungsklassifizierung für	zeigt die Ergebnisse der Kreuzvalidierungsklassifizierung für jede Beobachtung an.

Optionsname	Beschreibung
jede Beobachtung anzeigen	**Hinweis:** Diese Option ist nur verfügbar, wenn Sie das Kontrollkästchen **Zusammenfassende Ergebnisse der Kreuzvalidierungsklassifizierung** aktiviert haben.
Ergebnisse der Rückersetzungsklassifizierung	wählen Sie diese Option, um die Ergebnisse der Resubstitutionsklassifizierung für jede Beobachtung anzuzeigen.
Schätzer für Fehlerrate für Posteriori-Wahrscheinlichkeit	wählen Sie diese Option, um die Fehlerrateschätzungen der Posteriori-Wahrscheinlichkeit für das Klassifizierungskriterium basierend auf den Klassifizierungsergebnissen anzuzeigen.
Univariater Test auf Gleichheit der Klassenmittelwerte	wählen Sie diese Option, um die univariaten Statistiken anzuzeigen, die zum Testen der Hypothese verwendet werden, nach der die Klassenmittelwerte gleich der Grundgesamtheit für jede Variable sind.
Multivariate Tests auf Gleichheit der Klassenmittelwerte	wählen Sie diese Option, um die multivariaten Statistiken anzuzeigen, die zum Testen der Hypothese verwendet werden, nach der die Klassenmittelwerte in der Grundgesamtheit gleich sind.
Ergebnisse der Rückersetzungsklassifizierung	zeigt die Ergebnisse der Rückersetzungsklassifizierung für ausschließlich falsch klassifizierte Beobachtungen an. Diese Option ist nur für gewöhnliche SAS-Dateien verfügbar.
Ergebnisse der Kreuzvalidierungsklassifizierung	zeigt die Ergebnisse der Kreuzvalidierungsklassifizierung ausschließlich für falsch klassifizierte Beobachtungen an.
Kanonische Diskriminanzanalyse	führt eine kanonische Diskriminanzanalyse durch.
Anzahl der zu berechnenden kanonischen Variablen begrenzen	legt die Anzahl der zu berechnenden kanonischen Variablen fest. Die hier angegebene Anzahl muss kleiner gleich der Anzahl der Variablen sein. Wenn Sie 0 angeben, sind in der Ausgabe die kanonischen Korrelationen, nicht aber die kanon. Koeffizienten, Strukturen oder Mittelwerte enthalten. Angenommen *v* ist die Anzahl der Analysevariablen und *c* die Anzahl der Klassifizierungsvariablen. Wenn Sie die Anzahl der zu berechnenden kanonischen Variablen nicht begrenzen, können nur min(*v*,*c* - 1) kanonische Variablen generiert werden. Wenn Sie eine Ausgabedatei zur Ihren Ergebnissen hinzufügen, werden v kanonische Variablen erzeugt. In diesem Fall weisen die letzten *v* - (*c* - 1) kanonischen Variablen fehlende Werte auf.

Optionsname	Beschreibung
Priori-Wahrscheinlichkeiten	gibt die Methode zur Berechnung der Priori-Wahrscheinlichkeiten an. Diese Methoden sind verfügbar: ■ **Gleich** - gibt an, dass die Priori-Wahrscheinlichkeiten der Gruppenzugehörigkeit gleich sein soll ■ **Proportionale Stichprobengrößen** - gibt an, dass die Priori-Wahrscheinlichkeiten der Gruppenzugehörigkeit proportional zur Stichprobengröße sein soll ■ **Wahrscheinlichkeit für jede Ebene eingeben** - Ermöglicht Ihnen, die Priori-Wahrscheinlichkeit für jede Ebene der Klassifizierungsvariablen anzugeben. Wenn die Summe der Priori-Wahrscheinlichkeiten ungleich 1 ist, werden die Wahrscheinlichkeiten proportional auf die Summe 1 skaliert. **Hinweis:** Die Summe der von Ihnen angegebenen Wahrscheinlichkeiten muss 1 betragen. Klicken Sie auf **Wahrscheinlichkeiten zurücksetzen** , um die Wahrscheinlichkeiten auf ihre ursprünglichen Werte zurückzusetzen.

Diskriminanzanalyse: Klassifizieren neuer Daten

Festlegen der Dateneingabetabelle

Klicken Sie im Auswahlfenster auf **Neue Daten klassifizieren**, um auf diese Optionen zuzugreifen.

Sie können eine Tabelle auswählen (hier als "Eingabetabelle" bezeichnet), die Zeilen enthält, die nach dem Diskriminanzkriterium klassifiziert werden sollen, welches von der Tabelle, die Sie analysieren, abgeleitet wird. Die Eingabetabelle muss alle Variablen enthalten, die Sie der Rolle **Analysevariablen** zugewiesen haben.

Standardmäßig werden die Klassifizierungsergebnisse aus der Eingabetabelle in den Ausgabeergebnissen eingeschlossen. Bei der Eingabetabelle handelt es sich um die Eingabedaten, die Sie zum Öffnen der Anwendungsroutine verwendeten.

Um eine andere Eingabetabelle anzugeben, aktivieren Sie das Kontrollkästchen **Neue, zu klassifizierende Daten** und klicken auf **Durchsuchen**.

Wenn Sie der Rolle **Gruppieren nach Variable** eine Variable zugewiesen haben und Sie jede BY-Gruppe in den neuen Daten nach der Diskriminanzfunktion aus der

entsprechenden BY-Gruppe in der ursprünglichen Eingabedatei klassifizieren möchten, müssen die neuen Daten, die Sie wählen, in der gleichen Reihenfolge sortiert sein wie die Eingabedaten, die Sie zum Öffnen der Anwendungsroutine verwendeten.

Hinweis: Sie müssen der Rolle **Klassifizierungsvariable** eine Variable zuweisen, um das Kontrollkästchen **Neue, zu klassifizierende Daten** zu aktivieren.

Hinweis: Sie müssen das Kontrollkästchen **Neue, zu klassifizierende Daten** aktivieren, um eine Eingabetabelle anzugeben, oder um eine Testvariable anzugeben.

Verwenden Sie diese Optionen, um die Eingabetabellenoptionen festzulegen:

- Standardmäßig werden die Klassifizierungsergebnisse aus der Eingabetabelle in den Ausgabeergebnissen eingeschlossen. Deaktivieren Sie das Kontrollkästchen **Klassifizierungsergebnisse für die neuen Daten anzeigen**, um die Klassifizierungsergebnisse für die Eingabetabelle zu unterdrücken.
- Wählen Sie **`Ausgabetabelle für Klassifizierungsergebnisse der neuen Daten`**, um eine permanente Ausgabetabelle zu erstellen, die sämtliche Daten aus der Eingabetabelle enthält, plus die Posteriori-Wahrscheinlichkeiten und die Klasse, in die jede Zeile klassifiziert wird.
- Wählen Sie **Ausgabetabelle für Gruppendichteschätzungen**, um eine permanente Ausgabetabelle zu erstellen, die sämtliche Daten aus der Eingabetabelle enthält, plus die gruppenspezifischen Dichteschätzungen für jede Zeile.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Nachdem Sie eine Eingabetabelle angegeben haben, können Sie eine Testvariable angeben.

Angeben der Testvariable

Klicken Sie im Auswahlfenster auf **Neue Daten klassifizieren**, um auf diese Optionen zuzugreifen.

Nach Angabe einer Eingabetabelle, die Klassifizierungsdaten enthält, können Sie eine Testvariable angeben.

So geben Sie eine Testvariable an:

1 Wählen Sie in der Dropdown-Liste **Klassifizierungsvariable in den neuen Daten** einen Variablennamen aus. Diese Variable wird verwendet, um festzustellen, ob eine Beobachtung in der Eingabetabelle falsch klassifiziert ist. Die Variable, die Sie hier angeben, sollte vom gleichen Typ (alphanumerisch oder numerisch) sein wie die Variable, die der Rolle **Klassifizierungsvariable** zugewiesen ist.

   Wenn Sie hier keine Variable auswählen, wird die Variable verwendet, die der Rolle **Klassifizierungvariable** zugewiesen ist.

   Hinweis: Sie können diese Option verwenden, wenn die der Rolle **Klassifizierungsvariable** zugewiesene Variable in der Eingabetabelle vorhanden ist, aber einen anderen Namen als die ursprüngliche Datei besitzt. Beispielsweise ist die der Rolle 'Klassifizierungsvariable' zugewiesene Variable 'Sex', die in der Eingabetabelle enthaltene Klassifizierungsvariable heißt jedoch 'Gender'.

2 Wählen Sie ggfs.einen Variablennamen aus der Dropdown-Liste für **Häufigkeitszählung für Testvariable**. Die Variable, die Sie hier auswählen, erfüllt die gleiche Funktion für die Testvariable wie die Variable, die der Rolle **Häufigkeitszählung** zugewiesen ist.

3 Wählen Sie ggfs.einen Variablennamen aus der Dropdown-Liste für **ID-Etikett**. Die Variable, die Sie hier auswählen, erfüllt die gleiche Funktion für die Testvariable wie die Variable in der Rolle **ID-Etikett**. Die Auswahlen einer Variable an dieser Stelle ist nur wirksam, wenn das Kontrollkästchen **Klassifizierungsergebnisse für die neuen Daten anzeigen** markiert ist.

Diskriminanzanalyse: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Ausgabedaten speichern	gibt die Ergebnisse an, die in eine Ausgabetabelle aufgenommen werden sollen. Folgende Optionen können ausgewählt werden: ■ Die Tabelle **Posteriori-Wahrscheinlichkeiten & Rückersetzungsklassif.** enthält Ihre Eingabedaten, die Posteriorimöglichkeiten sowie die Klasse, in die jede Beobachtung durch die Rückersetzung eingestuft wird. ■ Die Tabelle **Posteriori-Wahrscheinlichkeiten & Kreuzvalidierungsklassif.** enthält Ihre Eingabedaten, die Posteriorimöglichkeiten sowie die Klasse, in die jede Beobachtung durch die Kreuzvalidierung eingestuft wird.

Optionsname	Beschreibung
	■ Die Tabelle **Beschreibende Statistiken und Diskriminanzfunktionen** enthält verschiedene Statistiken, z.B. Mittelwerte, Standardabweichungen und Korrelationen. ■ Die Tabelle **Spez. Dichteschätzungen für jede Beobachtung gruppieren** enthält Ihre Eingabedaten sowie die gruppenspezifischen Dichteschätzungen für jede Beobachtung. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Berichte unterdrücken	Bewirkt, dass keine Ergebnisse angezeigt werden.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

29

Verteilungsanalyse

Allgemeines über die Anwendungsroutine 'Verteilungsanalyse' ... 197
Verteilungsanalyse: Zuweisen von Variablen zu Analyserollen ... 198
Verteilungsanalyse: Auswählen der Verteilungen ... 199
Verteilungsanalyse: Anpassen der Verteilung ... 200
Verteilungsanalyse: Erzeugen von Plots ... 201
Verteilungsanalyse: Einfügungen hinzufügen ... 203
Verteilungsanalyse: Festlegen von Ergebnisoptionen ... 203
Festlegen von Titeln und Fußnoten ... 204
Anzeigen von Eigenschaften ... 204

Allgemeines über die Anwendungsroutine 'Verteilungsanalyse'

Die Anwendungsroutine **Verteilungsanalyse** stellt Tools für die Datenzusammenfassung sowie Informationen zur Verteilung numerischer Variablen bereit. Sie können diese Anwendungsroutine auch zum Erstellen verschiedenster Plots verwenden, wie z. B. Histogramme, Wahrscheinlichkeitsplots, Quantil-Quantil-Plots und Box-Plots.

Sie können diese Anwendungsroutine auch zum Erstellen der beschreibenden Statistiken für ein Produkt verwenden. Angenommen Sie haben die Beleihungssätze von 5840 Hypotheken in einer SAS-Datei gespeichert. Mithilfe der Anwendungsroutine **Verteilungsanalyse** könnten Sie diese Ergebnisse erhalten:

- Eine Tabelle mit beschreibenden Maßen, einschließlich Momentschätzern, und eine Tabelle mit extremen Beobachtungen
- Ein Histogramm, das Ihnen die Visualisierung der Verteilung der Beleihungssätze ermöglicht. Das Histogramm zeigt die Merkmale der Verteilung auf, wie beispielsweise deren Schiefe und Spitze.

- Eine Analyse der Datenverteilung. Diese Anwendungsroutine ermöglicht Ihnen die Ausführung von Tests für die Normalverteilung und das Erstellen von Diagrammen, wie z. B. einem Wahrscheinlichkeitsplot.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	UNIVARIATE, BOXPLOT
Erforderliche SAS-Produkte	Base SAS, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

Verteilungsanalyse: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Analysevariablen	gibt die zu analysierenden numerischen Variablen an. Eine Analysevariable ist erforderlich. Sie können dieser Rolle mehrere Variablen zuweisen.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Sie können Analysen nicht nach einer Variable gruppieren, die Sie als Analyserolle ausgewählt haben.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt die Variable an, die als relative Gewichtungsvariable verwendet wird. Jeder Wert in der Variable in dieser Rolle wird verwendet, um die entsprechenden Werte der Variablen zu

Optionsname	Beschreibung
	gewichten, die Sie als Analysevariablen definieren. Sie können dieser Rolle nur eine Variable zuweisen.
Klassifizierungsvariablen	gibt bis zu zwei Variablen an, die zum Gruppieren der Daten in Klassifizierungsausprägungen verwendet werden. Wenn Sie ein Histogramm, ein Wahrscheinlichkeitsplot oder einen Quantil-Quantil-Plot erzeugen, werden vergleichende Plots erstellt. Wenn Sie einen Box-Plot erzeugen, werden nebeneinander gestellte Box-Plots erstellt.

Verteilungsanalyse: Auswählen der Verteilungen

Klicken Sie im Auswahlfenster unter der Überschrift 'Verteilungen' auf **Zusammenfassung**, um auf diese Optionen zuzugreifen.

Sie können mehrere theoretische Verteilungen anpassen. Wenn Sie die zu berechnende Verteilung ausgewählt haben, wählen Sie den Namen der Verteilung im Auswahlfenster aus, um die Parameterwerte und Linieneigenschaften für die Verteilung anzugeben.

Verwenden Sie die Kontrollkästchen, um eine oder mehrere der folgenden geschätzten Dichtekurven auszuwählen:

- **Normal** passt eine normale Dichte mithilfe eines Mittelwerts (My) und einer Standardabweichung (Sigma) an.
- **Lognormal** passt eine Lognormal-Dichte mithilfe eines Skalierungsparameters (Zeta), eines Schwellenwertparameters (Theta) und eines Formparameters (Sigma) an.
- **Exponentiell** passt eine exponentielle Dichte mithilfe eines Skalierungsparameters (Sigma) und eines Schwellenwertparameters (Theta) an.
- **Weibull** passt eine Weibull-Dichte mit drei Parametern mit einem Skalierungsparameter (sigma), einem Schwellenwertparameter (theta) und einem Formparameter (c) an.
- **Beta** passt eine Beta-Dichte mit einem Skalierungsparameter (Sigma), einem Schwellenwertparameter (Theta) und Formparametern (Alpha und Beta) an.
- **Gamma** passt eine Gamma-Dichte mit einem Skalierungsparameter (Sigma), einen Schwellenwertparameter (Theta) und einen Formparameter (Alpha) an.
- **Kern** passt die nichtparametrischen Kerndichteschätzungen an.

Wählen Sie den Grafikstil, der in der Ausgabe verwendet werden soll. Sie haben die Auswahl zwischen üblichen SAS-Grafiken und ODS Statistical Graphs (statistischen ODS-Grafiken). Weitere Informationen zu ODS-Grafiken finden Sie unter *SAS Output Delivery System: User's Guide*.

Verteilungsanalyse: Anpassen der Verteilung

Klicken Sie im Auswahlfenster unter der Überschrift **Verteilungen** auf den Namen der Verteilung, die Sie anpassen möchten, um auf diese Optionen zuzugreifen.

Hinweis: Diese Verteilungsoptionen sind nur durch vorherige Auswahl der Verteilung auf der Seite **Zusammenfassung** verfügbar, oder durch Aktivieren des Kontrollkästchens ***Name der Verteilung*** auf der Seite dieser Verteilung.

Für die Normal-, Lognormal-, Exponential-, Weibull-, Beta- und Gamma-Verteilungen können Sie die Verteilungsparameter schätzen lassen oder angeben. Standardmäßig schätzt die Anwendungsroutine die Parameterwerte für alle Analysevariablen.

So legen Sie für jede Analysevariable Parameter fest:

1. Deaktivieren Sie das Kontrollkästchen **Verteilung auf alle Variablen anwenden**.
2. Aktivieren Sie das Kontrollkästchen der Variablen im Feld **Analysevariablen**.
3. Geben Sie die Parameterwerte an.

Hinweis: Um zu den geschätzten Parameterwerten zurückzukehren, klicken Sie auf **Schätzer verwenden**.

Bei nichtparametrischen Kerndichteschätzern können Sie die Bandbreite (c) und die Kernfunktion (k) angeben. Standardmäßig schätzt die Anwendungsroutine eine Bandbreite, die den integrierten Quadratfehler des approximierten Mittelwerts MISE (Mean Integrated Square Error) minimiert. Wenn Sie die Bandbreite nicht schätzen möchten, deaktivieren Sie das Kontrollkästchen **MISE minimieren**. Um den standardisierten Bandbreitenparameter für Kerndichten festzulegen, geben Sie die Bandbreite in das Textfeld ein. Sie können bis zu fünf Werte angeben. Trennen Sie die Werte durch Leerstellen. Geben Sie zum Beispiel folgende Werte in das Textfeld ein: `0,5 1,0 1,5`.

Hinweis: Wenn Sie das Kontrollkästchen **MISE minimieren** nicht deaktivieren, wird MISE als einer der fünf Bandbreitenwerte gezählt.

Sie können auch die Kernfunktion angeben (normal, quadratisch oder dreieckig), die zur Berechnung der Kerndichteschätzung verwendet wird. Die Standardkernfunktion ist **Normal**.

Sie können Kernfunktionen für bis zu fünf Schätzungen angeben. Wenn Sie mehr als eine Kernfunktion angeben, ist die Reihenfolge der Funktionen immer: normal - quadratisch - dreieckig. Um Funktionen hinzuzufügen oder die Reihenfolge zu ändern, müssen Sie den Code der Anwendungsroutine ändern.

Wenn Sie mehr Kernfunktionen als Bandbreiten angeben, wird die letzte Bandbreite für die restlichen Schätzungen wiederholt. Wenn Sie z.B. 0,5 und 1,0 als Bandbreiten und normal, quadratisch und dreieckig als Kernfunktionen angeben, wird die Bandbreite 1,0 für die dreieckige Funktion wiederholt.

Wenn Sie mehr Bandbreiten als Kernfunktionen angeben, wird die letzte Kernfunktion für die restlichen Bandbreiten wiederholt. Wenn Sie z.B. 0,5, 1,0 und 1,5 als Bandbreiten und normal und quadratisch als Kernfunktionen angeben, wird die quadratische Funktion für die Bandbreite 1,5 verwendet.

Verteilungsanalyse: Erzeugen von Plots

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Plots**, um auf diese Optionen zuzugreifen.

Hinweis: Wenn einer Analyserolle eine relative Gewichtungsvariable zugewiesen ist, können Sie keine Plots erstellen.

Optionsname	Beschreibung
Histogrammplot	erstellt ein Histogramm und überlagert optional Dichtekurven für kontinuierliche theoretische Verteilungen und für Kerndichteschätzungen. Um eine Verteilung auszuwählen, klicken Sie im Auswahlfenster auf **Zusammenfassung**. Wenn Sie Klassifizierungsvariablen zuweisen, wird ein Vergleichshistogramm erstellt, das aus den Komponentenhistogrammen besteht. Die Komponentenhistogramme ermöglichen den Vergleich der Verteilung einer Analysevariablen über die Ausprägungen der Klassifizierungsvariablen hinweg.
Wahrscheinlichkeitsplot	erstellt ein Wahrscheinlichkeitsplot für eine stetige theoretische Verteilung und eine Referenzlinie, die entweder dem angegebenen oder dem geschätzten Lageparameter und Skalierungsparameter für die theoretische Verteilung entspricht. Ein Wahrscheinlichkeitsplot wird verwendet, um die geordneten Werte einer Variablen mit den Perzentilen einer angegebenen theoretischen Verteilung zu vergleichen. Wenn

Optionsname	Beschreibung
	die Datenverteilung der theoretischen Verteilung entspricht, bilden die Punkte im Plot ein lineares Muster. Wenn Sie Klassifizierungsvariablen zuweisen, wird ein vergleichendes Wahrscheinlichkeitsplot erstellt, das aus Komponentenwahrscheinlichkeitsplots besteht. Wenn Sie keine Verteilung auswählen, wird die Normalverteilung verwendet.
Quantil-Plot	erstellt einen Quantil-Quantil-Plot (Q-Q-Plot) für eine stetige theoretische Verteilung und eine Referenzlinie, die entweder dem angegebenen oder dem geschätzten Lageparameter und Skalierungsparameter für die theoretische Verteilung entspricht. Der Q-Q-Plot wird verwendet, um die geordneten Werte einer Variablen mit den Quantilen einer angegebenen theoretischen Verteilung zu vergleichen. Wenn die Datenverteilung der theoretischen Verteilung entspricht, bilden die Punkte im Plot ein lineares Muster. Wenn Sie Klassifizierungsvariablen zuweisen, wird ein vergleichender Quantil-Quantil-Plot erstellt, der aus Komponenten-Q-Q-Plots besteht. Wenn Sie keine Verteilung auswählen, wird die Normalverteilung verwendet. **Hinweis:** Q-Q-Plots ähneln Wahrscheinlichkeitsplots. Q-Q-Plots sind bei der grafischen Schätzung von Verteilungsparametern vorzuziehen, wohingegen Wahrscheinlichkeitsplots bei der grafischen Schätzung von Perzentilen geeigneter sind.
Box-Plot	erstellt einen Box-Plot. Wenn Sie Klassifizierungsvariablen zuweisen, werden nebeneinander angeordnete Box-Plots erstellt. Ein Box-Plot zeigt den Mittelwert, die Quartile, die kleinste Beobachtung und die größte Beobachtung für eine Gruppe von Beobachtungen.
Textbasierte Plots	erstellt die folgenden Plots: ■ einen Stem-and-Leaf-Plot oder ein Balkendiagramm (je nach der Anzahl der Beobachtungen) ■ einen Box-Plot ■ einen normalen Wahrscheinlichkeitsplot ■ einen vergleichenden Plot, wenn der Rolle **Analyse gruppieren nach** eine Variable zugewiesen wurde
Achsenfarbe, **Hintergrundfarbe**, **Achsenbreite**	passt die Darstellung des Plots an. **Hinweis:** Für den Stamm-und-Blatt-Plot sind keine Optionen verfügbar, da es sich um einen einfachen Textplot handelt.

Verteilungsanalyse: Einfügungen hinzufügen

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Einfügung**, um auf diese Optionen zuzugreifen.

Sie können einen Plot erweitern, indem Sie direkt in der Grafik ein Feld oder eine Tabelle (eine so genannte Einfügung) der beschreibenden Statistiken hinzufügen. Wenn Sie ein Vergleichshistogramm erstellen, erscheint in jedem Komponentenhistogramm eine Einfügung.

Verteilungsanalyse: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

Standardmäßig enthält der Verteilungsanalysebericht die folgenden Tabellen: eine mit Grundkonfidenzintervallen, eine mit Grundmaßen, eine mit Tests auf Lageparametern und eine mit Momenten.

Optionsname	Beschreibung
Keine Anzeige von beschreibenden Statistiken und Prozessfähigkeitsindizes	unterdrückt alle beschreibenden Statistiken und Prozessfähigkeitsindextabellen in der Ausgabe. Folgende Tabellen werden durch Auswahl dieser Option nicht unterdrückt: ■ Tabellen, die von der ausgewählten Verteilung erstellt wurden. Sie können diese Tabellen unterdrücken, indem Sie im Fenster für diese Verteilung das Kontrollkästchen **Verteilungstabellen unterdrücken** aktivieren. ■ Tabellen, die mit der INTERVALS-Anweisung erstellt wurden. Um diese Tabellen zu unterdrücken, können Sie die Option NOPRINT angeben. Weitere Informationen zur Anweisung INTERVALS finden Sie in der Hilfe zur Prozedur CAPABILITY.

Optionsname	Beschreibung
Ausgabestatistiken in einer Datei speichern	erzeugt eine Ausgabedatei. Die in der endgültigen Ausgabe enthaltenen Statistiken sind von den Tabellen abhängig, die im Feld **Tabellen** ausgewählt wurden. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die

Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

30

Ringdiagramm

Allgemeines über die Anwendungsroutine 'Ringdiagramm' ... **207**

Ringdiagramm: Auswählen des Diagrammtyps ... **208**

Ringdiagramm: Zuweisen von Variablen zu Analyserollen ... **208**

Ringdiagramm: Festlegen von Darstellungsoptionen ... **209**
- Festlegen der Optionen für Ringdiagramme ... 209
- Festlegen des Layouts ... 211
- Festlegen erweiterter Optionen ... 211
- Weitere Darstellungsoptionen ... 212

Festlegen von Titeln und Fußnoten ... **212**

Anzeigen von Eigenschaften ... **212**

Allgemeines über die Anwendungsroutine 'Ringdiagramm'

Die Anwendungsroutine **Ringdiagramm** erstellt einfache oder Gruppendiagramme, die den relativen Beitrag der Teile zum Ganzen darstellen. Die Daten erscheinen als keilförmige "Segmente" eines Kreises. Jedes Segment stellt eine Datenkategorie dar. Die Segmentgröße stellt den Beitrag der Daten zur Gesamtdiagrammstatistik dar. Beispielsweise kann ein Ringdiagramm die Verkaufszahlen jedes Ladens als Bruchteil der Gesamtverkaufszahlen der Ladenkette darstellen. Ein Ringdiagramm ähnelt einem Kreisdiagramm, mit dem Unterschied, dass ein Ringdiagramm in der Mitte ein Loch aufweist.

Weitere Informationen, siehe “Auswählen des Grafikausgabeformats” auf Seite 6.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GCHART
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Ringdiagramm: Auswählen des Diagrammtyps

Klicken Sie im Auswahlfenster auf **Ringdiagramm**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Einfacher Ring	erstellt ein Diagramm, das den relativen Beitrag der Teile zum Ganzen aufzeigt. Die Daten erscheinen als keilförmige "Segmente" eines Kreises. Jedes Segment stellt eine Datenkategorie dar. Die Segmentgröße stellt den Beitrag der Daten zur Gesamtdiagrammstatistik dar.
Gruppenring	erstellt ein separates Ringdiagramm für jeden eindeutigen Wert der Variablen, die der Rolle **Ringe gruppieren nach** zugewiesen wurde.

Ringdiagramm: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im

Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Hinweis: Die jeweils verfügbaren Rollen hängen vom ausgewählten Diagrammtyp ab.

Optionsname	Beschreibung
Spalte für Diagramm	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die verschiedenen Segmente in Ringdiagrammen. Die Spalte kann alphanumerisch oder numerisch sein.
Ringe gruppieren nach	Für jeden eindeutigen Wert der Spalte, die Sie dieser Rolle zuweisen, wird ein Ringdiagramm erstellt. Alle Diagramme werden auf der gleichen Seite angezeigt.
Summe von	Die Spalte, die Sie dieser Rolle zuweisen, bestimmt die Größen der Abschnitte im Ringdiagramm. Wenn Sie dieser Rolle keine Spalte zuweisen, wird die Größe der Abschnitte über die Häufigkeit jedes Werts der Spalte **Spalte für Diagramm** verwendet.
Analyse gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Diagramme, die erstellt werden. Für jeden eindeutigen Wert wird eine separate Grafik erzeugt. Jedes Ringdiagramm wird auf einer neuen Seite angezeigt.

Ringdiagramm: Festlegen von Darstellungsoptionen

Festlegen der Optionen für Ringdiagramme

Klicken Sie im Auswahlfenster unter der Überschrift "Darstellung" auf **Ringe**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Farben angeben	gibt die Farbe der Segmente an. Folgende Optionen können ausgewählt werden: ■ **Standardfarbschema** - Verwendet für die Segmente das Standardfarbschema. Das Standardfarbschema wird durch den aktuellen Stil bestimmt. Sie können den Stil mit Hilfe des Stil-Managers oder des Dialogfelds **Eigenschaften** für das Ergebnis festlegen. In SAS Enterprise Guide können Sie im Dialogfeld **Optionen** auch den Stil für jeden Ergebnistyp festlegen. Im SAS Add-In for Microsoft Office können Sie im Dialogfeld **SAS Add-In - Optionen** Stile auf die SAS Report- und HTML-Formate anwenden. ■ **Existierende Farben (in einer früheren Version angegeben)** ermöglicht das Verwenden von Farben, die in einer früheren Version von SAS Enterprise Guide oder SAS Add-In for Microsoft Office festgelegt wurden. Diese Option ist nur für SAS-Inhalte verfügbar, die Sie von einer früheren Version migriert haben. ■ **Benutzerdefinierte Farben** - Ermöglicht die Angabe benutzerdefinierter Farben für die Segmente. Mit den Dropdown-Listen können Sie bis zu 12 Farben für das Balkendiagramm bestimmen. wenn das Diagramm mehr als 12 Farben erfordert, werden die zusätzlichen Farben aus dem aktuellen Standardfarbschema ausgewählt. Klicken Sie auf **Zurücksetzen**, um die benutzerdefinierten Farben wieder auf den Standard zurückzusetzen. **Hinweis:** Je nach Ringdiagrammtyp werden eventuell nicht alle benutzerdefinierten Farben verwendet.
Anzahl der Segmente angeben	gibt die Anzahl der Segmente im Diagramm an. Standardmäßig wird die Anzahl der Segmente automatisch bestimmt. Bei stetigen Spalten werden die Werte in Bereiche geteilt, wobei für jeden Bereich ein Segment zur Verfügung steht. Segmente, die 4 Prozent oder weniger des Gesamtbereichs darstellen würden, werden in einem Segment mit der Bezeichnung "Sonstige" zusammengefasst. Um die Anzahl der Segmente anzugeben, die im Diagramm angezeigt werden sollen, aktivieren Sie das Kontrollkästchen **Anzahl der Segmente angeben**.

Festlegen des Layouts

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Layout**, um auf diese Optionen zuzugreifen.

Wählen Sie aus, ob im Etikett jedes Segments der Namen, der Prozentwert, der Statistikwert oder eine Kombination aus allem angezeigt werden soll. Wählen Sie für jede Komponente des Etiketts eine der folgenden Positionen:

- im Segment
- außerhalb des Kreises oder Rings
- außerhalb des Segments mit einem darauf zeigenden Pfeil

Sie können für die Schrift der Segmentetiketten und für die Segmente die gleiche Farbe wählen. Diese Option sollten Sie nicht verwenden, wenn Sie die Etiketten innerhalb der Segmente selbst anzeigen möchten.

Sie können den Prozentwert und das Etikett für das Segment "Sonstige" angeben. Dieses Segment ist eine Sammlung aller Midpoints, deren statistische Werte des Diagramms kleiner oder gleich dem Gesamtprozentwert sind. Der Standardwert für den prozentualen Anteil an der Summe beträgt 4. Das bedeutet, dass jedes Segment, das 4 Prozent oder weniger von der Gesamtprozentzahl enstpricht, im Segment "Sonstige" dargestellt wird.

Sie können für das Etikett des Segments "Sonstige" einen Text mit max. 16 Zeichen eingeben. Das Standardetikett lautet "Sonstige". Das Segment "Sonstige" ist das letzte Kreissegment, unabhängig von der Reihenfolge der Segmente.

Wenn nur ein Midpoint in die Kategorie "Sonstige" fällt, bleibt das Segment für diesen Midpoint an seiner ursprünglichen Position im Kreis und behält sein ursprüngliches Etikett bei. Angenommen ein Kreis zeigt die Getränkeumsätze in Prozentwerten und enthält die folgenden Segmente und Prozentwerte: Kaffee 35%, Tee 15%, Mineralwasser 5% und Milch 45%. Wenn der Gesamtprozentwert 5 beträgt, bleibt das Segment für Wasser das dritte Segment und wird nicht zum letzten Segment, und dieses Segment bekommt das Etikett "Wasser" und nicht "Sonstige".

Verwenden Sie bei gruppierten Diagrammen den Bereich **Gruppiertes Layout**, um die Diagramme anzuordnen. Geben Sie die Anzahl der Spalten oder Zeilen im Ausgabebereich an.

Festlegen erweiterter Optionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Erweitert**, um auf diese Optionen zuzugreifen.

Wählen Sie aus der Dropdown-Liste die Statistik, mit der die Diagrammsegmente berechnet werden sollen. Wenn Sie der Rolle **Summe von** keine Variable zuweisen,

können Sie die Häufigkeit oder den Prozentwert für jedes Segment berechnen. Standardmäßig wird die Häufigkeit berechnet. Wenn Sie der Rolle **Summe von** eine Variable zuweisen, können Sie die Summe oder den Mittelwert für jedes Segment berechnen. Standardmäßig wird die Summe berechnet.

Weitere Darstellungsoptionen

Einige Darstellungsoptionen gelten für alle Grafik-Anwendungsroutinen. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Festlegen von Legendenoptionen” auf Seite 749
- “Anpassen des Diagrammbereichs” auf Seite 750

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

31

Herunterladen von Dateien auf den PC

Allgemeines über die Anwendungsroutine 'Dateien auf PC herunterladen' 215
Dateien auf PC herunterladen: Wählen Sie die herunterzuladenden SAS-Dateien aus . 216
Dateien auf PC herunterladen: Angeben der Optionen zum Herunterladen 216

Allgemeines über die Anwendungsroutine 'Dateien auf PC herunterladen'

Mit der Anwendungsroutine **Dateien auf PC herunterladen** können Sie eine oder mehrere SAS-Dateien (*.sas7dbat) aus einem oder mehreren SAS-Servern auswählen und sie in Ihr lokales Microsoft Windows-Dateisystem kopieren.

Die Anwendungsroutine erstellt ein Log, das eine Zusammenfassung folgender Informationen enthält: die kopierten Datendateien, die Größe in Byte, Zeiten für jede Datei und für alle Dateien als Gruppe sowie Fehler, die während der Übertragung aufgetreten sind.

Dateien auf PC herunterladen: Wählen Sie die herunterzuladenden SAS-Dateien aus

1 Klicken Sie auf **Hinzufügen**, um die SAS-Dateien auszuwählen, die Sie kopieren möchten. Das Dialogfeld **Öffnen** wird geöffnet.

2 Navigieren Sie zum entsprechenden Server und zur Bibliothek und wählen Sie den Namen der zu kopierenden Datei aus. Um mehr als eine Datei auszuwählen, wählen Sie bei gedrückt gehaltener Strg-Taste die Dateien aus, die hinzugefügt werden sollen.

3 Klicken Sie auf **Öffnen**, um die Dateien in die Liste der zu kopierenden Dateien aufzunehmen.

Hinweis: Sie können Dateien von mehreren SAS-Servern auswählen, indem Sie nochmals auf **Hinzufügen** klicken und die Daten von einem anderen Server auswählen.

Hinweis: Alle kopierten Dateien erhalten im Zielordner Namen in Kleinbuchstaben.

4 Klicken Sie auf **Weiter**, um die Optionen zum Herunterladen festzulegen.

Dateien auf PC herunterladen: Angeben der Optionen zum Herunterladen

1 Bevor Sie die Optionen zum Herunterladen angeben können, müssen Sie zuerst die herunterzuladenden Dateien auswählen.

2 Geben Sie den Pfad des Ordners ein, in den Sie die Dateien kopieren möchten. Der Standardspeicherort ist der Ordner "Eigene Dokumente".

3 Um auf Ihrem lokalen Computer bereits vorhandene Dateien mit demselben Namen wie die zu kopierenden Dateien zu überschreiben, aktivieren Sie das

Kontrollkästchen **Vorhandene Dateien mit demselben Namen überschreiben**. Standardmäßig werden vorhandene Dateien nicht überschrieben.

4 Klicken Sie auf **Fertig stellen**, um die Dateien zu kopieren.

Hinweis: Wenn es sich bei der Quellbibliothek um eine Base SAS Engine-Bibliothek mit einem einzelnen Dateipfad handelt, erfolgt die Dateiübertragung direkt von diesem Bibliothekspfad in den Zielordner. Ist die Zielbibliothek keine Base SAS Engine-Bibliothek oder handlet es sich um eine verkettete Bibliothek, werden die Dateien zuerst in eine Staging Area in WORK übertragen und dann mit PROC SQL in den Zielordner kopiert.

32

Faktoranalyse

Allgemeines über die Anwendungsroutine 'Faktoranalyse' 219
Faktoranalyse: Zuweisen von Variablen zu Analyserollen 220
Faktoranalyse: Festlegen von Optionen für Faktormethoden 221
Faktoranalyse: Festlegen von Kommunalitätsoptionen 223
Faktoranalyse: Festlegen von Rotations- und Plotoptionen 225
Faktoranalyse: Festlegen von Ergebnisoptionen 227
Festlegen von Titeln und Fußnoten 229

Allgemeines über die Anwendungsroutine 'Faktoranalyse'

Die Faktoranalyse führt eine Vielzahl allgemeiner Faktor- und Komponentenanalysen und Rotationen durch. Als Eingabe können multivariate Daten, eine Korrelationsmatrix, eine Kovarianzmatrix, ein Faktormuster oder eine Matrix mit Scoring-Koeffizienten verwendet werden.

Sie können diese Anwendungsroutine verwenden, um eine allgemeine Faktoranalyse für eine jährliche Angestelltenüberprüfung durchzuführen. Angenommen, es wurden 103 Polizeibeamte von ihren Vorgesetzten auf 14 Skalen (Variablen) bewertet. Anhand dieser Daten können Sie eine allgemeine Faktoranalyse für diese Variablen durchführen, um die latenten Faktoren hinter diesen Bewertungen zu erkennen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	FACTOR

Name der Anforderung	Prozedur- und Produktnamen
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	Ohne

Faktoranalyse: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Analysevariablen	gibt die zu analysierenden Variablen an. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Partielle Variablen	gibt die Variable an, mit der die zu analysierenden Variablen aussortiert werden. Auf diese Weise basiert die Analyse auf einer partiellen Korrelations- oder Kovarianzmatrix.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können die Tabelle nicht nach einer Variablen gruppieren, die bereits für **Analysevariablen** ausgewählt ist.
Häufigkeitszählung	gibt eine Variable an, welche die Häufigkeit für die anderen Werte innerhalb der Beobachtung darstellt. Dies führt dazu, dass die Anwendungsroutine jede Zeile so behandelt, als

Optionsname	Beschreibung
	würde sie n-mal vorkommen. Dabei entspricht n dem Wert der Häufigkeitszählung für diese Beobachtung. Als Gesamtanzahl der Beobachtungen wird ein Wert angenommen, der der Summe der Variablen entspricht, wenn die Freiheitsgrade für Signifikanzwahrscheinlichkeiten bestimmt werden. Wenn der Wert der Häufigkeitsvariablen fehlt oder kleiner als eins ist, wird die Beobachtung nicht in der Analyse verwendet. Wenn der Wert keine Ganzzahl ist, wird nur der ganzzahlige Teil verwendet.
Relative Gewichtung	gibt eine Variable an, die als relative Gewichtung für jede Beobachtung in der Eingabedatei verwendet wird. Dies ist häufig der Fall, wenn sich die Varianz, die den einzelnen Beobachtungen zugewiesen wurde, unterscheidet, und die Werte der relativen Gewichtungsvariablen proportional zu den Reziprokwerten der Varianzen sind. Wenn ein relativer Gewichtungswert negativ ist oder fehlt, wird er aus der Analyse ausgeschlossen.

Faktoranalyse: Festlegen von Optionen für Faktormethoden

Klicken Sie im Auswahlfenster auf **Faktormethode**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Faktormethode	gibt die Faktormethode an. Folgende Optionen können ausgewählt werden: ■ Die Hauptkomponentenanalyse führt eine Hauptkomponentenanalyse durch, wenn alle Priori gleich eins sind. Andernfalls wird eine Hauptfaktoranalyse durchgeführt. Diese Methode ist der Standard, außer für Dateien des Typs FACTOR. ■ Mit der Option **Harris-Komponentenanalyse** wird eine Harris-Komponentenanalyse der Matrix S-1RS-1 durchgeführt (eine nicht iterative Annäherung an die kanonische Komponentenanalyse).

Optionsname	Beschreibung
	■ **Bildkovarianzmatrix** führt eine Hauptkomponentenanalyse der Bildkovarianzmatrix durch. Es ist eine Korrelationsmatrix ohne Singularität erforderlich. ■ **Alphafaktoranalyse** führt eine Alphafaktoranalyse durch. ■ **Maximum-Likelihood-Faktoranalyse** führt eine Maximum-Likelihood-Faktoranalyse mit einem Algorithmus nach Fuller (mit Ausnahme kleinerer Details) durch. Es ist eine Korrelationsmatrix ohne Singularität erforderlich. ■ **Iterierte Hauptfaktoranalyse** führt eine iterierte Hauptfaktoranalyse durch. ■ **Ungewichtete Kleinste-Quadrate-Faktoranalyse** führt eine ungewichtete Kleinste-Quadrate-Faktoranalyse durch. ■ **Bestehendes Faktormuster** liest ein Faktormuster aus einer Datei des Typs FACTOR, CORR, UCORR, COV oder UCOV. Diese Methode ist der Standard für Dateien des Typs FACTOR. ■ **Faktor-Scoring-Koeffizienten** liest die Scoring-Koeffizienten aus einem der Dateitypen FACTOR, CORR, UCORR, COV oder UCOV. Die Datei muss darüber hinaus eine Korrelations- oder Kovarianzmatrix enthalten. Nach Auswahl der Methode können Sie die Optionen **Numerische Eigenschaften** und **Anzahl der Faktoren** in der Tabelle festlegen. Wenn Sie zwei oder mehr der Optionen **Anzahl der Faktoren** ändern, ist die verbleibende Anzahl an Faktoren mindestens gleich der Anzahl für die ein beliebiges der folgenden Kriterien erfüllt ist: ■ Kleinster Eigenwert, für den ein Faktor verwendet wird ■ Anzahl der verwendeten Faktoren ■ Anteil der allgemeinen Varianz, die von verwendeten Faktoren erklärt wird
Matrixoptionen	legt die Matrixoptionen fest ■ **Faktorkovarianzmatrix statt Korrelationsmatrix** verwendet die Kovarianzmatrix anstelle der Korrelationsmatrix für Faktoren, wenn die Faktormethode die Hauptkomponentenanalyse, Bildkovarianzmatrix, iterierte Hauptfaktoranalyse oder ungewichtete Kleinste-Quadrate-Faktoranalyse ist. ■ **Gewichtete Matrix faktorisieren** verwendet eine gewichtete Korrelations- oder Kovarianzmatrix für Faktoren, wenn die Faktormethode die

Optionsname	Beschreibung
	Hauptkomponentenanalyse, Bildkovarianzmatrix, iterierte Hauptfaktoranalyse oder ungewichtete Kleinste-Quadrate-Faktoranalyse ist. **Hinweis:** Diese Option ist nur verfügbar, wenn die Faktormethoden "Hauptkomponentenanalyse", "Bildkovarianzmatrix", "Iterierte Hauptfaktoranalyse" oder "Ungewichtete Kleinste-Quadrate-Faktoranalyse" verwendet werden und die Eingabedatei vom Typ CORR, UCORR, COV, UCOV oder FACTOR ist. ■ **Resultierende Matrix auf Mittelwert korrigieren** verwendet die Konstante für die Analyse. Kovarianzen oder Korrelationen werden für den Mittelwert korrigiert. Diese Option ist standardmäßig ausgewählt.
Teiler für Varianz	gibt den Teiler an, der bei der Berechnung von Varianzen und Kovarianzen verwendet wird. Standardmäßig entspricht die Berechnung des Teilers für Varianz den Freiheitsgraden.

Faktoranalyse: Festlegen von Kommunalitätsoptionen

Klicken Sie im Auswahlfenster auf **Kommunalität**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Priori-Kommunalitätsschätzer	Wählen Sie eine dieser Optionen aus: ■ **Proportional zu quadrierten multiplen Korrelationen** legt die Priori-Kommunalitätsschätzer proportional zu quadrierten multiplen Korrelationen fest, passt diese jedoch so an, dass ihre Summe gleich der maximalen absoluten Korrelation ist. ■ **Quadrierte multiple Korrelation mit allen anderen Spalten** definiert den Priori-Kommunalitätsschätzer für jede Variable als quadrierte multiple Korrelation mit allen anderen Variablen. ■ **Maximale absolute Korrelation mit anderen Spalten** definiert den Priori-Kommunalitätsschätzer für jede Variable als maximale Absolutkorrelation mit allen anderen Variablen.

Optionsname	Beschreibung
	■ **Alle Priori-Wahrscheinlichkeiten auf eins setzen** legt für alle Priori-Kommunalitäten den Wert 1,0 fest. ■ **Eingabe aus PRIORS-/COMMUNAL-Zeile in Eingabetabelle** liest die Priori-Kommunalitätsschätzer aus der ersten Beobachtung, wenn die Variable _TYPE_ die Werte PRIORS oder COMMUNAL enthält und als Dateityp FACTOR angegeben wurde. **Hinweis:** Diese Option ist nur verfügbar, wenn die Faktormethoden "Hauptkomponentenanalyse", "Alphafaktoranalyse", "Maximum-Likelihood-Faktoranalyse", "Iterierte Hauptfaktoranalyse" oder "Ungewichtete Kleinste-Quadrate-Faktoranalyse" verwendet werden und die Eingabedatei vom Typ FACTOR ist. ■ **Priori-Wahrscheinlichkeiten auf Zufallszahlen zwischen 0 und 1** legt für die Priori-Kommunalitätsschätzer Priori-Wahrscheinlichkeiten auf Zufallszahlen fest, die gleichmäßig zwischen 0 und 1 verteilt sind. ■ **Priori-Werte für jede Spalte explizit angeben** ermöglicht Ihnen das Auswählen des Priori-Kommunalitätsschätzers für jede Variable aus dem Eingabefeld sowie die anschließende Angabe eines neuen Werts. Standardmäßig lautet der Priori-Kommunalitätsschätzer für jede Variable `0`. **Hinweis:** Die voreingestellten Priori-Kommunalitätsschätzer hängen von der Faktormethode ab, die Sie verwenden.
Heywood-Optionen	Standardmäßig beenden die Faktorverfahren Alpha-Faktoranalyse, Maximum-Likelihood-Faktoranalyse, Iterierte Hauptfaktoranalyse und Ungewichtete Kleinste-Quadrate-Faktoranalyse die Iteration und legen die Zahl der Faktoren auf 0 fest, sobald eine geschätzte Kommunalität 1 überschreitet. Außerdem können folgende Optionen ausgewählt werden: ■ **Kommunalitäten über eins auf eins setzen** legt für alle Kommunalitäten größer als 1 den Wert 1 fest, sodass die Iterationen fortgesetzt werden können. ■ **Kommunalitäten können eins überschreiten** ermöglicht, dass Iterationen fortgesetzt werden, wenn die Kommunalitäten den Wert 1 überschreiten. Hierbei können Konvergenzprobleme auftreten, da die Kommunalitäten sehr groß werden und schlecht konditionierte Hessesche Matrizen auftreten können.

Faktoranalyse: Festlegen von Rotations- und Plotoptionen

Klicken Sie im Auswahlfenster auf **Rotation und Plots**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Rotationsmethode	legt die Rotationsmethode fest. Folgende Methoden können ausgewählt werden. Je nach ausgewählter Methode, sind weitere Optionen verfügbar. ■ **Keine Rotation** gibt an, dass keine Rotation durchgeführt und die ursprüngliche orthogonale Lösung belassen wird. Dies ist die Standardmethode. ■ **Orthogonales Equamax** gibt die Orthogonal-Equamax-Rotation an. Dies entspricht der Verwendung eines allgemeinen Orthomaxes mit einer Orthomax-Gewichtung von *n*/2, wobei *n* der Anzahl der Faktoren entspricht. ■ **Harris-Kaiser Fall II Orthoblique** gibt die Harris-Kaiser Fall II Orthoblique-Rotation an. Sie können die Potenz der Quadratwurzeln der Eigenwerte angeben, die zum erneuten Skalieren der Eigenvektoren verwendet werden. Dabei wird davon ausgegangen, dass die Faktoren durch die Hauptfaktormethode extrahiert werden. Bei anderen Extraktionsmethoden werden die Eigenvektoren durch die normalisierten Spalten der nicht gedrehten Faktormatrix ersetzt, während die Eigenwerte durch die Konstanten zur Spaltennormalisierung ersetzt werden. Der Standardeigenwert von Null führt zur unabhängigen Clusterlösung, in der alle Variablen dazu tendieren, nur einen Faktor stark zu belasten. ■ **Allgemeines Orthomax** gibt die Orthomax-Rotation mit einer Standard-Orthomax-Gewichtung von null an. Hinsichtlich der gültigen Werte für die Orthomax-Gewichtung gelten keine Einschränkungen, doch meist liegen die gängigen Werte zwischen null und der Anzahl der Variablen. Eine Orthomax-Gewichtung entspricht einem der Ergebnisse in der Varimax-Rotation. ■ **Orthogonales Parsimax** gibt die Orthogonal-Parsimax-Rotation an. Dies entspricht der Verwendung eines allgemeinen Orthomax mit einer Orthomax-Gewichtung

Optionsname	Beschreibung
	von [(*nvar* ×(*n* - 1))/(*nvar* + *n* - 2)], wobei *nvar* die Anzahl der Variablen und *n* die Anzahl der Faktoren ist. ■ **Oblique Procrustes** gibt eine Oblique-Procrustes-Rotation an, deren Zielmuster durch die Datei TARGET zur Verfügung gestellt wird. Es wird die unbeschränkte Methode kleinster Quadrate verwendet, wobei die Faktoren nach der Rotation zu einer Einheitenvarianz skaliert werden. ■ **Oblique Promax** gibt die Oblique-Promax-Rotation an. Sie können die gewünschte Prärotationsmethode mit orthogonal oder oblique angeben. Außerdem können Sie die Potenz ändern, die zum Berechnen des Zielmusters aus dem Standardwert 3 verwendet wird. ■ **Orthogonal Quartimax** gibt die Orthogonal-Quartimax-Rotation an. Dies entspricht der Verwendung eines allgemeinen Orthimaxes mit einer Orthomax-Gewichtung von 0. ■ **Orthogonal Varimax** gibt die Orthogonal-Varimax-Rotation an. Dies entspricht der Verwendung eines allgemeinen Orthimaxes mit einer Orthomax-Gewichtung von 1.
Normalisierungsmethode für Faktormusterzeilen	gibt die Methode für die Normalisierung der Zeilen an. Folgende Optionen können ausgewählt werden: ■ **Kaisers Normalisierung** verwendet Kaisers Normalisierung , welche die Standardmethode ist. ■ **Nach Cureton-Mulaik-Technik gewichten** gewichtet die Zeilen mithilfe der Cureton-Mulaik-Technik. ■ **Neu skalieren, um Kovarianzen darzustellen** skaliert die Zeilen der Mustermatrix neu, um die Kovarianzen anstelle der Korrelationen darzustellen. ■ **Nicht normalisieren** verhindert die Normalisierung.
Anzuzeigende Plots	gibt die Plots an, die in den Ergebnisse enthalten sein sollen. Folgende Optionen können ausgewählt werden: ■ **Faktormuster darstellen (nach Rotation)** zeigt einen Plot des Faktormusters nach der Rotation an. Über die Option **Darzustellende Faktoren** können Sie die Zahl, *n*, für die darzustellenden Faktoren angeben, sodass alle Paare der ersten *n* Faktoren dargestellt und insgesamt *n*(*n* - 1)/2 Plots erstellt werden. Geben Sie direkt einen ganzzahligen Wert ein, oder passen Sie den Wert mithilfe der Pfeiltasten um eins nach oben oder unten an. Der Standardwert beträgt 2. ■ Plot des Faktormusters vor der Rotation zeigt einen Plot des Faktormusters vor der Rotation an.

Optionsname	Beschreibung
	■ Scree-Plot der Eigenwerte anzeigen zeigt einen Scree-Plot der Eigenwerte an. Standardmäßig werden keine Plots angezeigt. Nach Auswahl der anzuzeigenden Plots können Sie auch die Anzahl der darzustellenden Faktoren angeben. Folgende Optionen können ausgewählt werden: ■ **Alle Faktoren verwenden** stellt alle Faktoren dar. Hierbei handelt es sich um die Standardeinstellung. ■ **Anzahl darzustellende Faktoren** ermöglicht Ihnen anzugeben, wieviele Faktoren dargestellt werden sollen (*n*). Der kleinste Wert, den Sie angeben können, ist 2. Wenn Sie einen Wert angeben, werden alle Paare der ersten n Faktoren werden dargestellt. Die gesamte Anzahl der Plots ergibt sich aus $n(n - 1)/2$.

Faktoranalyse: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Ausgabedaten speichern	Markieren Sie im Bereich **Ausgabedaten speichern** die Kontrollkästchen **Faktoren** oder **Statistiken** für die Ausgabedatei, die Sie permanent speichern möchten. Wenn Sie **Statistiken** auswählen, können Sie die Korrelationsmatrix aus der Ausgabedatei ausschließen. Die Korrelationsmatrix wird nur ausgeschlossen, wenn sie nicht für eine andere angeforderte Ausgabe erforderlich ist (wenn beispielsweise Scorewerte oder Residuen-Korrelationen angezeigt werden), und wenn Sie **Bestehendes Faktormuster** als Faktormethode auswählen. Diese Methode ist der Standard für Dateien des Typs FACTOR. **Hinweis:** Die Option **Faktoren** ist für Eingabedateien des Typs CORR, UCORR, COV, UCOV oder FACTOR nicht verfügbar. Diese Option ist auch nicht verfügbar, wenn Sie der Rolle **Partielle Variablen** eine Variable zugewiesen haben.

Optionsname	Beschreibung
	Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Ausgabe anzeigen	Wählen Sie mindestens eine der folgenden Optionen im Bereich **Ausgabe anzeigen**: ■ **Korrelationsmatrix der Residuen** zeigt die Korrelationsmatrix der Residuen und die verwandte partielle Korrelationsmatrix an. Die diagonalen Elemente der Korrelationsmatrix der Residuen entsprechen den eindeutigen Varianzen. ■ **Eigenvektoren** zeigt die Eigenvektoren der verringerten Korrelationsmatrix an, in der die diagonalen Elemente durch die Kommunalitätsschätzer ersetzt werden. Wenn Sie die Faktormethode "Maximum-Likelihood-Faktoranalyse" verwenden, gelten die Eigenvektoren für die gewichtete, reduzierte Korrelationsmatrix. ■ **Faktor-Scoring-Koeffizienten** zeigt die Faktor-Scoring-Koeffizienten an. Die quadrierte multiple Korrelation jedes Faktors mit den Variablen wird im Falle nicht gedrehter Hauptkomponenten ebenfalls angezeigt. ■ **Matrixzeilen nach höchster absoluter Ladung neu ordnen** ordnet die Zeilen (Variablen) der verschiedenen Faktormatrizes in der Ausgabe neu an. Variablen, deren höchste absolute Ladung (Referenzstrukturladung für Oblique-Rotationen) am ersten Faktor vorliegt, werden zuerst absteigend von der größten zur kleinsten Ladung angezeigt. Danach folgen die Variablen, deren höchste absolute Ladung am zweiten Faktor vorliegt usw. ■ **Mittelwerte und Standardabweichungen der Eingabespalten** zeigt Mittelwerte, Standardabweichungen und die Anzahl der Beobachtungen an. ■ **Korrelationsmatrix der Eingabespalten** zeigt die Korrelationsmatrix oder die partielle Korrelationsmatrix an. ■ **Kaisers Maß der Stichprobengüte** erstellt die partiellen Korrelationen zwischen den einzelnen Variablenpaaren, sodass alle anderen Variablen (negative Anti-Bild-Korrelationen) und Kaisers Maß der Stichprobengüte kontrolliert werden.

Optionsname	Beschreibung
	■ **Eingabefaktormuster** zeigt das Eingabefaktormuster und verwandte Statistiken an. Bei Oblique-Fällen werden die Referenz- und Faktorstrukturen berechnet und angezeigt. **Hinweis:** Diese Option ist nur verfügbar, wenn Sie die Methode Bestehendes Faktormuster ausgewählt haben. ■ **Scoring-Koeffizienten** zeigt die Scoring-Koeffizienten und verwandten Statistiken an. Bei Oblique-Fällen werden die Referenz- und Faktorstrukturen berechnet und angezeigt. **Hinweis:** Diese Option ist nur verfügbar, wenn Sie die Methode Faktor-Scoring-Koeffizienten ausgewählt haben. Standardmäßig ist kein Faktorergebnis und keine entsprechende Statistikberechnung ausgewählt.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

33

Forecast Studio - Projekt erstellen

Allgemeines über den Assistenten 'Forecast Studio - Projekt erstellen' 231
Forecast Studio - Projekt erstellen: Überprüfen der Daten 232
Forecast Studio - Projekt erstellen: Auswählen der Prognosevariablen 232
Forecast Studio - Projekt erstellen: Auswählen der Klassifizierungsvariablen 233
Forecast Studio - Projekt erstellen: Speichern des Projekts 235

Allgemeines über den Assistenten 'Forecast Studio - Projekt erstellen'

Der Assistent **Forecast Studio - Projekt erstellen** hilft Ihnen bei der Angabe von Prognosevariablen, bei der Entscheidung, Ihre Daten hierarchisch vorherzusagen, und bei der Angabe des Prognosehorizonts für neue SAS Forecast Studio-Projekte. **SAS Forecast Studio** ist eine Anwendung, die durch Implementierung automatisierter Abläufe die Zeit, die für das Erstellen und Verarbeiten von Prognosen benötigt wird, wesentlich verkürzt. Die Software ermöglicht die automatische Auswahl von Zeitreihenmodellen, die zur Prognose von mit Zeitstempeln versehenen Daten verwendet werden. Um ein Projekt zu speichern, ist der Zugriff auf SAS Forecast Server erforderlich.

Hinweis: Wenn Sie ein Prognoseprojekt erstellen, in dem SAS-Daten von Ihrem PC verwendet werden sollen, wählen Sie die Datenquelle über das Dateisystem, nicht durch Auswahl der Datenquelle auf Ihrem lokalen Rechner aus.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SAS Forecast Server-Prozeduren
Erforderliche SAS-Produkte	SAS Forecast Server
Empfohlene zusätzliche SAS-Produkte	Ohne

Forecast Studio - Projekt erstellen: Überprüfen der Daten

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Klicken Sie auf **Weiter**, um die Prognosevariablen auszuwählen.

Forecast Studio - Projekt erstellen: Auswählen der Prognosevariablen

1 Bevor Sie die Prognosevariablen auswählen können, müssen Sie die Daten überprüfen.

2 Weisen Sie folgenden Rollen Variablen zu:

Zeit-ID
Diese Variable legt die Zeit-ID für die Daten fest. Die Zeit-ID ist eine Variable in der Eingabedatei, die den SAS-Datums-, -Zeit- oder -Datetime-Wert jeder Beobachtung enthält. Diese Variable wird verwendet, um die Häufigkeit und die Reihenfolge der Daten zu bestimmen, und um die Zeit-ID-Werte für die Prognosen zu extrapolieren. Sie können dieser Rolle nur eine Variable zuweisen, die eine Datumsvariable, eine Datetime-Variable oder eine numerische Variable sein muss, die Datums- oder Datetime-Werte enthält.

Prognose (Abhängige Variablen)
Die Spalten, die Sie dieser Rolle zuweisen, legen fest, welche Variablen für die Modellierung und Prognose verwendet werden. Die Prognosevariablen werden auch als abhängige Variablen bezeichnet. Sie müssen dieser Rolle mindestens eine numerische Variable zuweisen. Legen Sie mit dem Nach-oben- bzw. Nach-unten-Pfeil die Reihenfolge der Variablen fest.

3 (Optional) Weisen Sie der Rolle **Einflussfaktoren (unabhängie Variablen)** eine oder mehrere numerische Variablen zu. Die Spalte, die Sie dieser Rolle zuweisen, legt eine erklärende, Eingabe-, Prädiktor oder Kausalfaktor-Variable fest. Die Einflussfaktoren-Variablen werden auch als unabhängige Variablen bezeichnet. Legen Sie die Reihenfolge der Variablen fest.

4 Klicken Sie auf **Weiter**, um die Klassifizierungsvariablen auszuwählen.

Forecast Studio - Projekt erstellen: Auswählen der Klassifizierungsvariablen

1 Bevor Sie die Klassifizierungsvariablen auswählen können, müssen Sie die Prognosevariablen auswählen.

2 Weisen Sie der Rolle **Für jeden Wert von: (Klassifizierungsvariable)** mindestens eine Variable zu. Variablen, die dieser Rolle zugewiesen werden, werden als Klassifizierungsvariablen bezeichnet. Beobachtungen mit demselben Wert für die Klassifizierungsvariable werden gruppiert. Durch die Zuweisung einer Klassifizierungsvariablen erhalten Sie separate Analysen für gruppierte Beobachtungen. Sie können dieser Rolle alphanumerische und numerische Variablen zuweisen.

Die Reihenfolge der Klassifizierungsvariablen beschreibt die Struktur der Hierarchie. Legen Sie mit dem Nach-oben- bzw. Nach-unten-Pfeil die Reihenfolge der Variablen fest.

3 Legen Sie fest, wie die Ausgabe dargestellt werden soll. Folgende Optionen können ausgewählt werden:

- **Listenansicht** erstellt keine Hiearchie der Daten. Die prognostizierten Zeitreihen werden in Listenform angeordnet.
- **Hierarchie** verwendet die Werte der Klassifizierungsvariablen, um die Daten in Hierarchien anzuordnen. Diese Option ist standardmäßig aktiviert.

Hinweis: Wenn Sie der Rolle **Prognose (abhängige Variablen)** mehrere Variablen zuweisen, steht die Option **Hierarchie** nicht zur Verfügung.

4 Wenn Sie in Ihrer Ausgabe eine Hierarchie darstellen möchten, klicken Sie auf **Einstellungen**, um die Hierarchieeinstellungen festzulegen. Das Fenster **Hierarchieeinstellungen** wird geöffnet.

Folgende Optionen können ausgewählt werden:

Aggregationsmethode

Geben Sie die Statistik zum Aggregieren von Daten in verschiedenen Hierarchieebenen an. Standardmäßig werden alle abhängigen und unabhängigen Variablen in der Tabelle aufgeführt. Wählen Sie in der Spalte **Aggregation** die zu verwendende Aggregationsstatistik für jede Variable aus.

Beispiel: In Ihrer Datei sind die Umsätze für eine Produktgruppe enthalten. Sie möchten nun die Gesamtumsätze für eine bestimmte Kategorie ermitteln. In diesem Fall wählen Sie als Aggregationsstatistik die Option **Summe** aus. Da in Ihren Daten auch der Preis je Produkt enthalten ist, sind Sie daran interessiert, den durchschnittlichen Preis für eine Produktlinie zu ermitteln. Verwenden Sie dazu die Aggregationsstatistik **Durchschnitt**.

Konsolidierungsebene

Wählen Sie die Ebene in der Hierarchie aus, in der die Daten für die Prognose enthalten sind. Folgende Optionen können ausgewählt werden:

- **Bottom-Up** - verwendet die Daten auf der untersten Ebene der Hierarchie, um die Prognosen zu generieren. Auf Basis dieser Prognosen werden anschließend die konsolidierten Prognosen für die höheren Hierarchieebenen erstellt.

  Mit der Bottom-Up-Methode können Sie beliebige Muster (z. B. Saisonalität) in den Daten anzeigen. Da jedoch die Daten der untersten Hierarchieebene betrachtet werden, ist die Wahrscheinlichkeit für ein starkes Rauschen und Zufälligkeiten in den Daten gegeben. Darüber hinaus können Prognosen fehlschlagen, da es sich bei Daten auf der untersten Ebene der Hierarchie um vereinzelte Daten mit starker Streuung handeln kann.

- **Middle-Out** - summiert die Daten der unteren Ebenen und verwendet diese Werte anschließend zum Erstellen der Prognosen für die mittlere Ebene. Da einige Hierarchien mehr als eine mittlere Ebene aufweisen, müssen Sie die gewünschte Ebene explizit angeben. Die Prognosen auf der mittleren Ebene werden zur Erstellung von Prognosen für die höheren als auch für die unteren Ebenen verwendet.

- **Top-Down** - summiert die Daten der untersten Ebenen und verwendet diese Werte anschließend zum Erstellen der Prognosen. Die Prognose auf der höchsten Ebene wird verwendet, um die Prognosen für die unteren Ebenen der Hierarchie zu erstellen.

  Die Top-Down-Methode eignet sich zum Entfernen des statistischen Rauschens auf den unteren Ebenen der Hierarchie. Es kann jedoch vorkommen, dass hierbei auch Datenmuster (z.B. Saisonalität) in der Prognose verloren gehen.

Disaggregationsverfahren

Wählen Sie die Disaggregationsmethode für die Hierarchie aus. Folgende Optionen können ausgewählt werden:

- **Proportionale Aufteilung** - konsolidiert Prognosen auf Basis des historischen Mittelwerts. Hierbei handelt es sich um die Standardeinstellung.

- **Gleichmäßige Aufteilung der Differenz** - konsolidiert Prognosen auf Basis von prognostizierten Abweichungen zwischen Ebenen.

5 Legen Sie im Feld **Anzahl der zu prognostizierenden Perioden (Horizont)** die Anzahl der Perioden in der Zukunft fest, für die Mehr-Schritt-Prognosen erstellt werden sollen. Der Standardwert ist `12`.

6 Klicken Sie auf **Weiter**, um das Projekt zu speichern.

Forecast Studio - Projekt erstellen: Speichern des Projekts

1 Bevor Sie ein Projekt speichern können, müssen Sie die Klassifizierungsvariablen auswählen.

2 Geben Sie im Feld **Speichern unter** einen Namen für das Projekt ein. Der Projektname muss ein gültiger SAS-Name sein. Der Projektname kann bis zu 32 Zeichen lang sein und muss mit einem Buchstaben (A–Z) beginnen. Die weiteren Zeichen im Namen können Buchstaben oder Zahlen (0–9) sein. Es ist sowohl Groß- als auch Kleinschreibung zulässig. Weitere Informationen zu SAS-Namenskonvention finden Sie in *SAS Language Reference: Concepts*

3 Klicken Sie auf **Fertig stellen**, um das SAS Forecast Studio-Projekt zu erstellen. Sie können das Projekt nun in SAS Forecast Studio öffnen.

34

Forecast Studio - Projekt öffnen

Allgemeines über den Assistenten 'Forecast Studio - Projekt öffnen' 237
Forecast Studio - Projekt öffnen: Auswählen des Prognoseprojekts 238
Forecast Studio - Projekt öffnen: Auswählen der Prognosezeitreihe 238
Forecast Studio - Projekt öffnen: Auswählen des Ergebnistyps 239
Forecast Studio - Projekt öffnen: Festlegen von Titeln und Fußnoten 239

Allgemeines über den Assistenten 'Forecast Studio - Projekt öffnen'

Der Assistent **Forecast Studio - Projekt öffnen** hilft Ihnen beim Öffnen der gewählten Zeitreihen aus einem bestehenden SAS Forecast Studio-Projekt und beim Anzeigen der Ergebnisse. SAS Forecast Studio ist eine Anwendung, die durch Implementierung automatisierter Abläufe die Zeit, die für das Erstellen und Verarbeiten von Prognosen benötigt wird, wesentlich verkürzt. Die Software ermöglicht die automatische Auswahl von Zeitreihenmodellen, die zur Prognose von mit Zeitstempeln versehenen Daten verwendet werden.

Mit dieser Anwendungsroutine können Sie ein in SAS Forecast Studio erstelltes Projekt öffnen und weitere Analysen durchführen. Um ein Prognoseprojekt zu speichern, ist der Zugriff auf SAS Forecast Server erforderlich.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SAS Forecast Server-Prozeduren
Erforderliche SAS-Produkte	SAS Forecast Server

Name der Anforderung	Prozedur- und Produktnamen
Empfohlene zusätzliche SAS-Produkte	Ohne

Forecast Studio - Projekt öffnen: Auswählen des Prognoseprojekts

Im Fenster **Prognoseprojekte** wird eine Liste aller verfügbaren Projekte inklusive des letzten Datums der letzten Änderung angezeigt. Wählen Sie das Projekt aus, das Sie öffnen möchten. Es ist nur möglich, jeweils ein Objekt gleichzeitig zu öffnen.

Die Details zu einem Projekt werden nicht standardmäßig angzeigt. Um die Beschreibung, abhängigen Variablen und Klassifizierungsvariablen (BY-Variablen) anzuzeigen, die dem Projekt zugewiesen wurden, klicken Sie auf **Details**. Da diese Informationen auf dem Forecasting Server abgelegt sind, kann das Aufrufen der Details einen kleinen Augenblick dauern. Wenn Sie die Details wieder ausblenden möchten, klicken Sie auf **Ausblenden**.

Klicken Sie auf **Weiter**, um die Prognosezeitreihe auszuwählen.

Forecast Studio - Projekt öffnen: Auswählen der Prognosezeitreihe

1. Bevor Sie die Prognosezeitreihe auswählen können, müssen Sie das Prognoseprojekt auswählen.
2. Aktivieren Sie im Fenster **Reihen für die Anzeige auswählen** das Kontrollkästchen für die zu öffnende Zeitreihe.

   Die Zeitreihen werden in Listenform dargestellt, wenn die Projektdaten nicht hierarchisch prognostiziert wurden. Die Zeitreihen werden in Hierarchieform dargestellt, wenn die Daten hierarchisch prognostiziert wurden. Klicken Sie auf **Alle aufklappen**, um alle Knoten der Hierarchie zu öffnen. Um die Knoten zuzuklappen, klicken Sie auf **Alle zuklappen**.
3. Klicken Sie auf **Weiter**, um die Ergebnistypen auszuwählen.

Forecast Studio - Projekt öffnen: Auswählen des Ergebnistyps

1 Bevor Sie die Ergebnistypen auswählen können, müssen Sie die Prognosezeitreihe auswählen.

2 Wählen Sie den gewünschten Ergebnistyp aus. Folgende Optionen können ausgewählt werden:

- **Prognose-Plot** zeigt die historischen, die angepassten Werte sowie die konsolidierte Prognose für das gewählte Modell an.
- **Prognosetabelle** zeigt das Modell für die statistische Prognose, die Konsolidierungsanpassung, die konsolidierte Prognose, die Werte der manuellen Wertekorrekturen sowie die endgültige Prognose an.
- **Modellauswahltabelle** zeigt alle Modelle an, die für das aktuelle Projekt definiert wurden.
- **Tabelle für Modell-Parameterschätzung** zeigt Schätzungen, Standardfehler und Signifikanztests für jeden der Modellparameter an.
- **Komponententabelle** zeigt Tabellen für jede der im Modell verwendeten Komponenten an.
- **Anpassungsgütetabelle** zeigt alle verfügbaren Anpassungsgütestatistiken für das Prognosemodell an.
- **Zusammenfassungstabelle für Prognose** zeigt Prognosewerte der abhängigen Variable für den Zeitraum des Prognosehorizonts an.

3 Klicken Sie auf **Weiter**, um Titel und Fußnoten anzugeben.

Forecast Studio - Projekt öffnen: Festlegen von Titeln und Fußnoten

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

35

Forecast Studio - Wertekorrektur für Projekt

Allgemeines über den Assistenten 'Forecast Studio - Wertekorrekturen für Projekt' . . 241

Forecast Studio - Wertekorrektur für Projekt: Auswählen des Prognoseprojekts 242

Forecast Studio - Wertekorrektur für Projekt: Auswählen der Prognosezeitreihe und Übertragen der Wertekorrekturen 243

Auswählen der Prognosezeitreihe 243

Erstellen und übertragen der Wertekorrekturen 243

Berechnen der Wertekorrekturen für die gewählten Zeitperioden 244

Lösen von Wertekorrekturkonflikten 246

Allgemeines über den Assistenten 'Forecast Studio - Wertekorrekturen für Projekt'

SAS Forecast Studio ist eine Anwendung, die durch Implementierung automatisierter Abläufe die Zeit, die für das Erstellen und Verarbeiten von Prognosen benötigt wird, wesentlich verkürzt. Die Software ermöglicht die automatische Auswahl von Zeitreihenmodellen, die zur Prognose von mit Zeitstempeln versehenen Daten verwendet werden. Um ein Prognoseprojekt in SAS Enterprise Guide oder SAS Add-In to Microsoft Office öffnen zu können, ist der Zugriff auf SAS Forecast Server erforderlich.

Der Assistent **Forecast Studio - Wertekorrekturen für Projekt** ermöglicht es Ihnen, Wertekorrekturen für die Prognosedaten in einem SAS Forecast Studio-Projekt zu übertragen. Bei einer Wertekorrektur wird ein prognostizierter Wert durch einen zuvor angegebenen ersetzt. Wertekorrekturen können nur für Werte, die in der Zukunft liegen, übertragen werden.

Nach der Übertragung einer Wertekorrektur wird das Projekt neu in SAS Forecast Server konsolidiert. Wie die Wertekorrekturen angewendet werden, ist davon abhängig ob die Daten im Projekt hierarchisch prognostiziert werden.

- Werden die Daten in einem Projekt hierarchisch prognostiziert, wird die Anpassung der Wertekorrekturen unter Berücksichtigung der konsolidierten statistischen Prognose vorgenommen.
- Werden die Daten in einem Projekt nicht hierarchisch prognostiziert, wird die Anpassung der Wertekorrekturen nur unter Berücksichtigung der statistischen Prognose vorgenommen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SAS Forecast Server-Prozeduren
Erforderliche SAS-Produkte	SAS Forecast Server
Empfohlene zusätzliche SAS-Produkte	Ohne

Forecast Studio - Wertekorrektur für Projekt: Auswählen des Prognoseprojekts

Im Fenster **Prognoseprojekte** wird eine Liste aller verfügbaren Projekte inklusive des letzten Datums der letzten Änderung angezeigt. Wählen Sie das Projekt aus, das Sie öffnen möchten. Es ist nur möglich, jeweils ein Objekt gleichzeitig zu öffnen.

Die Details zu einem Projekt werden nicht standardmäßig angzeigt. Um die Beschreibung, abhängigen Variablen und Klassifizierungsvariablen (BY-Variablen) anzuzeigen, die dem Projekt zugewiesen wurden, klicken Sie auf **Details**. Da diese Informationen auf dem Forecasting Server abgelegt sind, kann das Aufrufen der Details einen kleinen Augenblick dauern. Wenn Sie die Details wieder ausblenden möchten, klicken Sie auf **Ausblenden**.

Klicken Sie auf **Weiter**, um die Prognosezeitreihe auszuwählen.

Forecast Studio - Wertekorrektur für Projekt: Auswählen der Prognosezeitreihe und Übertragen der Wertekorrekturen

Auswählen der Prognosezeitreihe

1. Bevor Sie die Prognosezeitreihe auswählen können, müssen Sie das Prognoseprojekt auswählen.
2. Wählen Sie in der Baumansicht die Zeitreihe aus, für die Sie eine Wertekorrektur übertragen möchten.

   Wurden die Projektdaten hierarchisch prognostiziert, werden die Zeitreihen in einer Hierarchie dargestellt. Klicken Sie auf **Alle aufklappen**, um alle Knoten der Hierarchie zu öffnen. Um die Knoten zusammenzuklappen, klicken Sie auf **Alle zusammenklappen**.

   Die Zeitreihen werden in Listenform dargestellt, wenn die Projektdaten nicht hierarchisch prognostiziert wurden.
3. Klicken Sie nach der Auswahl der Zeitreihe auf **Wertekorrektur für Prognosedaten**, um die Wertekorrekturen zu erstellen und übertragen.
4. Wiederholen Sie die Schritte 2 und 3 für alle Zeitreihen, für die Sie eine Wertekorrektur übertragen möchten.
5. Klicken Sie nach der Erstellung und der Übertragung Ihrer Wertekorrekturen auf **Fertig stellen**, um den Assistenten zu beenden.

Erstellen und übertragen der Wertekorrekturen

Im Dialogfeld **Wertekorrektur für Prognosedaten** werden oben zwei Grafiken angezeigt. In der linken Grafik stellt die blaue Linie die Prognose dar, die historischen Daten werden durch einen Kreis dargestellt. In der rechten Grafik zeigt die blaue Linie die Prognose an, die hellblauen Linien stellen die Konfidenzintervalle für die Prognose und die schwarzen Pluszeichen alle angegebenen Wertekorrekturen dar.

So erstellen und übertragen Sie Wertekorrekturen:

1 Wählen Sie in der Tabelle **Wertekorrekturen bearbeiten** die Zeitperiode aus, für die Sie eine Wertekorrektur hinzufügen oder bearbeiten möchten. Halten Sie die STRG-Taste gedrückt und wählen Sie verschiedenen Zeilen in der Tabelle aus, um ggf. mehrere Zeitperioden auf einmal auszuwählen.

2 Um eine Wertekorrektur für eine einzelne Zeitperiode zu erstellen, klicken Sie in der Spalte **Wertekorrekturen** für die gewünschte Periode auf eine Zelle. Geben Sie den Wert für die Wertekorrektur ein und drücken Sie die Eingabetaste.

   Um Wertekorrekturen für ausgewählte Zeitperioden zu berechnen, klicken Sie auf **Wertekorrekturen berechnen**. Das Dialogfeld **Wertekorrektur-Editor** wird angezeigt.

   Um Wertekorrekturen, die Sie während einer Sitzung festgelegt haben, zu entfernen, klicken Sie auf **Zurücksetzen**. Wertekorrekturen, die Sie davor bereits festgelegt hatten oder die in SAS Forecast Studio angegeben wurden, werden nicht entfernt.

3 Klicken Sie nach der Erstellung der Wertekorrekturen für die aktuelle Zeitreihe auf **Übertragen**. Das Projekt wird von SAS Forecast Server konsolidiert. Nach der Konsolidierung wird das Fenster **Wertekorrektur für Prognosedaten** geschlossen und im Assistenten wird der Schritt zum Auswählen der Prognosezeitreihen aufgerufen. In der Baumansicht wird neben den Zeitreihen, für die Sie Wertekorrekturen übertragen haben, das Symbol Überschreiben angezeigt.

   Hinweis: Erkennt SAS Forecast Server bei der Konsolidierung eines Projekts Wertekorrekturkonflikte, erhalten Sie einen entsprechenden Warnhinweis darüber, dass das Projekt nicht konsolidiert werden kann. Verwenden Sie in diesem Fall SAS Forecast Studio, um die Konflikte zu konsolidieren.

Berechnen der Wertekorrekturen für die gewählten Zeitperioden

Folgende Optionen stehen zum Berechnen der Wertekorrekturen für eine oder mehrere Zeitperioden zur Verfügung:

Optionsname	Beschreibung
Relativ zum Prognosewert festlegen	ermöglicht es Ihnen, die Werte für die Wertekorrektur auf Basis der angegebenen Erhöhung bzw. Verringerung der konsolidierten statistischen anzugeben. **Hinweis:** Falls Sie keine hierarchische Prognose für Ihre Daten erstellt haben oder keine Konsolidierungsprognose verfügbar ist, können Sie diese Option auswählen, um die Werte für die Wertekorrekturen auf der Basis der

Optionsname	Beschreibung
	festgelegten Erhöhung bzw. Verringerung der statistischen Prognose anzugeben. Verwenden Sie die Dropdown-Listen und das Textfeld, um die Anzahl der Einheiten oder den Prozentwert festzulegen, um die bzw. den der aktuelle Wert erhöht bzw. verringert werden soll. Wählen Sie die Option **Erhöhen um** oder **Vermindern um**. Geben Sie im Textfeld für die gewählte Option den gewünschten Wert ein und legen Sie anschließend fest, ob es sich bei dem Wert um eine Einheit oder einen Prozentwert handelt. Beispiel: Sie möchten eine Wertekorrektur vornehmen, bei der die korrigierten Werte am Ende um 10% über der konsolidierten statistischen Prognose liegen sollen. Wählen Sie die Option **Erhöhen um**. Geben Sie in das Textfeld den Wert 10 ein und wählen Sie aus der Dropdown-Liste die Option **% (Prozent)**. Die korrigierten Werte liegen um 10% über den Werten der konsolidierten statistischen Prognose.
Einen Wert über ausgewählte Perioden verteilen	Verteilt einen einzelnen Wert für die Wertekorrektur über alle ausgewählten Perioden. Geben Sie den Wert in das Textfeld ein. Wie der Wert für die Wertekorrektur angewendet wird, ist abhängig von der gewählten Option. Folgende Optionen können ausgewählt werden: ■ **Proportional zu aktuellem Prognosewert auf jede Periode verteilen** - Der angegebene Wert wird proportional auf die ausgewählten Zeitperioden aufgeteilt. Beispiel: Sie haben zwei Zeitperioden mit den Prognosewerten 550 und 430 ausgewählt. Als Wert für die Korrektur wird 100 angegeben und Sie möchten diesen Wert proportional auf die beiden Zeitperioden aufteilen. Daraus ergeben sich die Prognosewerte 606,12 und 473,88 ■ **Gleichmäßig über Perioden verteilen** - Der angegebene Wert wird gleichmäßig auf die ausgewählten Zeitperioden aufgeteilt. Beispiel: Sie haben zwei Zeitperioden mit den Prognosewerten 550 und 430 ausgewählt. Als Wert für die Korrektur wird 100 angegeben und Sie möchten diesen Wert gleichmäßig auf die beiden Zeitperioden aufteilen. Daraus ergeben sich die Prognosewerte 600 und 480.
Jeder ausgewählten Periode einen Wert zuweisen	Wendet einen einzelnen Wert für die Wertekorrektur auf alle ausgewählten Perioden an. Geben Sie den Wert in das Textfeld ein.

Klicken Sie nach der Berechnung der Werte für die Wertekorrektur für die gewählten Zeitperioden auf **OK**, um die Wertekorrektur zu übertragen.

Lösen von Wertekorrekturkonflikten

Erkennt SAS Forecast Server bei der Konsolidierung eines Projekts Wertekorrekturkonflikte, erhalten Sie einen entsprechenden Warnhinweis darüber, dass das Projekt nicht konsolidiert werden kann. Verwenden Sie in diesem Fall SAS Forecast Studio, um die Konflikte zu konsolidieren. In SAS Forecast Studio werden alle übertragenen Wertekorrekturen angezeigt und Sie können auswählen, welche Wertekorrekturen für die Prognosedaten übernommen werden sollen.

So lösen Sie Wertekorrekturkonflikte:

1. Klicken Sie auf **Abbrechen**, um den **Forecast Studio-Assistenten zum Übertragen von Wertekorrekturen** zu beenden und Microsoft Excel zu schließen.
2. Öffnen Sie das Projekt in SAS Forecast Studio. Beim Öffnen von SAS Forecast Studio wird die Warnmeldung "Die Konsolidierung ist nicht mehr aktuell." im Arbeitsbereich angezeigt.
3. Klicken Sie neben der Warnung auf **Aktualisieren**. Das Dialogfenster **Wertekorrekturen - Konflikte** wird angezeigt.
4. Um die Konflikte zu lösen, können Sie die Wertekorrekturen im Dialogfeld **Wertekorrekturen - Konflikte** ändern. Darüber hinaus haben Sie die Möglichkeit, dieses Dialogfeld zu schließen und die Wertekorrekturen in der Datentabelle zu aktualisieren.

Weitere Informationen zum Lösen von Wertekorrekturkonflikten finden Sie in der Hilfe zu SAS Forecast Studio.

36

Verallgemeinerte lineare Modelle

Allgemeines über die Anwendungsroutine 'Verallgemeinerte lineare Modelle' 247
Verallgemeinerte lineare Modelle: Zuweisen von Variablen zu Analyserollen 248
Verallgemeinerte lineare Modelle: Erstellen eines Modells 250
Verallgemeinerte lineare Modelle: Festlegen von Modelloptionen 250
Verallgemeinerte lineare Modelle: Auswählen von Post-Hoc-Tests 252
Verallgemeinerte lineare Modelle: Erzeugen von Plots 252
Verallgemeinerte lineare Modelle: Festlegen von Vorhersageoptionen 254
Festlegen von Titeln und Fußnoten 255
Anzeigen von Eigenschaften 255

Allgemeines über die Anwendungsroutine 'Verallgemeinerte lineare Modelle'

Die Anwendungsroutine **Verallgemeinerte lineare Modelle** ist eine Erweiterung für herkömmliche lineare Modelle, mit der Sie Daten modellieren können, die nicht normal verteilt sind, wie z. B. Zählungen oder gemessene Proportionen, die als diskret angesehen werden. Sie können die Anwendungsroutine auch verwenden, um Daten zu modellieren, deren Mittelwert auf einen Werte- oder Datenbereich beschränkt wurde, bei dem die Varianz für alle Beobachtung nicht konstant ist.

Mit dieser Anwendungsroutine können folgende Aufgaben erledigt werden:

- Vergleichen der Lebensspanne von Produkten zweier verschiedener Hersteller. Angenommen, Ihnen liegen die Ausfallzeiten von Maschinenteilen vor, die zum Teil von Hersteller A und zum Teil von Hersteller B gefertigt wurden. Mithilfe dieser Anwendungsroutine können Sie die Unterschiede des Lebenszyklus der Teile beider Hersteller testen.

- Anpassen eines Modells an Daten, die auf einer ordinalen Skala gemessen werden. Ihnen liegt beispielsweise eine Datei mit den Ergebnissen eines Geschmackstests dreier Eiskremmarken vor. Der Geschmack der drei Marken kann auf einer Skala mit fünf Punkten von sehr gut (sg) bis sehr schlecht (ss) bewertet werden Für die Auswertung der Unterschiede bei der Bewertung der drei Marken wird eine Analyse durchgeführt.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GENMOD
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Verallgemeinerte lineare Modelle: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Abhängige Variablen	gibt die Variablen an, die als abhängige (Response) Variablen verwendet werden. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Quantitative Variablen	gibt die Variablen an, die als kontinuierliche, unabhängige (erklärende) Variablen (auch Effekte genannt) verwendet werden.

Optionsname	Beschreibung
Klassifizierungsvariablen	gibt die Variablen an, die als diskrete unabhängige Effekte verwendet werden. Variablen, die Sie dieser Rolle zuweisen, können numerisch oder alphanumerisch sein, sollten aber eine begrenzte Anzahl an diskreten Werten aufweisen.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits als abhängige Variable oder als unabhängige Variable ausgewählt ist.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt eine Variable an, die als relative Gewichtung für die entsprechenden Werte der Variablen verwendet wird, die Sie als abhängige Variable ausgewählt haben. Sie können dieser Rolle maximal eine Variable zuweisen.
Relativer Offset	gibt eine Variable an, die als Offset-Variable verwendet werden soll. **Hinweis:** Sie können dieser Rolle maximal eine Variable zuweisen. Sie können keine Variable als Offset-Variable verwenden, die Sie bereits als abhängige oder unabhängige Variable, oder als Klassifizierungsvariable ausgewählt haben.

Verallgemeinerte lineare Modelle: Erstellen eines Modells

Sie können Modelle in mehreren SAS-Anwendungsroutinen erstellen. Weitere Informationen, siehe Anhang 3, “Erstellen eines Modells” auf Seite 765.

Verallgemeinerte lineare Modelle: Festlegen von Modelloptionen

Klicken Sie im Auswahlfenster auf **Modelloptionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Verteilung	gibt die Wahrscheinlichkeitsverteilung für die abhängige Variable in Ihrem Modell an. **Hinweis:** Binomial-Verteilungen sind nur verfügbar, wenn die abhängige Variable zwei Ebenen hat. **Hinweis:** Wenn der Rolle **Abhängige Variable** eine alphanumerische Variable zugewiesen wird, sind nur die multinomialen und binomialen Verteilungen verfügbar.
Kopplungsfunktionen	beschreibt, inwieweit der erwartete Wert mit dem linearen Prädiktor in Beziehung steht. Abhängig vom erstellten Modell stehen verschiedene Kopplungsfunktionen zur Verfügung. In der Dropdown-Liste **Kopplungsfunktion** können Sie die folgenden Kopplungsfunktionen wählen: ■ **Ohne** ■ **Identität** ■ **Logit** ■ **Probit** ■ **Log** ■ **Komplementäres Log-Log** ■ **Potenz mit Nummer**

Optionsname	Beschreibung
	Hinweis: Die Kopplungsfunktionen **Logit kumulativ**, **Log-Log kumulativ komplementär** und **Probit kumulativ** sind nur bei Auswahl einer multinomialen Verteilung verfügbar. Wenn Sie in der Dropdown-Liste **Kopplungsfunktion** keine Kopplungsfunktion auswählen und eine benutzerdefinierte Kopplungsfunktion vorhanden ist, wir diese benutzerdefinierte Kopplungsfunktion verwendet. Falls Sie keine benutzerdefinierte Kopplungsfunktion angegeben haben, wird die Standard-Kopplungsfunktion für das Modell entsprechend der folgenden Tabelle verwendet. ■ Bei einer Binomialverteilung ist die Standardkopplungsfunktion logit. ■ Bei einer Gamma-Verteilung ist die Standardkopplungsfunktion invers (Potenz(-1)). ■ Bei einer inversen Gauß-Verteilung ist die Standardkopplungsfunktion invers quadrat (Potenz(-2)). ■ Bei einer multinomialen Verteilung ist die Standardkopplungsfunktion kumulativ logit. ■ Bei einer negativen Binomialverteilung ist die Standardkopplungsfunktion log. ■ Bei einer Normalverteilung lautet die Standardkopplungsfunktion Identität. ■ Bei einer Poisson-Verteilung ist die Standardkopplungsfunktion log. **Hinweis:** Wenn Sie für die Verteilung und die Kopplungsfunktion **Keine** auswählen, wird in der Analyse die Normalverteilung mit der Identitätskopplungsfunktion verwendet.
Effekttests	legt die Statistiken fest, die in die Analyse einbezogen werden sollen. ■ Aktivieren Sie das Kontrollkästchen **Typ 1 oder sequenzielle Analyse berechnen**, um die Likelihood-Ratio-Statistik zwischen aufeinanderfolgenden Modellpaaren anzuzeigen. Die nacheinander angepassten Modelle beginnen mit dem Nullmodell (nur konstanter Begriff) und werden bis zu den von Ihnen definierten Modellen fortgesetzt. ■ Aktivieren Sie das Kontrollkästchen **Statistiken für Typ-3-Kontraste berechnen**, um die Kontraststatistiken für jeden Effekt in den definierten Modellen anzuzeigen. Bei der Standardanalyse wird die Likelihood-Ratio-Statistik für die Kontraste berechnet. Wald-Statistiken werden auch dann berechnet, wenn das Kontrollkästchen **Wald-Statistik berechnen** ausgewählt ist.

Optionsname	Beschreibung
Konfidenzgrenzen	gibt das Konfidenzniveau für die vorhergesagten Werte an. Der Standardwert ist ein Konfidenzniveau von 95%.
Details schätzen	gibt an, ob die Parameterschätzer-Korrelationsmatrix, die Parameterschätzer-Kovarianzmatrix oder beie angezeigt werden sollen.
Konfidenzintervalle	gibt an, ob ein zweiseitiges Konfidenzintervall verwendet werden soll. ■ Aktivieren Sie das Kontrollkästchen **Zweiseitig**, um zweiseitige Konfidenzintervalle für alle Modellparameter, die auf der Profil-Likelihood-Funktion basieren (auch bekannt als teilweise maximierte Likelihood-Funktion) zu berechnen. ■ Aktivieren Sie das Kontrollkästchen **Zweiseitig - Wald**, um zweiseitige Wald-Konfidenzintervalle für alle Modellparameter basierend auf der asymptotischen Normalität von Parameterschätzern zu berechnen.

Verallgemeinerte lineare Modelle: Auswählen von Post-Hoc-Tests

Klicken Sie im Auswahlfenster auf **Post-Hoc-Tests**, um auf diese Optionen zuzugreifen.

Geben Sie für jeden Effekt im Feld **Zu schätzende Effekte** Optionen im Feld **Optionen für Tests auf Mittelwerte** an.

Verallgemeinerte lineare Modelle: Erzeugen von Plots

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Standardmäßig werden alle für die aktuelle Datenauswahl geeigneten Plots in der Ausgabe berücksichtigt. Sie können jedoch durch Auswahl der Option

Benutzerdefinierte Plot-Listen festlegen, welche Plots Sie in die Ausgabe aufnehmen möchten. Sie können folgende Optionen wählen:

Optionsname	Beschreibung
Cook's D-Plot	erstellt einen Plot der Cook's D-Statistik gegen die Beobachtungszahl.
DFBETA-Plots	erstellt Plots für die Differenz der Parameterschätzer durch Löschen der Beobachtung.
DFBETAS-Plots	stellt den standardisierten Unterschied im Parameterschätzer aufgrund der Löschung der Beobachtung dar. Es ist hilfreich, den Effekt einer einzelnen Beobachtung für jeden einzelnen Parameterschätzer des angepassten Modells zu bestimmen.
Leverage-Plot	stellt den Wert des entsprechenden diagonalen Elements in der H-Matrix dar. Die diagonalen Elemente der Matrix helfen beim Auffinden von Extrempunkten im Entwurfsbereich, in dem die Werte gewöhnlich größer sind.
Vorhergesagter-Wert-Plot	erstellt einen Plot der vorhergesagten Werte.
Pearson-Residuum-Plot	erstellt einen Plot des Pearsonschen Residuums (Chi), der zum Identifizieren von Beobachtungen, die vom Modell nur unwesentlich berücksichtigt werden, gegen vorhergesagte Werte dient.
Abweichungsresiduum-Plot	erstellt einen Plot des Abweichungsresiduums, der benutzt wird, um schlecht angepasste Werte gegen die vorhergesagten Werte zu ermitteln.
Likelihood-Residuum-Plot	erstellt einen Plot des Likelihood-Residuums, der benutzt wird, um schlecht angepasste Werte gegen vorhergesagte Werte zu ermitteln.
Gewöhnliches-Residuum-Plot	erstellt einen Plot der gewöhnlichen Residuen gegen vorhergesagte Werte.
Standardisiertes Pearsonsches Residuum	erstellt einen Plot des standardisierten Pearsonschen Residuums (Chi), der zum Identifizieren von Beobachtungen, die vom Modell nur unwesentlich berücksichtigt werden, gegen vorhergesagte Werte dient.
Standardisiertes Abweichungsresiduum	erstellt einen Plot des standardisierten Abweichungs-Residuums, der zum Identifizieren von schlecht angepassten Werten gegen vorhergesagte Werte verwendet wird.

Verallgemeinerte lineare Modelle: Festlegen von Vorhersageoptionen

Klicken Sie im Auswahlfenster auf **Vorhersagen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Daten für Vorhersage	identifiziert die Datenquellen, die Sie zur Vorhersage der Ausgabedaten verwenden möchten. Wenn Sie das Kontrollkästchen **Weitere Daten** aktivieren sind diese Daten eingeschlossen. Um den Speicherort für Ihre zusätzlichen Daten zu bestimmen, klicken Sie auf **Durchsuchen**.
Ausgabedaten speichern	gibt an, ob die Vorhersagen und Diagnosestatistiken in einer Ausgabedatei gespeichert werden sollen: **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Weitere Statistiken	gibt an, ob weitere Statistiken gespeichert werden sollen. ■ **Residuen** speichert das Likelihood-Residuum, das zum Identifizieren schlecht angepasster Beobachtungen verwendet wird, das standardisierte Pearson-Residuum (Chi), das zum Identifizieren von Beobachtungen dient, die im Modell nur unwesentlich berücksichtigt werden, und das standardisierte Abweichungsresiduum, das zum Identifizieren schlecht angepasster Beobachtungen dient. ■ **Vorhersagegrenzen** speichert die obere und untere Konfidenzgrenze und die Konfidenzgrenze für den vorhergesagten Wert.
Vorhersagen anzeigen	zeigt die Vorhersagen in den Ergebnissen an.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

37

High-Performance - Lineare Regression

Allgemeines über die Anwendungsroutine 'High Performance - Lineare Regression' . . 257
High-Performance - Lineare Regression: Zuweisen von Variablen zu Analyserollen . . 258
High-Performance - Lineare Regression: Erstellen eines Modells 260
Erstellen eines Modells 260
Angeben einer Modellauswahlmethode 260
High-Performance - Lineare Regression: Festlegen von Modelloptionen 263
High-Performance - Lineare Regression: Festlegen der Partitionsoptionen 264
High-Performance - Regression: Festlegen von Performance-Optionen 265
High-Performance - Regression: Festlegen von Vorhersageoptionen 266
Festlegen von Titeln und Fußnoten 267

Allgemeines über die Anwendungsroutine 'High Performance - Lineare Regression'

Die Anwendungsroutine **High-Performance - Lineare Regression** ist für die Anpassung und die Durchführung der Modellselektion für geordnete lineare Kleinste-Quadratemodelle geeignet. Im Gegensatz zur Anwendungsroutine **Lineare Regression** können Sie hier allgemeine lineare Modelle angeben, die Klassifizierungsvariablen enthalten. Die Anwendungsroutine **High-Performance - Lineare Regression** unterstützt darüber hinaus LARS- (LARS, Least Angle Regression) sowie LASSO-Selektionsverfahren (LASSO, Least Absolute Shrinkage and Selection Operator).

In der Anwendungsroutine **High-Performance - Lineare Regression** stehen zwei Parametrisierungen für Klassifizierungsvariablen zur Verfügung (GLM und Referenz). Im Gegensatz zu anderen Anwendungsroutinen, die noch weitere Parametrisierungen bieten, haben die Klassifizierungsvariablen in dieser Anwendungsroutine dieselbe Parametrisierung. Standardmäßig wird die GLM-Parametrisierung verwendet.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	HPREG
Erforderliche SAS-Produkte	Base SAS, SAS High-Performance Analytics Server, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

High-Performance - Lineare Regression: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Abhängige Variablen	gibt die Variablen an, die als abhängige (Response) Variablen verwendet werden. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Erklärende Effekte	gibt die Variablen an, die als Regressorvariablen (erklärende Variablen) verwendet werden. Wenn Sie keine Regressorvariablen angeben, wird jede abhängige Variable an ein Nur-Mittelwert-Modell (Mean-Only) angepasst. Das bedeutet, dass jede abhängige Variable an ein Modell mit nur einem konstanten Term angepasst wird.
Klassifizierungsvariablen	gibt die Klassifizierungsvariablen an, die in der Analyse verwendet werden sollen. Hierbei kann es sich um alphanumerische oder numerische Variablen handeln. Bei numerischen Variablen müssen Einzelwerte statt kontinuierlicher Werte vorliegen. Diese

Optionsname	Beschreibung
	Variablen werden in die Analyse einbezogen durch Ausprägungen, denen eindeutige Werte zugewiesen werden. Bei der "Nivellierung" der Werte werden Beobachtungen, die denselben Wert aufweisen, derselben Ausprägung zugeordnet. Sie können den Kodierungsstil für alle oder für einzelne Klassifizierungsvariablen festlegen. Verwenden Sie die Optionen im Bereich **Globale Optionen für Klassifizierungsvariablen**, um denselben Kodierungsstil für alle Klassifizierungsvariablen festzulegen. ■ Verwenden Sie die Option **GLM** für eine Less-Than-Full-Rank-Referenzzellen-Kodierung. Diese Parametrisierung wird in den Prozeduren GLM, MIXED und GLIMMIX in SAS/STAT verwendet. ■ Verwenden Sie die Option **Referenz** für eine Referenzzellen-Kodierung. Legen Sie fest, ob Sie den ersten oder den letzten geordneten Wert als Referenz verwenden möchten. Wenn Sie die Option **Referenz** auswählen, können Sie bei einer einzelnen Klassifizierungsvariablen die Referenzzellen-Kodierung außer Kraft setzen. Mit dem Kontrollkästchen **Teilen** wird festgelegt, welche Spalten der Designmatrix einem Effekt entsprechen, der eine geteilte Klassifizierungsvariable (Split) enthält. Sie können für jede Spalte angeben, ob sie ins Modell aufgenommen bzw. dort belassen werden soll, unabhängig von den anderen Designspalten für den Effekt. Dies können Sie als globale Option oder als Option für eine einzelne Klassifizierungsvariable festlegen. Angenommen die Variable "Temp" weist drei Ausprägungen mit den Werten "Heiß", "Warm" und "Kalt" auf. Es gibt zwei Ausprägungen für das Geschlecht, die Werte "M" und "W". ■ Wenn die Option **Teilen** als globale Option festgelegt ist, werden die beiden Effekte in acht unabhängige Effekte aufgeteilt. Für das Geschlecht entstehen diese beiden Effekte: Geschlecht_M und Geschlecht_W. Der Effekt Geschlecht*Temp wird in sechs Effekte aufgeteilt: Geschlecht_M*Temp_Heiß, Geschlecht_W*Temp_Heiß, Geschlecht_M*Temp_Warm, Geschlecht_W*Temp_Warm, Geschlecht_M*Temp_Kalt und Geschlecht_F*Temp_Kalt ■ Wenn die Option **Teilen** als Option für die Variable "Temp" ausgewählt wurde, wird der Effekt Geschlecht*Temp in drei Effekte aufgeteilt: Geschlecht*Temp_Heiß, Geschlecht*Temp_Warm und Geschlecht*Temp_Kalt. Jeder dieser drei geteilten Effekte weist nun zwei Parameter auf, die den beiden Ausprägungen von "Geschlecht" entsprechen.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n*

Optionsname	Beschreibung
	Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt die Variable mit den Werten an, die relative Gewichtungen für eine gewichtete Kleinste-Quadrate-Anpassung darstellen. Die Werte der relativen Gewichtungsvariablen müssen positiv sein. Sie können dieser Rolle maximal eine Variable zuweisen.

High-Performance - Lineare Regression: Erstellen eines Modells

Erstellen eines Modells

Sie können Modelle in mehreren SAS-Anwendungsroutinen erstellen. Weitere Informationen, siehe Anhang 3, “Erstellen eines Modells” auf Seite 765.

Angeben einer Modellauswahlmethode

Klicken Sie im Auswahlfenster unter der Überschrift 'Modell' auf **Auswählen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Modellauswahlmethode	Wählen Sie aus der Dropdown-Liste eine Modellauswahlmethode aus. Wenn Sie keine Methode auswählen, wird das volle Modell verwendet. Je nach Modellauswahlmethode sind weitere Optionen für diese Methode verfügbar. Folgende Methoden stehen zur Verfügung: ■ **Volles Modell (keine Auswahl)** ist der Standard und ermöglicht keine Auswahl von Effekten. Dies ist das Modell, das beim Zuweisen der abhängigen und quantitativen (erklärenden) Variablen erstellt wird. ■ **Vorwärtsverfahren** startet ohne Effekte im Modell. Diese Methode berechnet für jede erklärende Variable *F*-Statistiken, die

Optionsname	Beschreibung
	den Beitrag der Variable wiedergeben, wenn diese ins Modell einbezogen wird. Die *p*-Werte für diese *F*-Statistiken werden mit dem Signifikanzniveau verglichen, das für den Einschluss einer Variable in das Modell angegeben wird. Standardmäßig beträgt dieser Wert `0,05`. Um das Signifikanzniveau zu ändern, geben Sie den Wert in das Textfeld **Für Aufnahme ins Modell** ein. Wenn keine *F*-Statistik einen größeren Wert als diesen aufweist, hält das Vorwärtsverfahren an. Andernfalls fügt das Vorwärtsverfahren die Variable mit der größten *F*-Statistik zum Modell hinzu. Das Vorwärtsverfahren berechnet dann erneut *F*-Statistiken für die Variablen, die nicht im Modell verwendet werden und das Bewertungsverfahren wird wiederholt. Die Variablen werden also solange nacheinander zum Modell hinzugefügt, bis keine verbleibende Variable mehr eine signifikante *F*-Statistik erzeugt. Eine Variable, die zum Modell hinzugefügt wurde, verbleibt im Modell. ■ Das **Rückwärtsverfahren** berechnet zuerst *F*-Statistiken für ein Modell, einschließlich aller erklärenden Variablen. Die Variablen werden anschließend solange nacheinander aus dem Modell gelöscht, bis alle im Modell verbleibenden Variablen *F*-Statistiken erzeugen. Das Signifikanzniveau wird im Textfeld Für Verbleiben im Modell angegeben. Standardmäßig beträgt dieser Wert `0,05`. Bei jedem Schritt wird die am wenigsten zum Modell beitragende Variable gelöscht. ■ **Schrittweise Auswahl** ist eine Variante des Vorwärtsverfahrens. Bei der schrittweisen Auswahl verbleiben die bereits enthaltenen Variablen jedoch nicht notwendigerweise im Modell. Wie auch bei der Vorwärtsselektion werden Variablen nacheinander zum Modell hinzugefügt, wenn eine signifikante *F*-Statistik vorliegt. Das Signifikanzniveau wird im Textfeld Für Aufnahme ins Modell angegeben. Nachdem eine Variable hinzugefügt wurde, werden alle bereits im Modell enthaltenen Variablen überprüft. Weist eine Variable keine signifikante *F*-Statistik auf, wird sie aus dem Modell gelöscht. Das Signifikanzniveau wird im Textfeld **Für Verbleiben im Modell** angegeben. Erst nach dieser Prüfung und nach Ausführung der notwendigen Löschvorgänge, kann eine andere Variable dem Modell hinzugefügt werden. Die schrittweise Auswahl ist beendet, wenn eine der folgenden Bedingungen erfüllt ist: □ keine Variable außerhalb des Modells besitzt eine signifikante *F*-Statistik und jede Variable im Modell ist hinsichtlich des für das Verbleiben im Modell Signifikanzniveaus signifikant. □ die Variable, die zum Modell hinzugefügt werden soll, ist die zuvor aus dem Modell gelöschte Variable. ■ Die **Vorwärts-Swap-Auswahl** ist eine Erweiterung basierend auf der Vorwärtsselektion. Bei diesem Verfahren werden Effekte

Optionsname	Beschreibung
	paarweise in das Modell auf- bzw. herausgenommen (Swap), durch die das Selektionskriterium verbessert wird. ■ Die **Least-Angle-Regression** startet ohne Effekte im Modell und fügt diese später hinzu. Die Parameterschätzer sind bei jedem Schritt kleiner im Vergleich zu den entsprechenden Kleinste-Quadrate-Schätzern. Enthält das Modell Klassifizierungsvariablen, werden diese Klassifizierungsvariablen geteilt. ■ Bei der **Lasso-Methode** werden Parameter hinzugefügt und gelöscht, basierend auf einer Version der Kleinsten-Quadrate (OLS, Ordinary Least Squares), wobei die Summe der absoluten Regressionskoeffizienten beschtänkt (constrained) ist. Enthält das Modell Klassifizierungsvariablen, werden diese Klassifizierungsvariablen geteilt.
Signifikanzniveaus	Beim Vorwärtsverfahren und der Least-Angle-Regression geben Sie ein Signifikanzniveau an, um eine quantitative (erklärende) Variable in das Modell aufzunehmen. Beim Rückwärtsverfahren geben Sie ein Signifikanzniveau an, damit eine quantitative Variable im Modell verbleibt. Bei den Methoden Schrittweise Auswahl und Lasso-Methode geben Sie beide Signifikanzniveaus an.
Modell mit dem besten Wert auswählen	Verwendet das Kriterium um (bei jedem Schritt des Auswahlprozesses) das Modell auszuwählen, mit dem der beste Wert erzielt wird. Wenn der optimale Wert für das angegebene Kriterium bei Modellen in mehr als einem Schritt auftritt, wird das Modell mit der geringsten Anzahl an Parametern ausgewählt. Wenn kein Kriterium angegeben wird, wird das Modell ausgewählt, das im letzten Schitt des Auswahlprozesses verwendet wird. Folgende Kriterien stehen zur Verfügung: ■ **AIC** - Akaikes Informationskriterium ■ **AICC** - korrigiertes Akaikes Informationskriterium ■ **BIC** - Schwarz-Bayessches Kriterium ■ **ADJRSQ** - korrigiertes R-Quadrat ■ **CP** - Mallows C_p-Statistik ■ **PRESS** - prognostizierte Quadratesumme der Residuen ■ **VALIDATE** - durchschittlicher Quadratefehler für die Validierungsdaten
Kriterien für das Beenden des Selektionsprozesses auswählen	Verwendet das Kriterium, um den Auswahlprozess zu beenden. Wenn Sie **Keine** angeben, wird der Auswahlprozess beendet, sobald kein geeigneter Kandidat zum Hinzufügen oder Entfernen gefunden werden kann oder wenn ein größenbasiertes Limit erreicht wurde. Folgende Kriterien stehen zur Verfügung: ■ **SL** - die Signifikanz des Scoretests

Optionsname	Beschreibung
	■ **AIC** - Akaikes Informationskriterium ■ **AICC** - korrigiertes Akaikes Informationskriterium ■ **BIC** - Schwarz-Bayessches Kriterium ■ **ADJRSQ** - korrigiertes R-Quadrat ■ **CP** - Mallows C_p-Statistik ■ **PRESS** - prognostizierte Quadratesumme der Residuen ■ **VALIDATE** - durchschittlicher Quadratefehler für die Validierungsdaten
Für das Modell erzwungene Effekte	Wählen Sie den Effekt aus, den Sie zum Modell hinzufügen möchten. Die Variable wird in das Feld **Für das Modell erzwungene Effekte** eingefügt. Sie können diese Variablen umordnen, indem Sie die Variable markieren und mit den Pfeiltasten verschieben. **Hinweis:** Diese Option ist nicht verfügbar, wenn Sie die Modellselektionsverfahren **Volles Modell (keine Auswahl)**, **Least-Angle-Regression** oder **Lasso-Methode** verwenden.

High-Performance - Lineare Regression: Festlegen von Modelloptionen

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Details zu Schätzern	Folgende Statistiken können Sie im Bereich **Details zu Schätzern** auswählen: ■ **Standardisierte Regressionskoeffizienten** erstellt standardisierte Regressionskoeffizienten. Ein standardisierter Regressionskoeffizient wird durch Dividieren eines Parameterschätzers durch das Verhältnis der Stichproben-Standardabweichung der abhängigen Variablen zur Stichproben-Standardabweichung des Regressors berechnet. ■ **Konfidenzgrenzen für Parameterschätzer** erfordert die oberen und unteren Konfidenzgrenzen für die Parameterschätzer. Standardmäßig werden die 95%-Grenzen berechnet.

Optionsname	Beschreibung
Kollinearitätsanalyse	Folgende Diagnosen stehen zur Auswahl: ■ **Toleranzwerte für Schätzer** erzeugt die Toleranzwerte für die Schätzer. Toleranz für einen Parameter ist definiert als 1 – R^2, wobei R^2 aus der Regression des Parameters zu allen anderen Parametern im Modell ermittelt wird. ■ **Varianzinflationswerte** - erzeugt die Varianzinflationswerte mit den Parameterschätzungen. Die Varianzinflation ist der Kehrwert der Toleranz. **Hinweis:** Diese Optionen sind nicht verfügbar, wenn Sie die Least-Angle-Regression oder die Lasso-Methode als Modellselektionsmethode verwenden.

High-Performance - Lineare Regression: Festlegen der Partitionsoptionen

Klicken Sie im Auswahlbereich auf **Partition**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Daten für Training, Validierung und Testen des Modells nicht partitionieren	legt fest, dass die Beobachtungen nicht partitioniert werden.
Proportionen für zufällige Zuweisung von Beobachtungen festlegen	legt fest, wie Beobachtungen in der Eingabedatei logisch in disjunkte Teilmengen für Modell-Training, Validierung und Test partitioniert werden. Sie können angeben, dass Anteile der Beobachtungen den Trainings- und Validierungsrollen zufällig zugewiesen werden. ■ **Trainings-Prozentsatz** - zeigt den Anteil der Beobachtungen in der Eingabedatei an, der zufällig zur Trainingsrolle zugewiesen wurde. Beobachtungen, die nicht der Validierungs- oder Testrolle zugewiesen werden, werden der Trainingsrolle zugewiesen. ■ **Validierungs-Prozentsatz** - zeigt den Anteil der Beobachtungen in der Eingabedatei an, der zufällig zur Validierungsrolle zugewiesen wurde.

Optionsname	Beschreibung
	■ **Test-Prozentsatz** - zeigt den Anteil der Beobachtungen in der Eingabedatei an, der zufällig zur Testrolle zugewiesen wurde.

High-Performance - Regression: Festlegen von Performance-Optionen

Klicken Sie im Auswahlbereich auf **Performance**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Standardeinstellungen für diese Anwendungsroutine verwenden	verwendet die Standardeinstellungen.
Benutzerdefinierte Einstellungen für diese Anwendungsroutine verwenden	ermöglicht die Angabe von Knoten, Grid-Host, Installationsverzeichnis und Datenserver für die Anwendungsroutine. ■ **Knoten** legt die Anzahl der Knoten in der verteilten Umgebung fest, vorausgesetzt die Daten werden nicht entlang der Datenbank verarbeitet. ■ **Grid Host** legt den Namen des Appliance-Host fest. ■ **Installationsverzeichnis** legt das Verzeichnis fest, in dem die gemeinsamen Bibliotheken für High-Performance Analytics installiert sind. ■ **Datenserver** legt den Namen des Servers auf dem Teradata-System fest, so wie er über die Datei "hosts" definiert wurde und wie er in der LIBNAME-Anweisung für Teradata angegeben ist.
Detaillierte Prozedurschritte anzeigen	erstellt eine Tabelle, in der die einzelnen Prozedurschritte dargestellt werden.

High-Performance - Regression: Festlegen von Vorhersageoptionen

Klicken Sie im Auswahlfenster auf **Vorhersagen**, um auf diese Optionen zuzugreifen.

Plot-Name	Beschreibung
Ausgabedaten speichern	legt fest, ob Prognosen und Diagnosestatistiken dauerhaft in einer Ausgabedatei gespeichert werden sollnen. ■ **Vorhersagen** speichert die vorhergesagte Wahrscheinlichkeit aller abhängigen Ausprägungen. Für eine abhängige Variable Y mit drei Ausprägungen, 1, 2 und 3, lauten die individuellen Wahrscheinlichkeiten Pr(Y=1), Pr(Y=2) und Pr(Y=3). ■ **Diagnosestatistiken** speichert DFBETAS-Diagnosestatistik, die Verschiebungsdiagnose der Konfidenzintervalle (C und CBAR), die *H*-Matrix und die Schätzung des linearen Prädiktors mit Standardfehlerschätzungen. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Weitere Statistiken	■ Wählen Sie **Residuen**, um folgende Statistiken zu speichern: □ Änderungen der Abweichung □ Abweichungsresiduum □ pearsonsches Residuum (Chi-Quadrat) □ die Änderung der Abweichung aufgrund der Löschung der individuellen Beobachtung □ die Änderung in der Chi-Quadrat-/Güte-der-Anpassung-Statistik aufgrund der Löschung einer individuellen Beobachtung ■ Die Auswahl von **Vorhersagegrenzen** schließt die obere und untere Grenze eines Konfidenzintervalls für jede Vorhersage und den erwarteten Mittelwert der abhängigen Variable in die Ausgabe ein.

Plot-Name	Beschreibung
Ausgabevariablen auswählen	legt die Spalten fest, die in der Ausgabedatei enthalten sein sollen.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

38

High-Performance - Logistische Regression

Allgemeines über die Anwendungsroutine 'High Performance - Logistische Regression' ... **269**

High-Performance - Logistische Regression: Zuweisen von Variablen zu Analyserollen ... **270**

High-Performance - Logistische Regression: Erstellen eines Modells ... **271**
- Angeben der Ausprägung ... 271
- Erstellen eines Modells ... 273
- Angeben einer Modellauswahlmethode ... 273

High-Performance - Logistische Regression: Festlegen von Modelloptionen ... **276**

High-Performance - Regression: Festlegen von Performance-Optionen ... **277**

High-Performance - Regression: Festlegen von Vorhersageoptionen ... **278**

Festlegen von Titeln und Fußnoten ... **279**

Allgemeines über die Anwendungsroutine 'High Performance - Logistische Regression'

Die Anwendungsroutine **High-Performance - Logistische Regression** passt logistische Regressionsmodelle für binäre, binomische und multinomiale Daten in der SAS High-Performance Analytics-Umgebung an. In der Anwendungsroutine High-Performance - Logistische Regression stehen zwei Parametrisierungen für Klassifizierungsvariablen zur Verfügung (GLM und Referenz). Im Gegensatz zu anderen Anwendungsroutinen, die noch weitere Parametrisierungen bieten, haben die Klassifizierungsvariablen in dieser Anwendungsroutine dieselbe Parametrisierung. Standardmäßig wird die GLM-Parametrisierung verwendet.

Die Anwendungsroutine High-Performance - Logistische Regression wurde speziell für den Einsatz innerhalb der verteilten High-Performance-Umgebung entwickelt. Standardmäßig werden mit ihr Berechnungen in mehreren Threads durchgeführt. In der Anwendungsroutine Logistische Regression wird hingegen ein einzelner Thread verwendet.

Hinweis: Beim Öffnen der Anwendungsroutine **High-Performance - Logistische Regression** kann es ggf. zu einer kleinen zeitlichen Verzögerung kommen, da SAS Enterprise Guide im Hintergrund ermittelt, welche Variablen binäre Werte aufweisen. Nach Abschluss der Analyse bietet die Anwendungsroutine Informationen über die Variable, die Sie der Rolle **Abhängige Variable** zuweisen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	HPLOGISTIC
Erforderliche SAS-Produkte	Base SAS, SAS High-Performance Analytics Server, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

High-Performance - Logistische Regression: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Abhängige Variable	gibt die Variable an, die als abhängige (Response) Variable verwendet wird. Sie können dieser Rolle nur eine Variable zuweisen.
Quantitative Variablen	gibt die Variablen an, die als unabhängige (erklärende) Variablen verwendet werden.

Optionsname	Beschreibung
Klassifizierungsvariablen	gibt die Klassifizierungsvariablen an, die in der Analyse verwendet werden sollen. Hierbei kann es sich um alphanumerische oder numerische Variablen handeln. Bei numerischen Variablen müssen Einzelwerte statt kontinuierlicher Werte vorliegen.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt eine Variable an, die als relative Gewichtung für die entsprechenden Werte der Variablen verwendet wird, die Sie als abhängige Variable ausgewählt haben. Negative oder fehlende Werte werden aus der Analyse ausgeschlossen.

High-Performance - Logistische Regression: Erstellen eines Modells

Angeben der Ausprägung

Klicken Sie im Auswahlfenster unter der Überschrift **Modell** auf **Ausprägung**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Typ abhängiger Ausprägungen	Welche Typen abhängiger Ausprägungen verfügbar sind, hängt von der Anzahl der Ausprägungen in der abhängigen Variable ab. ■ Wenn Ihre abhängige Variable nur zwei Ausprägungen hat, ist der Typ abhängiger Ausprägungen binär. ■ Wenn Ihre abhängige Variable mehr als zwei Ausprägungen hat, können Sie als Response-Typ Geordnet oder Ungeordnet auswählen.

Optionsname	Beschreibung
Modelltyp	Sie können aus folgenden Modelltypen auswählen: ■ **Logit** zeigt die Ausgabe aus der Log-Odds-Funktion (Standard) an. Das binäre Logit-Modell wird verwendet, wenn zwei abhängige Kategorien vorhanden sind. Das kumulative Logit-Modell wird verwendet, wenn mehr als zwei abhängige Kategorien vorhanden sind. ■ **Probit** zeigt die Ausgabe aus der inversen Integralfunktion für die Standard-Normalwahrscheinlichkeit an. Das binäre Probit-Modell wird verwendet, wenn zwei abhängige Kategorien vorhanden sind. Das kumulative Probit-Modell wird verwendet, wenn mehr als zwei abhängige Kategorien vorhanden sind. ■ **Komplementäres Log-Log** zeigt die Ausgabe aus der komplementären Log-Log-Funktion an. Das binäre komplementäre Log-Log-Modell wird verwendet, wenn zwei abhängige Kategorien vorhanden sind. Das kumulatve komplementäre Log-Log-Modell wird verwendet, wenn mehr als zwei abhängige Kategorien vorhanden sind. ■ **Log-Log** zeigt die Ausgabe aus der Log-Log-Funktion an. Das binäre Log-Log-Modell wird verwendet, wenn zwei abhängige Kategorien vorhanden sind. Das kumulatve Log-Log-Modell wird verwendet, wenn mehr als zwei abhängige Kategorien vorhanden sind. ■ **glogit** zeigt die Ausgabe der generalisierten Logit-Funktion an. Im generalisierten Logit-Modell wird jeder Kategorie ohne Verweis eine Verweiskategorie gegenübergestellt. **Hinweis:** Die Option ist nur dann verfügbar, wenn Sie **Ungeordnet** als Typ der abhängigen Ausprägung angegeben haben.
Abhängige Ausprägungen	Im Feld "Abhängige Ausprägungen" werden in aufsteigender Reihenfolge die formatierten Ebenen für die abhängige Variable aufgeführt.
Modell an Niveau anpassen	Sie können angeben, wie das Modell an ein Niveau angepasst wird. ■ Wenn Sie als Typ abhängiger Ausprägungen **Binär** angegeben haben, wählen Sie in der Dropdown-Liste **Modell an Niveau anpassen** die Ereigniskategorie für das binäre Response-Modell. ■ Wenn Sie als Typ abhängiger Ausprägungen **Ungeordnet** angegeben haben, wählen Sie in der Dropdown-Liste **Modell an Niveau anpassen** die Referenzkategorie für das verallgemeinerte Logit-Modell. Für jeden Logit wird einer Referenzkategorie eine Nicht-Referenzkategorie gegenübergestellt.

Optionsname	Beschreibung
	■ Wenn Sie als Typ abhängiger Ausprägung **Geordnet** angegeben haben, können Sie mit den Werten in der Dropdown-Liste **Modell an Niveau anpassen** die Reihenfolge der Response-Kategorien festlegen. Wenn Sie den ersten geordneten Wert auswählen, werden die Response-Kategorien in der Reihenfolge sortiert, in der sie im Feld **Abhängige Ausprägungen** aufgeführt sind. Wenn Sie den letzten geordneten Wert auswählen, werden die Response-Kategorien in der umgekehrten Reihenfolge sortiert.

Erstellen eines Modells

Sie können Modelle in mehreren SAS-Anwendungsroutinen erstellen. Weitere Informationen, siehe Anhang 3, “Erstellen eines Modells” auf Seite 765.

Angeben einer Modellauswahlmethode

Klicken Sie im Auswahlfenster unter der Überschrift 'Modell' auf **Auswählen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Modellauswahlmethode	Wählen Sie aus der Dropdown-Liste eine Modellauswahlmethode aus. Wenn Sie keine Methode auswählen, wird das volle Modell verwendet. Je nach Modellauswahlmethode sind weitere Optionen für diese Methode verfügbar. Folgende Methoden stehen zur Verfügung: ■ **Volles Modell (keine Auswahl)** ist der Standard und ermöglicht keine Auswahl von Effekten. Dies ist das Modell, das beim Zuweisen der abhängigen und quantitativen (erklärenden) Variablen erstellt wird. ■ **Vorwärtsverfahren** startet ohne Effekte im Modell. Diese Methode berechnet für jede erklärende Variable *F*-Statistiken, die den Beitrag der Variable wiedergeben, wenn diese ins Modell einbezogen wird. Die *p*-Werte für diese *F*-Statistiken werden mit dem Signifikanzniveau verglichen, das für den Einschluss einer Variable in das Modell angegeben wird. Standardmäßig beträgt dieser Wert `0,05`. Um das

Optionsname	Beschreibung
	Signifikanzniveau zu ändern, geben Sie den Wert in das Textfeld **Für Aufnahme ins Modell** ein. Wenn keine *F*-Statistik einen größeren Wert als diesen aufweist, hält das Vorwärtsverfahren an. Andernfalls fügt das Vorwärtsverfahren die Variable mit der größten *F*-Statistik zum Modell hinzu. Das Vorwärtsverfahren berechnet dann erneut *F*-Statistiken für die Variablen, die nicht im Modell verwendet werden und das Bewertungsverfahren wird wiederholt. Die Variablen werden also solange nacheinander zum Modell hinzugefügt, bis keine verbleibende Variable mehr eine signifikante *F*-Statistik erzeugt. Eine Variable, die zum Modell hinzugefügt wurde, verbleibt im Modell. ■ Das **Rückwärtsverfahren** berechnet zuerst *F*-Statistiken für ein Modell, einschließlich aller erklärenden Variablen. Die Variablen werden anschließend solange nacheinander aus dem Modell gelöscht, bis alle im Modell verbleibenden Variablen *F*-Statistiken erzeugen. Das Signifikanzniveau wird im Textfeld Für Verbleiben im Modell angegeben. Standardmäßig beträgt dieser Wert `0,05`. Bei jedem Schritt wird die am wenigsten zum Modell beitragende Variable gelöscht. ■ **Schrittweise Auswahl** ist eine Variante des Vorwärtsverfahrens. Bei der schrittweisen Auswahl verbleiben die bereits enthaltenen Variablen jedoch nicht notwendigerweise im Modell. Wie auch bei der Vorwärtsselektion werden Variablen nacheinander zum Modell hinzugefügt, wenn eine signifikante *F*-Statistik vorliegt. Das Signifikanzniveau wird im Textfeld Für Aufnahme ins Modell angegeben. Nachdem eine Variable hinzugefügt wurde, werden alle bereits im Modell enthaltenen Variablen überprüft. Weist eine Variable keine signifikante *F*-Statistik auf, wird sie aus dem Modell gelöscht. Das Signifikanzniveau wird im Textfeld **Für Verbleiben im Modell** angegeben. Erst nach dieser Prüfung und nach Ausführung der notwendigen Löschvorgänge, kann eine andere Variable dem Modell hinzugefügt werden. Die schrittweise Auswahl ist beendet, wenn eine der folgenden Bedingungen erfüllt ist: □ keine Variable außerhalb des Modells besitzt eine signifikante *F*-Statistik und jede Variable im Modell ist hinsichtlich des für das Verbleiben im Modell Signifikanzniveaus signifikant. □ die Variable, die zum Modell hinzugefügt werden soll, ist die zuvor aus dem Modell gelöschte Variable. ■ **Rückwärtsverfahren (schnell, ohne erneute Modellanpassung)** führt eine rückwärts gerichtete Eliminierung durch. Zu Beginn sind alle Effekte im Modell

Optionsname	Beschreibung
	enthalten; Effekte werden danach gelöscht, ohne dass das Modell erneut angepasst wird.
Signifikanzniveaus	■ Beim Vorwärtsverfahren geben Sie ein Signifikanzniveau an, um eine quantitative (erklärende) Variable in das Modell aufzunehmen. ■ Bei den Methoden Rückwärtsverfahren und Rückwärtsverfahren (schnell, ohne erneute Modellanpassung) geben Sie ein Signifikanzniveau an, um zu erreichen, dass eine quantitative Variable im Modell verbleibt. ■ Bei der Methode **Schrittweise Auswahl** geben Sie beide Signifikanzniveaus an.
Modell mit dem besten Wert auswählen	Verwendet das Kriterium um (bei jedem Schritt des Auswahlprozesses) das Modell auszuwählen, mit dem der beste Wert erzielt wird. Wenn der optimale Wert für das angegebene Kriterium bei Modellen in mehr als einem Schritt auftritt, wird das Modell mit der geringsten Anzahl an Parametern ausgewählt. Wenn kein Kriterium angegeben wird, wird das Modell ausgewählt, das im letzten Schitt des Auswahlprozesses verwendet wird. Folgende Kriterien stehen zur Verfügung: ■ **AIC** - Akaikes Informationskriterium ■ **AICC** - korrigiertes Akaikes Informationskriterium ■ **BIC** - Schwarz-Bayessches Kriterium
Kriterien für das Beenden des Selektionsprozesses auswählen	Verwendet das Kriterium, um den Auswahlprozess zu beenden. Wenn Sie **Keine** angeben, wird der Auswahlprozess beendet, sobald kein geeigneter Kandidat zum Hinzufügen oder Entfernen gefunden werden kann oder wenn ein größenbasiertes Limit erreicht wurde. Folgende Kriterien stehen zur Verfügung: ■ **SL** - die Signifikanz des Scoretests ■ **AIC** - Akaikes Informationskriterium ■ **AICC** - korrigiertes Akaikes Informationskriterium ■ **BIC** - Schwarz-Bayessches Kriterium
Für das Modell erzwungene Effekte	Wählen Sie den Effekt aus, den Sie zum Modell hinzufügen möchten. Die Variable wird in das Feld **Für das Modell erzwungene Effekte** eingefügt. Sie können diese Variablen umordnen, indem Sie die Variable markieren und mit den Pfeiltasten verschieben. **Hinweis:** Diese Option ist nicht verfügbar, wenn Sie die Modellauswahlmethode **Volles Modell (keine Auswahl)** wählen.

High-Performance - Logistische Regression: Festlegen von Modelloptionen

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Modellanpassungsbewertung	gibt folgende Modellanpassungsstatistiken an: ■ **Hosmer-und-Lemeshow-Test auf Güte der Anpassung** zeigt den Hosmer-und-Lemeshow-Test auf Güte der Anpassung für den Fall eines binären Response-Modells an. **Hinweis:** Diese Option ist nur verfügbar, wenn Sie der Rolle **Abhängige Variablen** eine binäre Variable zugewiesen haben. ■ **Verallgemeinertes R-Quadrat** zeigt einen verallgemeinerten Bestimmtheitskoeffizienten für das angepasste Modell an.
Konfidenzniveau	gibt an, ob die Konfidenzintervalle basierend auf einzelnen Wald-Tests berechnet werden sollen. Sie können auch einen Wert aus der Dropdown-Liste **Konfidenzniveau** wählen, um das Signifikanzniveau für die Konfidenzintervalle der Regressionsparameter oder Odd Ratios festzulegen.
Anzeige der Klassifizierungsausprägungsinformationen unterdrücken	unterdrückt die Anzeige der Tabelle Klassifizierungsausprägungsinformationen. Wählen Sie die Option **Berechnung der logistischen Regressionskoeffizienten unterdrücken**, um die Berechnung der Kovarianzmatrix und der Standardfehler der logistischen Regressionskoeffizienten zu unterdrücken. Wenn ein Modell mehrere Tausend Variablen enthält, ist die Erstellung der inversen Hesse-Matrix zur Ableitung der Kovarianzmatrix und der Standardfehler der Regressionskoeffizienten sehr zeitaufwändig.
Codierungsstil für Klassifizierungsvariablen	legt die Parameterisierungsmethode für die Klassifizierungsvariablen fest. ■ **GLM** wird für eine Less-Than-Full-Rank-Referenzzellen-Kodierung verwendet. ■ **Referenz** wird für eine Referenzzellen-Kodierung verwendet.

Optionsname	Beschreibung
CLASS-Variablen in Großbuchstaben umwandeln	wählen Sie diese Option, wenn Sie die Werte von Klassifizierungsvariablen vor dem Nivellieren in Großbuchstaben umwandeln möchten. Beispiel: Wenn Sie diese Option wählen und die Werte einer Klassifizierungsvariablen lauten 'a', 'A' und 'b', entsprechen 'a' und 'A' derselben Ausprägung und die Klassifizierungsvariable wird daher behandelt, als lägen nur zwei Werte vor: 'A' und 'B'.

High-Performance - Regression: Festlegen von Performance-Optionen

Klicken Sie im Auswahlbereich auf **Performance**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Standardeinstellungen für diese Anwendungsroutine verwenden	verwendet die Standardeinstellungen.
Benutzerdefinierte Einstellungen für diese Anwendungsroutine verwenden	ermöglicht die Angabe von Knoten, Grid-Host, Installationsverzeichnis und Datenserver für die Anwendungsroutine. ■ **Knoten** legt die Anzahl der Knoten in der verteilten Umgebung fest, vorausgesetzt die Daten werden nicht entlang der Datenbank verarbeitet. ■ **Grid Host** legt den Namen des Appliance-Host fest. ■ **Installationsverzeichnis** legt das Verzeichnis fest, in dem die gemeinsamen Bibliotheken für High-Performance Analytics installiert sind. ■ **Datenserver** legt den Namen des Servers auf dem Teradata-System fest, so wie er über die Datei "hosts" definiert wurde und wie er in der LIBNAME-Anweisung für Teradata angegeben ist.
Detaillierte Prozedurschritte anzeigen	erstellt eine Tabelle, in der die einzelnen Prozedurschritte dargestellt werden.

High-Performance - Regression: Festlegen von Vorhersageoptionen

Klicken Sie im Auswahlfenster auf **Vorhersagen**, um auf diese Optionen zuzugreifen.

Plot-Name	Beschreibung
Ausgabedaten speichern	legt fest, ob Prognosen und Diagnosestatistiken dauerhaft in einer Ausgabedatei gespeichert werden sollnen. ■ **Vorhersagen** speichert die vorhergesagte Wahrscheinlichkeit aller abhängigen Ausprägungen. Für eine abhängige Variable Y mit drei Ausprägungen, 1, 2 und 3, lauten die individuellen Wahrscheinlichkeiten Pr(Y=1), Pr(Y=2) und Pr(Y=3). ■ **Diagnosestatistiken** speichert DFBETAS-Diagnosestatistik, die Verschiebungsdiagnose der Konfidenzintervalle (C und CBAR), die *H*-Matrix und die Schätzung des linearen Prädiktors mit Standardfehlerschätzungen. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Weitere Statistiken	■ Wählen Sie **Residuen**, um folgende Statistiken zu speichern: □ Änderungen der Abweichung □ Abweichungsresiduum □ pearsonsches Residuum (Chi-Quadrat) □ die Änderung der Abweichung aufgrund der Löschung der individuellen Beobachtung □ die Änderung in der Chi-Quadrat-/Güte-der-Anpassung-Statistik aufgrund der Löschung einer individuellen Beobachtung ■ Die Auswahl von **Vorhersagegrenzen** schließt die obere und untere Grenze eines Konfidenzintervalls für jede Vorhersage und den erwarteten Mittelwert der abhängigen Variable in die Ausgabe ein.

Plot-Name	Beschreibung
Ausgabevariablen auswählen	legt die Spalten fest, die in der Ausgabedatei enthalten sein sollen.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

39

Histogramme

Allgemeines über die Anwendungsroutine 'Histogramme' ... **281**

Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen ... **282**

Histogramme: Auswählen der Verteilung ... **284**

Histogramme: Anpassen der Verteilung ... **285**
- Zugriff auf die Verteilungsoptionen ... 285
- Festlegen the Parameter für die Verteilung ... 285
- Festlegen der Kerndichteschätzer ... 285

Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung ... **286**
- Festlegen von Achsenoptionen ... 286
- Hinzufügen von Referenzlinien ... 286
- Einfügungen hinzufügen ... 287
- Festlegen weiterer Plotoptionen ... 287
- Festlegen von Optionen für Spezifikationsgrenzen ... 290

Prozessfähigkeitsanalyse: Festlegen der Analysetypen ... **290**

Festlegen von Titeln und Fußnoten ... **291**

Anzeigen von Eigenschaften ... **291**

Allgemeines über die Anwendungsroutine 'Histogramme'

Eine Prozessfähigkeitsanalyse vergleicht die Verteilung der Ausgabe eines Prozesses in statistischer Kontrolle mit den Spezifikationsgrenzen, um die Konsistenz zu bestimmen, mit der die Spezifikationen erfüllt werden. Histogramme werden bei der Prozessfähigkeitsanalyse verwendet, um die Verteilung von Messwerten aus einem Prozess in statistischer Kontrolle mit den Spezifikationsgrenzen zu vergleichen.

Sie können die Anwendungsroutine **Histogramme** zum Analysieren eines Fertigungsprozesses verwenden. Angenommen, ein Halbleiterhersteller produziert

Platinen, von denen Stichproben genommen werden, um festzustellen, ob die Dicke der Verkupferung innerhalb einer unteren Spezifikationsgrenze von 3,45 Milli-Inch und einer oberen Spezifikationsgrenze von 3,55 Milli-Inch liegt. Jeder Balken im Histogramm zeigt den Prozentwert der Beobachtungen und die gemessene Dicke an. Die unteren und oberen Spezifikationsgrenzen werden im Histogramm ebenfalls dargestellt, sodass Sie sehen können, welcher Prozentsatz der Beobachtungen innerhalb dieser Spezifikationsgrenzen liegt.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	CAPABILITY
Erforderliche SAS-Produkte	Base SAS, SAS/QC, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Analysevariablen	gibt die zu analysierenden numerischen Variablen an. Sie müssen dieser Rolle mindestens eine Variable zuweisen. Bei allen Plots, außer dem P-P-Plot, können Sie eine Obergrenze, einen Zielwert und eine Untergrenze im Bereich **Spezifikationsgrenzen für *Variablenname*** für jede Analysevariable angeben. Wenn Sie diese Grenzen angeben, enthält die Ausgabe automatisch eine Tabelle mit den Tests auf

Optionsname	Beschreibung
	Normalverteilung, eine Tabelle mit den Spezifikationsgrenzen und eine Tabelle mit den Prozessfähigkeitsindizes.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt die Variable an, die als relative Gewichtungsvariable verwendet werden soll. Wenn Sie dieser Rolle eine Variable zuweisen, werden mit dem Wert der Variablen in jeder Beobachtung gewichtete Statistiken berechnet. Sie können dieser Rolle nur eine Variable zuweisen. **Hinweis:** Sie sollten der Rolle **Relative Gewichtung** keine Variable zuweisen, wenn Sie ein Histogramm, Wahrscheinlichkeitsplot, P-P-Plot, Q-Q-Plot oder CDF-Plot erstellen. Falls Sie dieser Rolle ein Variable zugewiesen haben, müssen Sie eine eventuelle Ausgabe mit Vorsicht interpretieren.
Klassifizierungsvariablen (nur verfügbar für die Anwendungsroutine Histogramm)	gibt eine oder zwei Variablen an, die als Klassifizierungsvariablen dienen. Wenn Sie dieser Rolle eine Variable zuweisen, erstellt die Anwendungsroutine ein Vergleichshistogramm, mit dem Sie die Verteilung einer Analysevariablen über Ausprägungen der Klassifizierungsvariablen hinweg vergleichen können. **Hinweis:** Um vergleichende Wahrscheinlichkeitsplots oder vergleichende Q-Q-Plots zu erstellen, verwenden Sie die Anwendungsroutine Verteilungsanalyse.

Histogramme: Auswählen der Verteilung

Klicken Sie im Auswahlfenster unter der Überschrift 'Verteilungen' auf **Zusammenfassung**, um auf diese Optionen zuzugreifen.

Sie können mehrere theoretische Verteilungen in einem einzigen Histogramm darstellen. Wenn Sie die zu berechnende Verteilung ausgewählt haben, wählen Sie den Namen der Verteilung im Auswahlfenster aus, um die Parameterwerte und Linieneigenschaften für diese Verteilung anzugeben.

Optionsname	Beschreibung
Verteilungen	Verwenden Sie die Kontrollkästchen, um eine oder mehrere der folgenden geschätzten Dichtekurven auszuwählen: ■ **Normal** passt eine normale Dichte mithilfe eines Mittelwerts (My) und einer Standardabweichung (Sigma) an. ■ **Lognormal** passt eine Lognormal-Dichte mithilfe eines Skalierungsparameters (Zeta), eines Schwellenwertparameters (Theta) und eines Formparameters (Sigma) an. ■ **Exponentiell** passt eine exponentielle Dichte mithilfe eines Skalierungsparameters (Sigma) und eines Schwellenwertparameters (Theta) an. ■ **Weibull** passt eine Weibull-Dichte mit drei Parametern mit einem Skalierungsparameter (sigma), einem Schwellenwertparameter (theta) und einem Formparameter (c) an. ■ **Beta** passt eine Beta-Dichte mit einem Skalierungsparameter (Sigma), einem Schwellenwertparameter (Theta) und Formparametern (Alpha und Beta) an. ■ **Gamma** passt eine Gamma-Dichte mit einem Skalierungsparameter (Sigma), einen Schwellenwertparameter (Theta) und einen Formparameter (Alpha) an. ■ **Kern** passt die nichtparametrischen Kerndichteschätzungen an. **Hinweis:** Bei Vergleichshistogrammen sind nur die **Normalverteilung** und die **Kerndichteschätzungen** verfügbar.
Grafikstil	Wählen Sie den Grafikstil, der in der Ausgabe verwendet werden soll. Sie haben die Auswahl zwischen üblichen SAS-Grafiken und ODS Statistical Graphs (statistischen ODS-Grafiken). Weitere Informationen zu ODS-Grafiken finden Sie unter *SAS Output Delivery System: User's Guide*.

Histogramme: Anpassen der Verteilung

Zugriff auf die Verteilungsoptionen

Klicken Sie im Auswahlfenster unter der Überschrift **Verteilungen** auf den Namen der Verteilung, die Sie anpassen möchten, um auf diese Optionen zuzugreifen.

Hinweis: Diese Verteilungsoptionen sind nur durch vorherige Auswahl der Verteilung auf der Seite **Zusammenfassung** verfügbar, oder durch Aktivieren des Kontrollkästchens ***Name der Verteilung*** auf der Seite dieser Verteilung.

Festlegen the Parameter für die Verteilung

Für die Normal-, Lognormal-, Exponential-, Weibull-, Beta- und Gamma-Verteilungen können Sie die Parameter für jede Analysevariable oder für alle Analysevariablen angeben. Standardmäßig schätzt die Anwendungsroutine die Parameterwerte für alle Analysevariablen.

So legen Sie für jede Analysevariable Parameter fest:

1. Deaktivieren Sie das Kontrollkästchen **Verteilung auf alle Variablen anwenden**.
2. Aktivieren Sie das Kontrollkästchen der Variablen im Feld **Analysevariablen**.
3. Geben Sie die Parameterwerte an.

Festlegen der Kerndichteschätzer

Bei nichtparametrischen Kerndichteschätzern können Sie die Bandbreite (c) und die Kernfunktion (k) angeben. Standardmäßig schätzt die Anwendungsroutine eine Bandbreite, die den integrierten Quadratfehler des approximierten Mittelwerts MISE (Mean Integrated Square Error) minimiert. Wenn Sie die Bandbreite nicht schätzen möchten, deaktivieren Sie das Kontrollkästchen **MISE minimieren**. Um den standardisierten Bandbreitenparameter für Kerndichten festzulegen, geben Sie die Bandbreite in das Textfeld ein. Sie können bis zu fünf Werte angeben. Trennen Sie die Werte durch Leerstellen. Geben Sie zum Beispiel folgende Werte in das Textfeld ein: `0,5 1,0 1,5`.

Hinweis: Wenn Sie das Kontrollkästchen **MISE minimieren** nicht deaktivieren, wird MISE als einer der fünf Bandbreitenwerte gezählt.

Sie können auch die Kernfunktion angeben (normal, quadratisch oder dreieckig), die zur Berechnung der Kerndichteschätzung verwendet wird. Die Standardkernfunktion ist **Normal**.

Sie können Kernfunktionen für bis zu fünf Schätzungen angeben. Wenn Sie mehr als eine Kernfunktion angeben, ist die Reihenfolge der Funktionen immer: normal - quadratisch - dreieckig. Um Funktionen hinzuzufügen oder die Reihenfolge zu ändern, müssen Sie den Code der Anwendungsroutine ändern.

Wenn Sie mehr Kernfunktionen als Bandbreiten angeben, wird die letzte Bandbreite für die restlichen Schätzungen wiederholt. Wenn Sie z.B. 0,5 und 1,0 als Bandbreiten und normal, quadratisch und dreieckig als Kernfunktionen angeben, wird die Bandbreite 1,0 für die dreieckige Funktion wiederholt. Wenn Sie mehr Bandbreiten als Kernfunktionen angeben, wird die letzte Kernfunktion für die restlichen Bandbreiten wiederholt.

Wenn Sie z.B. 0,5, 1,0 und 1,5 als Bandbreiten und normal und quadratisch als Kernfunktionen angeben, wird die quadratische Funktion für die Bandbreite 1,5 verwendet.

Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung

Festlegen von Achsenoptionen

Klicken Sie im Auswahlfenster unter **Darstellung** ⇨ **Achsen** auf **Achsen**, um auf die Optionen zuzugreifen.

Folgende Optionen können angegeben werden:

- die Farbe der Achsen
- die Farbe der Achsenmarkierungen
- ein Etikett für die horizontalen und vertikalen Achsen

Hinzufügen von Referenzlinien

Klicken Sie im Auswahlfenster unter **Darstellung** ⇨ **Achsen** auf **Horizontal** bzw. **Vertikal**, um auf die Optionen zuzugreifen.

Um Referenzlinien zu verwenden, wählen Sie **Referenzlinien verwenden**. Sie können die Referenzlinie formatieren und die Position jeder Referenzlinie festlegen.

Einfügungen hinzufügen

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Einfügung**, um auf diese Optionen zuzugreifen.

Sie können einen Plot erweitern, indem Sie direkt in der Grafik ein Feld oder eine Tabelle (eine so genannte Einfügung) der beschreibenden Statistiken hinzufügen. Wenn Sie ein Vergleichshistogramm erstellen, erscheint in jedem Komponentenhistogramm eine Einfügung.

Wählen Sie **Einfügung einschließen**, um eine beschreibende Statistik zum Plot hinzuzufügen. Wählen Sie dann die Statistiken, die Sie in die Einfügung einschließen möchten.

Hinweis: Die Statistiken für die Berechnung der Prozessfähigkeitsindizes sind nur verfügbar, wenn Sie Spezifikationsgrenzen festgelegt und dann der Rolle **Analysevariablen** zugewiesen haben. Sie können Spezifikationsgrenzen für Histogramm, CDF-Plots, Wahrscheinlichkeitsplots und Q-Q-Plots erstellen. Weitere Informationen zur Berechnung der Prozessfähigkeitsindizes finden Sie in der Hilfe zur Prozedur CAPABILITY.

Festlegen weiterer Plotoptionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Optionen**, um auf diese Optionen zuzugreifen.

Die verfügbaren Plotdarstellungsoptionen hängen vom Typ des Plots ab, den die Anwendungsroutine erzeugt.

Optionsname	Beschreibung
Hintergrund	bestimmt die Hintergrundfarbe des Plots.
Balkenumriss und **Balkenfüllung**	legt die Farben der Balken in einem Histogramm fest. **Hinweis:** Diese Optionen sind nur für Vergleichshistogramme verfügbar.
Beide Achsen umkehren (Perzentile auf entgegengesetzte Achsen)	kehrt die Achsen auf dem Plot um. Wenn Sie diese Option auswählen, funktionieren die von Ihnen ausgewählten Achsenoptionen und Referenzlinien eventuell nicht mehr, da sich die Ausrichtung des Plots ändert.

Optionsname	Beschreibung
	Hinweis: Diese Option ist nur für Wahrscheinlichkeitsplots und Q-Q-Plots verfügbar.
Keine Standardlegenden	verhindert die Anzeige der Standardlegende mit dem Plot.
Midpoints verwenden	geben Sie Midpoints für das Histogramm an. Die in der Dropdown-Liste verfügbaren Optionen hängen davon ab, ob Sie ein Histogramm oder ein Vergleichshistogramm erstellen. Wenn Sie der Rolle **Klassifizierungsvariablen** eine Variable zuweisen, erstellen Sie ein Vergleichshistogramm. Für Vergleichshistogramme stehen folgende Optionen zur Verfügung: ■ **Midpoints angeben** - Ermöglicht das Angeben der Start- und End-Midpoints. Sie können auch das Schrittintervall festlegen, das für die Berechnung der anderen Midpoints verwendet werden soll. Der Midpoint ist der Mittelwert des Wertebereichs, der von den einzelnen Balken dargestellt wird. ■ **Schlüssel** - Bestimmt die Midpoints für die Daten in der Schlüsselzelle. Standardmäßig werden die Ausprägungen der Klassifizierungsvariablen von oben nach unten (links nach rechts) in aufsteigender Reihenfolge der internen (nicht formatierten) Werte der ersten Klassifizierungsvariablen angezeigt. Wenn Sie nur eine Klassifizierungsvariable angegeben haben, ist die Schlüsselzelle die Ausprägung, die als erstes in dieser Reihenfolge erscheint. Wenn Sie zwei Klassifizierungsvariablen angegeben haben, ist die Schlüsselzelle die Kombination der Ausprägungen von Variable 1 und Variable 2, die als erstes in dieser Reihenfolge erscheint. Die Auswahl der Schlüsselzelle bestimmt die einheitliche Horizontalachse, die für alle Zellen verwendet wird. Die Midpoint-Liste für die Schlüsselzelle wird dann wie erforderlich in beide Richtungen erweitert, bis sie die Daten in den übrigen Zellen abdeckt. ■ **Einheitlich** - Bestimmt die Anzahl der Midpoints abhängig von der Gesamt-Stichprobengröße. Wenn Sie der Rolle **Klassifizierungsvariablen** keine Variable zuweisen, erstellen Sie ein Histogramm. Für Histogramme sind die folgenden Optionen verfügbar: ■ **Midpoints angeben** - Ermöglicht das Angeben der Start- und End-Midpoints. Sie können auch das Schrittintervall festlegen, das für die Berechnung der anderen Midpoints verwendet werden soll. Der Midpoint ist der Mittelwert des

Optionsname	Beschreibung
	Wertebereichs, der von den einzelnen Balken dargestellt wird. ■ **Anzahl der Intervalle angeben** - Ermöglicht, die Anzahl der Intervalle (auch Histogrammintervalle genannt) für die Daten anzugeben. Standardmäßig basiert die Anzahl der angezeigten Intervalle auf dem Datenbereich. Sie können jedoch eine Maximalzahl anzuzeigender Intervalle angeben. **Hinweis:** Sie können nur dann Midpoints angeben, wenn Sie die Anwendungsroutine **Histogramme** verwenden.
Größe ben.def. Diagramm angeben (in Pixel)	legt die Bildgröße fest. Zum Vergleich wird die Gesamtbildschirmgröße angezeigt in den Feldern **Breite** und **Höhe** angezeigt.

Hinweis: Vergleichsoptionen sind für Vergleichshistogramme verfügbar. Sie können ein Vergleichshistogramm erstellen, indem Sie in der Anwendungsroutine **Histogramme** der Rolle **Klassifizierungsvariablen** eine Variable zuweisen.

Optionsname	Beschreibung
Zeilen und **Spalten**	Wenn Sie einer Analyserolle mindestens eine Klassifizierungsvariable zuweisen, können Sie über die Felder **Zeilen pro Seite** und **Spalten pro Seite** die Anordnung der Spalten und Zeilen im Vergleichshistogramm ändern. Wenn Sie eine Klassifizierungsvariable zuweisen, ist die Standardanordnung zwei Zeilen und eine Spalte pro Seite. Wenn Sie zwei Klassifizierungsvariablen zuweisen, ist die Standardanordnung zwei Spalten und zwei Zeilen pro Seite. Sie können auch über die Dropdown-Listen **Rahmenseite** und **Rahmenoberseite** Farben zum Füllen des Rahmenbereichs der Zeilenetiketten und Spaltenetiketten angeben.
Anzahl der Intervalle	Standardmäßig basiert die Anzahl der angezeigten Intervalle auf dem Datenbereich. Sie können jedoch eine Maximalzahl anzuzeigender Intervalle angeben. Sie können auch eine Standardabweichung angeben, so dass die Anzahl der angezeigten Intervalle von einem Bereich des Standardabweichungswerts über und unter dem Mittelwert der Daten abhängt.

Festlegen von Optionen für Spezifikationsgrenzen

Hinweis: Diese Optionen sind für Histogramme, Wahrscheinlichkeitsplot, Q-Q-Plots und CDF-Plots verfügbar.

Klicken Sie im Auswahlfenster unter **Darstellung** auf **Spezifikationsgrenzen**, um die untere Spezifikationsgrenze, das Ziel und die obere Spezifikationsgrenze anzugeben.

Hinweis: Die Optionen für **Spezifikationsgrenzen** sind nur verfügbar, wenn Sie eine oder mehrere Grenzen für mindestens eine Analysevariable angeben. Um eine Spezifikationsgrenze für eine Analysevariable anzugeben, klicken Sie im Auswahlfenster auf **Daten** und wählen dann die Analysevariable aus.

Prozessfähigkeitsanalyse: Festlegen der Analysetypen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

Standardmäßig enthält der Prozessfähigkeitsanalysebericht die folgenden Tabellen: eine mit Grundkonfidenzintervallen, eine mit Grundmaßen, eine mit Tests auf Lageparametern und eine mit Momenten.

Um eine Tabelle zum Bericht hinzuzufügen, wählen Sie die Tabelle im Feld **Tabellen** aus.

Um die Analyseoptionen festzulegen, wählen Sie die Tabelle im Feld **Tabellen** aus. Sind Optionen verfügbar, erscheint rechts ein Feld, in dem Sie die Optionen angeben können.

Um alle Tabellen mit beschreibenden Statistiken im Bericht zu unterdrücken, aktivieren Sie die Option **Beschreibende Statistiken und Tabellen für Prozessfähigkeitsindizes unterdrücken**. Folgende Tabellen werden durch Auswahl dieser Option nicht unterdrückt:

- Tabellen, die von der ausgewählten Verteilung erstellt wurden. Sie können diese Tabellen unterdrücken, indem Sie im Fenster für diese Verteilung das Kontrollkästchen **Verteilungstabellen unterdrücken** aktivieren.
- Tabellen, die mit der INTERVALS-Anweisung erstellt wurden. Um diese Tabellen zu unterdrücken, können Sie die Option NOPRINT angeben. Weitere Informationen zur Anweisung INTERVALS finden Sie in der Hilfe zur Prozedur CAPABILITY.

Um Ihre Ergebnisse in eine Ausgabedatei zu speichern, aktivieren Sie das Kontrollkästchen **Ausgabestatistiken in einer Datei speichern**. Die Datei enthält univariate Statistiken und Prozessfähigkeitsindizes.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

   Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine

- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

40

Urwertkarte

Allgemeines über die 'Urwertkarte' ... 293
Urwertkarte: Zuweisen von Variablen zu Analyserollen ... 294
Regelkarten: Festlegen von Regelgrenzen ... 295
Regelkarte: Auswählen der durchzuführenden Tests ... 297
Regelkarten: Festlegen von Optionen für die Plotdarstellung ... 298
Regelkarten: Speichern von Analyseergebnissen ... 298
Festlegen von Titeln und Fußnoten ... 299
Anzeigen von Eigenschaften ... 300

Allgemeines über die 'Urwertkarte'

Die Shewhart-Regelkarte ist ein Grafik- und Analysetool, mit dem bestimmt werden kann, ob ein Prozess innerhalb der statistischen Kontrolle liegt. Die Anwendungsroutine **Urwertkarte** erstellt Regelkarten für Einzelmessungen und gleitende Spannweiten. Diese Regelkarten eignen sich nur, wenn eine Messung für jede Untergruppenstichprobe verfügbar ist und wenn die Messungen unabhängig und normalverteilt sind.

Sie können diese Anwendungsroutine auch zum Analysieren eines Fertigungsprozesses verwenden. Angenommen, ein Luftfahrtunternehmen, das Düsentriebwerke herstellt, misst den Innendurchmesser der vorderen Fläche der einzelnen Triebwerke (in Zentimetern). Die Durchmessermaße von 20 Motoren werden in einer SAS-Datei gespeichert. Jeder Punkt in der Urwertkarte gibt den Innendurchmesser eines bestimmten Triebwerks an. Die einzelnen Punkte auf dem Moving-Range-Diagramm geben den Bereich der beiden letzten Messungen an. Wenn alle Urwerte und Moving Ranges innerhalb der Regelgrenzen liegen, können Sie davon ausgehen, dass sich der Prozess innerhalb der statistischen Kontrolle befindet.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SHEWHART
Erforderliche SAS-Produkte	Base SAS, SAS/QC, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

Urwertkarte: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Prozessmaß	gibt die zu analysierenden Prozessvariablen an. Diese Variablen enthalten die Einzelmaße. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Maß-Identifier	gibt die Variable an, welche die Untergruppen in den Daten identifiziert. Die Werte der Variablen geben an, wie die Beobachtungen in der Eingabetabelle in rationale Untergruppen gegliedert werden. Dies sind die üblichen Werte: ■ die Indizes, welche die Reihenfolge angeben, in der die Untergruppenstichproben gesammelt wurden ■ die Daten oder Zeiten, zu denen die Untergruppenstichproben gesammelt wurden ■ die Etiketten, welche die Untergruppenstichproben eindeutig identifizieren. Ein Maß-Identifier ist erforderlich. Sie können dieser Rolle nur genau eine Variable zuweisen.

Optionsname	Beschreibung
	Um die Daten nach den identifizierten Untergruppen zu sortieren, aktivieren Sie das Kontrollkästchen **Nach Untergruppe sortieren**.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.
Blockvariablen	gibt bis zu zwei Variablen an, welche die Daten in Blöcke aufeinanderfolgender Untergruppen gruppieren. Diese Blöcke werden in einer Legende beschriftet. Jede Blockvariable liefert eine Etikettenausprägung in der Legende.

Regelkarten: Festlegen von Regelgrenzen

Klicken Sie im Auswahlfenster auf **Regelgrenzen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Sigma-Grenzen	gibt die Breite der Regelgrenzen als Mehrfaches des Standardfehlers (Sigma) der beschreibenden Untergruppenstatistik an, die im Diagramm dargestellt wird. Die Breite muss positiv sein. Der Standardwert für das Mehrfache ist 3.
Berechnungsmethode auswählen	gibt die Berechnungsmethode für die Regelgrenzen an. Folgende Optionen können ausgewählt werden: ■ **Regelgrenzen aus den aktiven Daten berechnen** ■ **Regelgrenzen aus der ausgewählten Datei berechnen** Klicken Sie auf **Durchsuchen**, um die Datei anzugeben, die die Regelgrenzen bzw. die Parameter zu deren

Optionsname	Beschreibung
	Berechnung enthält. Wählen Sie die gewünschte Datei aus und klicken Sie auf **Öffnen**. **Hinweis:** Bei der ausgewählten Datei muss es sich um eine LIMITS= Datei handeln. Weitere Informationen über die in einer LIMITS= Datei erforderlichen Variablen finden Sie in der Prozedurendokumentation zur erstellten Regelkarte. Wenn Sie die beschreibenden Statistiken für alle Untergruppen darstellen möchten, aktivieren Sie das Kontrollkästchen **Beschreibende Statistiken für alle Untergruppen darstellen, unabhängig davon, ob die Stichprobengröße der Untergruppe der nominalen Stichprobengröße für Regelgrenzen entspricht**. Sie können diese Statistiken beispielsweise darstellen, wenn fast alle Untergruppen eine gemeinsame Stichprobengröße haben und wenn Sie die festen (anstelle der variablen) Regelgrenzen, die der nominalen Stichprobengröße *n* entsprechen, anzeigen möchten. Wenn Sie diese Option auswählen und sich Ihre Stichprobengrößen für die Untergruppen deutlich unterscheiden, ist die Interpretation der Regelgrenzen nur für die Untergruppen sinnvoll, deren Stichprobengröße *n* entspricht. Um spezielle Marker für die Datenpunkte hinzuzufügen, bei denen die Stichprobengröße ungleich `n` ist, aktivieren Sie das Kontrollkästchen **Spezielle Marker für Punkte hinzufügen, die Stichprobengrößen entsprechen, welche ungleich nominalen Stichprobengrößen für feste Regelgrenzen sind**. ■ **Regelgrenzen angeben** Geben Sie in den Feldern **Oben**, **Mitte** und **Unten** die obere Regelgrenze, die Mittellinie bzw. die untere Regelgrenze an. Wenn die Stichprobengrößen Ihrer Untergruppe variieren, müssen Sie eine nominale Stichprobengröße angeben. **Hinweis:** Wenn Sie der Rolle **Analyse gruppieren nach** Variablen zuweisen, können Sie keine Werte für Regelgrenzen angeben.

Regelkarte: Auswählen der durchzuführenden Tests

Klicken Sie im Auswahlfenster auf **Tests**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Tests wenden nicht die Regelgrenzen an, die keine Drei-Sigma-Grenzen sind oder die mit der Untergruppenstichprobengröße variieren.

Optionsname	Beschreibung
Test auswählen	Fordern Sie über die Kontrollkästchen einen oder mehrere Tests auf besondere Ursachen an, die auch als Ausführungstests, Mustertests und Western-Electric-Regeln bekannt sind. Diese Tests entdecken bestimmte nichtzufällige Muster in den Punkten, die in der primären Regelkarte dargestellt werden. Das Vorkommen eines nichtzufälligen Musters wird als Signal bezeichnet. Die Tests können auf das Vorhandensein einer besonderen Variationsursache hinweisen. Wenn Sie die Tests auswählen, erscheint eine Beschreibung des jeweiligen Tests im Feld **Beschreibung**.
Etikett	Geben Sie in den Textfeldern die Etiketten für Punkte an, bei denen die Tests auf besondere Ursachen positiv sind. Die Etiketten dürfen maximal 16 Zeichen lang sein. Jedes Etikett erscheint bei allen Punkten im Test, an denen ein Signal vorliegt.
Text zum Identifizieren der Testsignale	■ Wählen Sie über die Dropdown-Liste eine Farbe für die angegebenen Etiketten aus. ■ Markieren Sie das Kontrollkästchen **Zonenlinien anzeigen**, um Linien einzuschließen, welche die Zonen A, B und C in einem primären Diagramm kennzeichnen. Diese Zonen werden gebildet, indem die Intervalle zwischen den Regelgrenzen in sechs gleichmäßig verteilte Unterintervalle unterteilt werden. ■ Verwenden Sie das Kontrollkästchen **3 Sigma-Grenzen überschreiben**, um nach speziellen Ursachen zu suchen, wenn Sie eine andere Sigma-Regelgrenze als das Standardvielfache von 3 angeben. Um die Sigma-Grenze zu ändern, klicken Sie im Auswahlfenster auf **Regelgrenzen**.

Optionsname	Beschreibung
	■ Über das Kontrollkästchen **Test auf überlappende Punktmuster anwenden** wenden Sie Tests auf besondere Ursachen auf überlappende Punktmuster an.

Regelkarten: Festlegen von Optionen für die Plotdarstellung

Einige der Darstellungsoptionen sind für unterschiedliche Regelkarten-Anwendungsroutinen gleich. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Festlegen von Achsenoptionen” auf Seite 759
- “Hinzufügen von Referenzlinien” auf Seite 760
- “Festlegen weiterer Plotoptionen” auf Seite 760
- “Festlegen von Blockvariablenoptionen” auf Seite 762

Regelkarten: Speichern von Analyseergebnissen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

Um die Analyseergebnisse zu speichern, können Sie Ausgabedateien mit diversen Statistiken speichern. Markieren Sie die Kontrollkästchen der Ausgabedaten, die Sie permanent speichern möchten.

Optionsname	Beschreibung
Datei mit Untergruppenstatistiken und Regelgrenzen	erstellt eine Ausgabedatei, welche die im Diagramm dargestellten Informationen enthält. Dies beinhaltet die Werte der Untergruppenvariablen und ihre entsprechenden beschreibenden Statistiken und Regelgrenzen. **Hinweis:** Diese Option erstellt eine Datei OUTTABLE=. Sie können diese Datei mit der Option TABLE= aus der

Optionsname	Beschreibung
	SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Untergruppenstatistiken	erstellt eine Ausgabedatei mit den beschreibenden Statistiken der Untergruppe. **Hinweis:** Diese Option erstellt eine Datei mit OUTHISTORY=. Sie können diese Datei mit der Option HISTORY= aus der SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Regelgrenzen	erstellt eine Ausgabedatei, welche die Regelgrenzen speichert. Sie können diese Datei verwenden, wenn Sie die Berechnungsmethode für die Regelgrenzen später in dieser Anwendungsroutine auswählen.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

   Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

41

Sterbetafeln

Allgemeines über die Anwendungsroutine 'Sterbetafeln' 301
Sterbetafeln: Zuweisen von Variablen zu Analyserollen 302
Sterbetafeln: Festlegen von Methodenoptionen 304
Sterbetafeln: Erzeugen von Plots 305
Sterbetafeln: Festlegen von Ergebnisoptionen 306
Festlegen von Titeln und Fußnoten 307
Anzeigen von Eigenschaften 308

Allgemeines über die Anwendungsroutine 'Sterbetafeln'

Die Anwendungsroutine **Sterbetafeln** berechnet nichtparametrische Schätzer der Überlebensverteilung von Daten, die aufgrund vorzeitigen Datenverlusts oder der Beendigung der Studie rechtszensiert sein könnten. Bei solchen Daten wissen Sie nur, dass die Überlebenszeit über dem angegebenen Wert lag. Im Allgemeinen vergleichen Sie Überlebenskurven, um zu bestimmen, ob zwei Stichproben aus derselben Überlebensverteilung stammen.

Diese Anwendungsroutine berechnet Rangtests und einen Likelihood Ratio-Test auf Homogenität der Überlebensfunktionen über Schichten hinweg. Die Anwendungsroutine erzeugt auch Plots, mit denen Sie die Überlebenskurven untersuchen können.

Eventuell möchten Sie diese Anwendungsroutine einsetzen, um die Überlebensverteilung für eine bestimmte Grundgesamtheit zu untersuchen. Angenommen, Sie möchten die Überlebenszeit von 40, mit einem Krebserreger infizierten Ratten in einem kleinen Zufallsversuch untersuchen und haben zu diesem Zweck ein oder zwei Behandlungsgruppen zugewiesen. Die Überlebenszeit ist die Zeit von der Randomisation bis zum Tod. Das Ereignis von Interesse ist der Tod durch den Krebs, der durch den Krebserreger herbeigeführt wurde. Mithilfe der

Anwendungsroutine **Sterbetafeln** können Sie bestimmen, ob sich die Überlebensverteilungen zwischen den beiden Behandlungsgruppen unterscheiden.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	LIFETEST
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Sterbetafeln: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Überlebenszeit	gibt die als Überlebenszeitvariable verwendete Variable an. Die Überlebenszeiten werden oft Fehlerzeiten genannt, während Ereigniszeiten nicht zensierte Überlebenszeiten sind. Überlebenszeit ist erforderlich. Sie können dieser Rolle nur eine Variable zuweisen.
Zensurvariable	gibt die Variable an, die als Zensorvariable verwendet wird. Wählen Sie im Feld **Rechts zensierende Werte** die Werte dieser Variablen aus, die angeben, ob eine Überlebenszeit zensiert wird. Die Werte der Zensorvariablen sollten numerische, nichtfehlende Werte sein.

Optionsname	Beschreibung
	Die Zensurvariable ist nicht erforderlich. Falls Sie sie jedoch verwenden, können Sie nur eine Variable angeben, und Sie müssen die Zensorwerte angeben.
Schichtvariablen	gibt die Variablen an, deren Werte die Schichtausprägungen bestimmen. Wenn Sie keine Endpunkte für die numerischen Variablen angeben, werden die Schichten durch die Kombination der Ausprägungen der eindeutigen Werte der Schichtvariablen bestimmt. Sie können das Kontrollkästchen markieren, um fehlende Werte als gültige Schichtausprägungen zuzulassen. Bei numerischen Schichtvariablen können Sie die Schichtintervalle angeben, indem Sie im Bereich **Schichtausprägungen** auf **Intervalle angeben** klicken. Anschließend definieren Sie die Intervalle, indem Sie Endpunkte hinzufügen. Standardmäßig bilden die eindeutigen Werte Schichtausprägungen. Geben Sie die Endpunkte an, welche die Schichtung bestimmen. Das Intervall unter einem Endpunkt enthält keinen Endpunkt, das Intervall über dem Endpunkt enthält den Endpunkt. Um Endpunkte für eine Variable anzugeben, geben Sie die Endpunktwerte ein und klicken auf **Hinzufügen**. Die Endpunkte sind nicht auf ganzzahlige Werte beschränkt. Sie können einzelne, durch Komma oder Leerstellen getrennte Werte oder Bereiche eingeben. Geben Sie beispielsweise `5 10 15` ein, um eindeutige Schichten für Werte zu erzeugen, die in die Intervalle {`[-infinity,5)`, `[5,10)`, `[10,15)`, `[15,infinity)`} fallen. Der Bereich `5 ... 10` erzeugt eindeutige Schichten für Werte, die in die folgenden Intervalle fallen: {`[-infinity,5)`, `[5,6)`, `[6,7)`, `[7,8)`, `[8,9)`, `[9,10)`, `[10,infinity)`} Wenn Sie dieser Rolle mehrere Variablen zuweisen, definieren Sie für jede Variable die Intervalle separat.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Test-Kovariate	gibt eine Liste numerischer (stetiger) Kovariate an, die beim Assoziationstest mit gleichen Ausfallzeiten verwendet werden. Es werden zwei Sätze von Rangstatistiken berechnet. Die Statistiken und ihre Varianzen werden über alle Schichten zusammengefasst. Univariate (marginale) Teststatistiken werden für alle Kovariate angezeigt. Teststatistiken werden auch für Joint-Effekte der Kovarianzen angezeigt.

Optionsname	Beschreibung
ID-Variablen	gibt eine oder mehrere Variablen an, die als ID-Variablen dienen. Die Werte von ID-Variablen werden verwendet, um die Beobachtungen der Produktgrenzen-Überlebensschätzer zu beschriften.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die Sie bereits als Analyserolle ausgewählt haben.

Sterbetafeln: Festlegen von Methodenoptionen

Klicken Sie im Auswahlfenster auf **Methode**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Schätzmethode	Um eine Schätzmethode anzugeben, klicken Sie auf die Schaltfläche **Produktgrenze (Kaplan-Meier)** oder **Sterbetafel**. Beide Methoden erzeugen ähnliche Ergebnisse. Die Produktgrenzmethode (auch Kaplan-Meier-Methode genannt) basiert jedoch auf tatsächlichen Überlebenszeiten, und die Sterbetafelmethode gruppiert Überlebenszeiten in Intervalle. Bei einer sehr großen Stichprobengröße kann es hilfreich sein, die Überlebenszeiten in Intervalle zu gruppieren und die Sterbetafelanalyse durchzuführen. Die Standardmethode ist Produktgrenze.
Intervalle	Wenn Sie **Produktgrenze** als Schätzmethode auswählen, können Sie die Anzahl der angezeigten Überlebensschätzer verringern, indem Sie für jeden Intervall Endpunkte definieren. Die Anwendungsroutine zeigt nur die Schätzer für die kleinste Zeit innerhalb der einzelnen angegebenen Intervalle an.

Optionsname	Beschreibung
	Wenn Sie **Sterbetafel** als Schätzmethode wählen, können Sie Angaben zum Aufbau der Intervalle machen, die zur Berechnung der Überlebensfunktion verwendet werden. Standardmäßig werden 10 Intervalle verwendet. Sie können einen anderen Wert angeben, indem Sie auf **Nummer** klicken und den gewünschten Wert eingeben. Sie können Intervalle auch erstellen, indem Sie die Intervallbreite angeben. Wählen Sie **Breite** und geben Sie Ihren Wert ein. Wenn Sie exakte Endpunkte definieren möchten, um die Intervalle zu bilden, klicken Sie auf **Endpunkte**.
Konfidenzniveau für Überlebensfunktion	Sie können einen Wert für das Konfidenzniveau eingeben oder aus der Dropdown-Liste auswählen. Der Standardwert ist ein Konfidenzniveau von 95%.

Sterbetafeln: Erzeugen von Plots

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Verwenden Sie diese Optionen, um Plots für geschätzte Überlebensfunktionen, Hazard-Funktionen und Wahrscheinlichkeitsdichtefunktionen gegenüber der Zeit zu erstellen.

Optionsname	Beschreibung
Überlebensfunktion splot anzeigen	Diese Funktion erstellt einen Plot der Überlebensverteilungsfunktion. Sie beschreibt die Lebensdauer der gewünschten Grundgesamtheit. ■ Um Konfidenzgrenzen einzufügen, aktivieren Sie das Kontrollkästchen **Punkteweise Konfidenzintervalle einschließen**. Die Konfidenzgrenzen für den Plot werden über die Einstellung Methoden des jeweiligen Konfidenzintervalls gesteuert. Der Standardwert ist ein Konfidenzniveau von 95%. ■ Um alle Plots für verschiedene Schichten in einem einzelnen Plot mit einer Legende darzustellen, aktivieren Sie das Kontrollkästchen **Schichten in Einzelplot überlagern**. Wenn Sie der Rolle **Schichtvariable** eine Variable zugewiesen haben und das Kontrollkästchen **Überlebensfunktionsplot anzeigen** aktiviert ist, werden standardmäßig alle Plots für verschiedne Schichten in einem einzigen Plot dargestellt.

Optionsname	Beschreibung
	Hinweis: Das Kontrollkästchen **Schichten in Einzelplot überlagern** lässt sich nur aktivieren, wenn einer Variablen die Rolle **Schichtvariablen** zugewiesen wurde. Wurde mehr als eine Variable zugewiesen, können Sie diese Option nicht auswählen.
Probit(1 - Überlebensfunktion)	Diese Option erstellt einen Plot der Probit von 1 minus der geschätzten Überlebensfunktion gegenüber dem Log der Zeit. Dieser Plottyp ist besonders geeignet für die Evaluierung von Lognormal-Verteilungen. Bei einem geeigneten Lognormal-Modell, ist die Kurve näherungsweise linear.
- Log(Überlebensfunktion)	Diese Option erstellt einen Plot des negativen Log der geschätzten Überlebensfunktion gegenüber der Zeit. Dieser Plottyp ist besonders geeignet für die Evaluierung von Exponentialverteilungen. Bei geeignetem Exponentialmodell ist die Kurve näherungsweise linear durch den Ursprung.
Log(-log(Überlebensfunktion))	Diese Option erstellt einen Plot des Log des negativen Log der geschätzten Überlebensfunktion gegenüber dem Log der Zeit. Dieser Plottyp ist besonders geeignet für die Evaluierung einer Weibull-Verteilung. Bei geeignetem Weibull-Modell ist die Kurve näherungsweise linear.
Schätzung der Sterbetafel	Um Plots der Hazard- und Wahrscheinlichkeitsdichtefunktion gegenüber der Zeit zu erstellen, müssen Sie zunächst die Sterbetafelmethode auswählen, um die Überlebensfunktionsschätzwerte zu berechnen. Anschließend können Sie das Kontrollkästchen **Hazard-Funktion** oder **Wahrscheinlichkeitsdichtefunktion** aktivieren.

Sterbetafeln: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Ausgabedaten speichern	erstellt Ausgabedateien mit den Überlebensschätzern und diversen Statistiken.

Optionsname	Beschreibung
	■ Aktivieren Sie das Kontrollkästchen **Überlebensschätzer**, um eine Ausgabedatei mit den Schätzern der Überlebensfunktion und der entsprechenden Konfidenzgrenzen für alle Schichten zu erstellen. ■ Aktivieren Sie das Kontrollkästchen **Statistiken**, um eine Ausgabedatei mit der allgemeinen Chi-Quadrat-Teststatistik für Assoziation mit gleicher Ausfallzeit für die Variablen in den **Testkovariatenrollen** zu erstellen, die Sie ausgewählt haben. Diese Datei enthält auch die Werte der univariaten Rangstatistiken und ihre geschätzte Kovarianzmatrix. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Anzeige für gesamte Ergebnisausgabe unterdrücken	Bewirkt, dass keine Ergebnisse angezeigt werden. Wenn Sie dieses Kontrollkästchen aktivieren, sind die Optionen im Plots-Fenster nicht verfügbar. Die Ergebnisse werden standardmäßig angezeigt.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

42

Liniendiagramm

Allgemeines über die Anwendungsroutine 'Liniendiagramm' ... 309
Liniendiagramm: Auswählen eines Diagrammtyps ... 310
Liniendiagramm: Zuweisen von Variablen zu Analyserollen ... 311
Liniendiagramm: Festlegen von Darstellungsoptionen ... 313
Festlegen weiterer Plotoptionen ... 313
Weitere Darstellungsoptionen ... 314
Festlegen von Titeln und Fußnoten ... 314
Anzeigen von Eigenschaften ... 315

Allgemeines über die Anwendungsroutine 'Liniendiagramm'

Die Anwendungsroutine **Liniendiagramm** erstellt Liniendiagramme, Streuungsdiagramme, Spline-Plots, Nadeldiagramme, Stufendiagramme, Regressions-Plots, Glättungsplots, STD-Plots, Lagrange-Interpolationsplots oder Überlagerungsplots, die die mathematischen Beziehungen zwischen Variablen durch Aufzeigen von Trends oder Mustern in den Datenpunkten anzeigen.

Beispielsweise kann ein Liniendiagramm anzeigen, ob eine Beziehung zwischen den Werbungskosten und dem Umsatz für ein bestimmtes Produkt besteht.

Weitere Informationen, siehe “Auswählen des Grafikausgabeformats” auf Seite 6.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GPLOT

Name der Anforderung	Prozedur- und Produktnamen
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Liniendiagramm: Auswählen eines Diagrammtyps

Klicken Sie im Auswahlfenster auf **Liniendiagramm**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Liniendiagramm	erstellt ein Diagramm, in dem alle Datenpunkte durch Geraden verbunden sind. Die Punkte werden in der Reihenfolge verbunden, in der sie in der Eingabedatei vorkommen.
Spline-Plot	gibt an, dass deine Splineroutine für die Interpolation der Linie verwendet wird.
Nadeldiagramm	Zieht eine vertikale Linie von jedem Datenpunkt zu einer horizontalen Linie am 0-Wert auf der vertikalen Achse oder am Minimalwert auf der vertikalen Achse, wenn dieser größer als 0 ist. Die horizontale Linie wird automatisch angezeigt.
Stufendiagramm	Legt fest, dass die Daten mit einer Stufenfunktion dargestellt werden. Im Diagramm werden die Stufen durch eine vertikale Linie verbunden. Die Daten werden vor dem Erstellen des Plots durch die unabhängige Variable sortiert.
Streuungsdiagramm mit Regressionslinie	Erstellt ein Streuungsdiagramm und passt die lineare Regressionslinie für die Daten optimal an. Standardmäßig werden Regressionslinien nicht durch Diagrammursprungspunkte erzwungen, und Konfidenzgrenzen werden nicht angezeigt.
Glättungsplot	Legt fest, dass eine glatte Linie mithilfe einer Splineroutine an die Daten angepasst wird. Diese Interpolationsmethode wird zum Glätten verrauschter Daten verwendet. Die Punkte auf dem Plot liegen nicht unbedingt auf der Linie. Die Daten

Optionsname	Beschreibung
	werden vor dem Erstellen des Plots durch die unabhängige Variable sortiert.
Standardabweichungsdiagramm	gibt an, dass eine durchgehende Linie den Mittelwert der Y-Variablen mit zwei Standardabweichungen für jeden Wert der X-Variablen verbindet.
Lagrange-Interpolationsplot	Legt fest, dass ein Lagrange-Polynomial des ersten Grades verwendet wird, um die Plotlinie zu glätten. Die Daten werden vor dem Erstellen des Plots durch die unabhängige Variable sortiert.
Mehrere vertikale sich überlagernde Liniendiagramme	erstellt ein separates Diagramm von jeder Y-Variablen vs. derselben X-Variablen. Alle Diagramme werden auf derselben Grafik angezeigt.
Mehrere Liniendiagramme nach Gruppenspalte	erstellt ein separates Diagramm für jeden Einzelwert der Gruppenvariablen. Alle Diagramme werden auf derselben Grafik angezeigt.
Streuungsdiagramm	Erstellt einen Streuungsplot, der die Beziehung zwischen zwei oder drei Variablen aufzeigt, in dem die Muster oder Konzentrationen von Datenpunkten dargestellt werden.

Liniendiagramm: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Horizontal	Die Spalte, die Sie dieser Rolle zuweisen, ist die horizontale oder X-Achsenvariable für das Diagramm.
Vertikal	Die Spalte, die Sie dieser Rolle zuweisen, ist die vertikale oder Y-Achsenvariable für das Diagramm. So erstellen Sie eine statistische Zusammenfassung der eindeutigen Werte der horizontalen Variablen. Wählen Sie die Variable unter der Rolle **Vertikal** aus. Wählen Sie **Zusammenfassung der einzelnen horizontalen Werte** und anschließend die Zusammenfassungsfunktion. Wenn z. B. Ihre Eingabedaten die Umsätze und Ausgaben jedes Landes nach Regionen aufführen. Weisen Sie der Variablen 'Country' die Rolle **Horizontal** und der Variablen 'Sales' die Rolle **Vertikal** zu. Wenn Sie die Summe der Umsätze für jedes Land darstellen möchten, aktivieren Sie das Kontrollkästchen und wählen aus der Dropdown-Liste **Summe** aus. Die Anwendungsroutine fügt die regionalen Umsätze für ein Land hinzu und stellt in der Endausgabe einen einzelnen Umsatzwert für jedes Land dar.
Vertikal (Rechts)	Die Spalte, die Sie dieser Rolle zuweisen, ist die vertikale bzw. Y-Achsenvariable für die rechte Seite der Grafik. Diese Variable wird gegen die X-Achsenvariable dargestellt, sodass ein anderer Plot auf der Grafik erstellt wird. So erstellen Sie eine statistische Zusammenfassung der eindeutigen Werte der horizontalen Variablen. Wählen Sie die Variable unter der Rolle **Vertikal** aus. Wählen Sie **Zusammenfassung der einzelnen horizontalen Werte** und anschließend die Zusammenfassungsfunktion. Wenn z. B. Ihre Eingabedaten die Umsätze und Ausgaben jedes Landes nach Regionen aufführen. Weisen Sie der Variablen 'Country' die Rolle **Horizontal** und der Variablen 'Expenses' die Rolle **Vertikal (Rechts)** zu. Wenn Sie die Summe der Ausgaben für jedes Land darstellen möchten, aktivieren Sie das Kontrollkästchen und wählen aus der Dropdown-Liste **Summe** aus. Die Anwendungsroutine fügt die regionalen Ausgaben für ein Land hinzu und stellt in der Endausgabe einen einzelnen Ausgabenwert für jedes Land dar. **Hinweis:** Die Rolle **Vertikal (Rechts)** ist nicht verfügbar, wenn Sie **Mehrere Liniendiagramme nach Gruppenspalte** als Liniendiagrammtyp auswählen.
Gruppe	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen, wie viele Plotlinien im gleichen Diagramm gezogen werden.

Optionsname	Beschreibung
	Für jeden eindeutigen Wert dieser Spalte wird ein separater Plot erstellt. Alle Linien verwenden dasselbe Achsenpaar. **Hinweis:** Die Rolle **Vertikal (Rechts)** ist nicht verfügbar, wenn Sie **Mehrere Liniendiagramme nach Gruppenspalte** als Liniendiagrammtyp auswählen.
Diagramme gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Diagramme, die erstellt werden. Für jeden eindeutigen Wert wird eine separate Grafik erzeugt. Aktivieren Sie das Kontrollkästchen **Gleiche Achsenskalierung für alle Diagramme**, wenn Sie möchten, dass für alle erzeugten Diagramme die gleiche Achsenskalierung verwendet wird. Standardmäßig basiert der Wertebereich für jede Achse auf den Minimal- und Maximalwerten der dargestellten Daten und könnte daher in den Diagrammen variieren. Wenn Sie diese Option auswählen haben alle Diagramme den gleichen Bereich für jede Achse.

Liniendiagramm: Festlegen von Darstellungsoptionen

Festlegen weiterer Plotoptionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Plots**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Optionen sind nur dann verfügbar, wenn Sie ein Liniendiagramm oder ein zweidimensionales Streuungsdiagramm auswählen.

So steuern Sie die Darstellung einer bestimmten Variablen im Plot:

1. Wählen Sie eine Variable aus der Liste aus. In der Liste werden die Variablen angezeigt, die Sie den Rollen **Vertikal** und **Vertikal (Rechts)** zugewiesen haben. Wenn Sie **Mehrere Liniendiagramme nach Gruppenspalte** als Typ für das Liniendiagramm ausgewählt haben, enthält diese Liste die eindeutigen Werte der Variablen, die Sie der Rolle **Gruppe** zugewiesen haben.
2. Geben Sie im Bereich **Linie** den Linienstil, die Linienbreite und die Linienarbe an.

3 Wählen Sie eine Umrissfarbe aus. Diese Option legt die Farben für folgende Elemente fest:

- den Umriss sämtlicher eingeschlossener Flächen
- Konfidenzgrenzen in einer Regressionsanalyse
- die Linien und Umrisse für Standardabweichungsplots

4 Geben Sie im Bereich **Datenpunktmarker** den Typ, das zu verwendende Symbol, die Höhe für jeden Marker und die Farbe der Datenpunktmarker an.

Hinweis: Manche Symbole werden nicht von allen Grafikausgabeformaten unterstützt. Wenn Sie ein nicht durch das Ausgabeformat unterstütztes Symbol auswählen, wird stattdessen ein anderes Symbol (ein unterstütztes Symbol) angezeigt.

5 Legen Sie fest, ob die Plotlinien bei fehlenden Werten unterbrochen werden soll.

Hinweis: Das Kontrollkästchen **Zweite Plotlinie bei fehlenden Werten unterbrechen** ist nur verfügbar, wenn Sie der Rolle **Vertikal (Rechts)** eine Variable zuweisen.

Weitere Darstellungsoptionen

Einige Darstellungsoptionen gelten für alle Grafik-Anwendungsroutinen. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Auswählen der Interpolationsmethode” auf Seite 754
- “Festlegen von Achsenoptionen” auf Seite 745
- “Hinzufügen von Referenzlinien” auf Seite 748
- “Festlegen von Legendenoptionen” auf Seite 749
- “Anpassen des Diagrammbereichs” auf Seite 750

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

43

Assistent für Liniendiagramme

Allgemeines über die Anwendungsroutine 'Liniendiagramm - Assistent' 317
Assistent für Liniendiagramme: Auswählen der Daten 317
Assistent für Liniendiagramme: Zuweisen von Variablen zu Rollen 318
Assistent für Liniendiagramme: Festlegen der Plotoptionen 319
Assistent für Liniendiagramme: Festlegen der Darstellungsoptionen 320
Assistent für Liniendiagramme: Festlegen von Titeln und Fußnoten 320

Allgemeines über die Anwendungsroutine 'Liniendiagramm - Assistent'

Der Assistent für **Liniendiagramme** unterstützt Sie bei der Erstellung zweidimensionaler Liniendiagramme mit Markern und Linieninterpolationen.

Assistent für Liniendiagramme: Auswählen der Daten

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im

Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Klicken Sie auf **Weiter**, um Variablen Rollen zuzuweisen.

Assistent für Liniendiagramme: Zuweisen von Variablen zu Rollen

1. Bevor Sie Variablen zu Rollen zuweisen können, müssen Sie die Daten auswählen.
2. So weisen Sie den folgenden Rollen Variablen zu:
   - **Horizontal** - Die Spalte, die Sie dieser Rolle zuweisen, ist die horizontale oder X-Achsenvariable für das Diagramm.
   - **Vertikal** - Die Spalte, die Sie dieser Rolle zuweisen, ist die vertikale oder Y-Achsenvariable für das Diagramm.
3. (Optional) Weisen Sie der Rolle **Gruppiert nach** eine Variable zu. Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen, wie viele Liniendiagramme erzeugt werden sollen. Für jeden eindeutigen Wert wird ein separates Diagramm erzeugt. Dieser Rolle können Sie mehrere Variablen zuweisen.
4. (Optional) Aktivieren Sie das Kontrollkästchen **Zusammenfassung der einzelnen horizontalen Werte**, um eine statistische Zusammenfassung der eindeutigen Werte der horizontalen Variable zu erstellen. Wählen Sie die Zusammenfassungsfunktion aus der Dropdown-Liste **Funktion** aus.

   Wenn z. B. Ihre Eingabedaten die Umsätze und Ausgaben jedes Landes nach Regionen aufführen. Weisen Sie der Variablen 'Country' die Rolle **Horizontal** und der Variablen 'Sales' die Rolle **Vertikal** zu. Wenn Sie die Summe der Umsätze für jedes Land darstellen möchten, aktivieren Sie das Kontrollkästchen und wählen aus der Dropdown-Liste **Summe** aus. Die Anwendungsroutine fügt die regionalen Umsätze für ein Land hinzu und stellt in der Endausgabe einen einzelnen Umsatzwert für jedes Land dar.
5. Klicken Sie auf **Weiter**, um die Diagrammoptionen festzulegen.

Assistent für Liniendiagramme: Festlegen der Plotoptionen

1. Bevor Sie diese Optionen festlegen können, müssen Sie den Rollen Variablen zuweisen.
2. Wählen Sie die Linieninterpolation aus der Dropdown-Liste **Typ** aus. Klicken Sie auf **Detail**, um Optionen für die ausgewählte Interpolation festzulegen. Folgende Interpolationen sind verfügbar:
   - **LaGrange** zeichnet eine glatte Kurve durch die Datenpunkte. Diese Methode wird in erster Linie verwendet, wenn die Daten aus exakten, durch Tabulatoren getrennten Werten bestehen. Ein Polynom des angegebenen Grads (1, 3 oder 5) wird durch die nächsten 2, 4 oder 6 Punkte angepasst. Im Allgemeinen ist die erste Ableitung nicht stetig. Wenn die Werte der X-Variablen nicht streng zunehmend sind, wird die entsprechende parametrische Methode verwendet.
   - **Linie** - Verbindet die Datenpunkte durch Geraden. Die Punkte werden in der Reihenfolge verbunden, in der sie in der Eingabedatei vorkommen.
   - **Nadel** zieht eine vertikale Linie von jedem Datenpunkt zu einer horizontalen Linie am 0-Wert auf der vertikalen Achse oder am Minimalwert auf der vertikalen Achse, wenn dieser größer als 0 ist. Die horizontale Linie wird automatisch gezeichnet.
   - **Regression** erstellt ein Regressionsanalyseplot. Diese Interpolationsmethode ist nur verfügbar, wenn Sie den Rollen Horizontal und Vertikal numerische Variablen zuweisen.
   - **Streuung** unterdrückt die Interpolation. Es wird eine Darstellung der Datenpunkte erstellt.
   - **Glättung** - passt eine glatte Linie mithilfe einer Splineroutine an die Daten an. Diese Methode dient zum Glätten verrauschter Daten. Die Punkte auf dem Plot liegen nicht unbedingt auf der Linie.
   - **Spline** - stellt die Linie mithilfe einer Splineroutine dar.
   - **STD** verbindet den mittleren Y-Wert mit ± 1, 2 oder 3 Standardabweichungen für jeden X-Wert mit einer durchgehenden Linie. Standardmäßig werden zwei Standardabweichungen verwendet.
   - **Stufe** - stellt die Daten mithilfe einer Stufenfunktion dar.
3. Geben Sie den Linienstil, die Linienbreite- und die Linienfarbe an.
4. Geben Sie das zu verwendende Symbol, die Höhe der einzelnenn Marker und die Farbe der Datenpunktmarker an.

Hinweis: Manche Symbole werden nicht von allen Grafikausgabeformaten unterstützt. Wenn Sie ein nicht durch das Ausgabeformat unterstütztes Symbol auswählen, wird stattdessen ein anderes Symbol (ein unterstütztes Symbol) angezeigt.

5 Klicken Sie auf **Weiter**, um die Darstellungsoptionen festzulegen.

Assistent für Liniendiagramme: Festlegen der Darstellungsoptionen

1 Vor dem Festlegen der Plotdarstellung müssen Sie die Plotoptionen festlegen.

2 Geben Sie an, ob der Plot Referenzlinien und Markierungen enthalten soll.

3 (Optional) Bearbeiten Sie die Etiketten für die vertikalen und horizontalen Achsen.

4 Klicken Sie auf **Weiter**, um die Titel und Fußnoten anzugeben.

Assistent für Liniendiagramme: Festlegen von Titeln und Fußnoten

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

1 Vor dem Angeben der Titel und Fußnoten müssen Sie die Darstellungsoptionen festlegen.

2 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld. Sie können in Titeln und Fußnoten Makrovariablen verwenden.

3 Klicken Sie auf **Fertig stellen**, um die Anwendungsroutine auszuführen.

44

Lineare Modelle

Allgemeines über die Anwendungsroutine 'Lineare Modelle' 321
Lineare Modelle: Zuweisen von Variablen zu Analyserollen 322
Lineare Modelle: Erstellen eines Modells 323
Lineare Modelle: Festlegen von Modelloptionen 324
Lineare Modelle: Festlegen erweiterter Optionen 325
Lineare Modelle: Auswählen von Post-Hoc-Tests 326
Informationen zu Post-Hoc-Tests 326
Festlegen von Optionen für Tests auf Mittelwerte für Kleinste-Quadrate-Mittelwerte 327
Festlegen von Optionen für Tests auf Mittelwerte für arithmetische Mittelwerte 328
Lineare Modelle: Erzeugen von Plots 330
Lineare Modelle: Festlegen von Vorhersageoptionen 331
Festlegen von Titeln und Fußnoten 332
Anzeigen von Eigenschaften 332

Allgemeines über die Anwendungsroutine 'Lineare Modelle'

Die Anwendungsroutine **Lineare Modelle** verwendet die Methode der kleinsten Quadrate, um allgemeine lineare Modelle anzupassen. Die Anwendungsroutine analysiert Daten im Rahmen allgemeiner linearer Modelle. Sie verwaltet Modelle, die mindestens eine kontinuierliche, abhängige Variable auf mindestens eine unabhängige Variable beziehen. Die unabhängigen Variablen können entweder Klassifizierungsvariablen sein, die die Beobachtungen in diskrete Gruppen unterteilen, oder es kann sich dabei um kontinuierliche Variablen handeln.

Mit dieser Anwendungsroutine können folgende Aufgaben erledigt werden:

- Durchführen einer Regressionsanalyse. Beispielsweise könnte ein Auto auf den Kraftstoffverbrauch bei unterschiedlichen Drehzahlen getestet werden, um festzustellen, bei welcher Drehzahl das Auto den höchsten Verbrauch hat. Es wird ein quadratisches Modell an die experimentellen Daten angepasst.
- Analysieren eines multivariaten wiederholten Messungsdesigns. Angenommen, zwei Reaktionen, Y1 und Y2, werden jeweils dreimal für jedes Subjekt gemessen (vor der Behandlung, nach der Behandlung und beim späteren Nachfassen). Jedes Subjekt erhält eine von drei Behandlungen: A, B oder die Kontrolle. In dieser Anwendungsroutine können Sie die verschiedenen Reaktionen und andere wiederholte Faktoren identifizieren, um die verschiedenen Messungslinien aufzuzeigen. Die wiederholte Messungsanalyse umfasst multivariate Tests auf die zeit- und behandlungsabhängigen Hauptauswirkungen sowie auf deren Interaktionen innerhalb der Reaktionen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GLM
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Lineare Modelle: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Abhängige Variablen	gibt die Variablen an, die als abhängige (Response) Variablen verwendet werden. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Quantitative Variablen	gibt die Variablen an, die als kontinuierliche, unabhängige (erklärende) Variablen (auch Effekte genannt) verwendet werden.
Klassifizierungsvariablen	gibt die Variablen an, die als diskrete unabhängige Effekte verwendet werden. Variablen, die Sie dieser Rolle zuweisen, können numerisch oder alphanumerisch sein, sollten aber eine begrenzte Anzahl an diskreten Werten aufweisen.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits als abhängige Variable oder als unabhängige Variable ausgewählt ist.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt eine Variable an, die als relative Gewichtung für die entsprechenden Werte der Variablen verwendet wird, die Sie als abhängige Variable ausgewählt haben. Sie können dieser Rolle maximal eine Variable zuweisen.

Lineare Modelle: Erstellen eines Modells

Sie können Modelle in mehreren SAS-Anwendungsroutinen erstellen. Weitere Informationen, siehe Anhang 3, “Erstellen eines Modells” auf Seite 765.

Lineare Modelle: Festlegen von Modelloptionen

Klicken Sie im Auswahlfenster auf **Modelloptionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Mit Konstante verknüpfte Tests anzeigen	zeigt die Ergebnisse aus Hypothesentests an, die an der Konstanten durchgeführt wurden.
Anzuzeigende Quadratsumme	zeigt die Quadratsummen an, die mit den schätzbaren Funktionen für jeden Effekt verknüpft sind. Schätzbare Funktionen sind lineare Kombinationen der Modellparameter. Quadratsummen werden für Hypothesentests berechnet. Typ I- und Typ III-Quadratsummen sind die Standardwerte. Sie können Quadratsummentypen durch Aktivieren des entsprechenden Kontrollkästchens hinzufügen oder entfernen. **Hinweis:** Es sollte mindestens ein Kontrollkästchen für **Anzuzeigende Quadratsumme** aktiviert sein.
Parameterschätzer anzeigen	zeigt die Parameterschätzer an. Ein Parameter ist eine Populationsgröße, die in der Regel fix, aber nicht bekannt ist, sodass sie mit Hilfe eines Modells aus Beispieldaten vorhergesagt wird. Die Parameterschätzer werden standardmäßig angezeigt. Sie können das Kontrollkästchen **Parameterschätzer anzeigen** nur dann deaktivieren, wenn der Rolle **Klassifizierungsvariablen** eine Variable zugewiesen wurde. Aktivieren Sie das Kontrollkästchen **Konfidenzgrenzen für Parameterschätzer**, um die Konfidenzintervalle für die Parameterschätzer anzuzeigen. Ein Konfidenzintervall ist eine Ober- und Untergrenze, die einen Bereich angemessener Werte für den Parameter angeben, der mit einem zuvor angegebenen Konfidenzniveau geschätzt wird. Sie können einen Wert für das Konfidenzniveau angeben oder einen Wert aus der Dropdown-Liste auswählen. Der Standardwert ist ein Konfidenzniveau von 95%.

Lineare Modelle: Festlegen erweiterter Optionen

Klicken Sie im Auswahlfenster auf **Weitere Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Schätzbare Funktionen für jeden Modelleffekt	■ Aktivieren Sie das Kontrollkästchen **Allg. Form aller schätzbaren Funktionen zeigen**, um die Matrixkoeffizienten der schätzbaren Funktionen von Typ I - Typ IV für jedes Modell anzuzeigen. ■ Aktivieren Sie jeweils das Kontrollkästchen neben dem Quadratsummentypen, um die entsprechenden Matrixkoeffizienten einer bestimmten schätzbaren Funktion anzuzeigen. ■ Aktivieren Sie das Kontrollkästchen **Schätzbare Funktionen als Aliasstruktur zeigen**, um die schätzbaren Funktionen als Aliasstruktur anzuzeigen. In einer Aliasstruktur stellt jede Zeile dar, welche lineare Parameterkombination durch jede schätzbare Funktion geschätzt wird. Der Tabelle mit den Parameterschätzungen wird eine Spalte hinzugefügt, die dieselben Informationen enthält. **Hinweis:** Sie können auch die schätzbaren Funktionen für die Konstante anzeigen, indem Sie das Kontrollkästchen **Mit Konstante verknüpfte Tests anzeigen** aktivieren.
Zu zeigende Matrizen	■ Aktivieren Sie das Kontrollkästchen **X'X**, um die erweiterte X'X-Kreuzproduktmatrix anzuzeigen. ■ Aktivieren Sie das Kontrollkästchen **X'X-Inverse**, um die erweiterte Inverse-X'X-Matrix anzuzeigen (verallgemeinerte Inverse).
Tuning-Parameter	**Hinweis:** Die Tuning-Parameter eignen sich zur Lösung schlecht konditionierter Probleme. ■ Um die Empfindlichkeit (Sensitivität) bei der Bildung von Typ III- oder IV-Funktionen zu ändern, geben Sie den neuen Wert in das Feld **Empfindlichkeit für Typ-III- und IV-Funktionen** ein. Alle Elemente in der schätzbaren Funktionsbasis, deren absoluter Wert kleiner als der

Optionsname	Beschreibung
	angegebene Wert ist, werden auf 0 gesetzt. Der Standardwert ist `1E-08`. ■ Um die Empfindlichkeit (Sensitivität) bei der Bestimmung einer Matrix zu ändern, geben Sie den gewünschten Wert in das Feld **Singularitätsprüfwert** ein. Diese Option regelt die Sensitivität der Regressionsroutine für lineare Abhängigkeiten im Design. Wenn ein diagonales Pivotelement kleiner ist als die korrigierte Quadratsumme für die Variable, multipliziert mit dem angegebenen Wert, dann wird die verknüpfte Spalte für die vorherigen Spalten als linear abhängig erklärt, der Verknüpfte Parameter wird auf 0 gesetzt. Der Standardwert ist `1E-07`. ■ Um die Toleranzen anzuzeigen, die verwendet werden, wenn die Empfindlichkeit der Regressionsroutine basierend auf dem Wert optimiert wird, der im Feld **Singularitätsprüfwert** angegeben wurde, aktivieren Sie das Kontrollkästchen **Toleranzen zeigen**.

Lineare Modelle: Auswählen von Post-Hoc-Tests

Informationen zu Post-Hoc-Tests

Klicken Sie im Auswahlfenster unter der Überschrift 'Post-Hoc-Tests' entweder auf **Kleinste Quadrate** oder **Arithmetisch**, um auf diese Optionen zuzugreifen.

1 Klicken Sie auf **Hinzufügen**, um dem Feld **Zu schätzende Effekte** einen Effekt hinzuzufügen.

2 Geben Sie für jeden Effekt in der Liste im Feld **Optionen für Tests auf Mittelwerte** Optionen an.

Hinweis: Im Feld **Optionen für Tests auf Mittelwerte** werden für Kleinste-Quadrate-Mittelwerte und arithmetische Mittelwerte unterschiedliche Optionen angezeigt.

Festlegen von Optionen für Tests auf Mittelwerte für Kleinste-Quadrate-Mittelwerte

Die folgenden Optionen sind im Feld **Optionen für Tests auf Mittelwerte** verfügbar.

Optionsname	Beschreibung
Zu verwendende Klasseneffekte	Wählen Sie die Klasseneffekte, die in der Analyse verwendet werden sollen. Standardmäßig werden keine Klasseneffekte verwendet.
Vergleiche	Geben Sie die Vergleichsoptionen an, die in der Analyse verwendet werden sollen. ■ **p-Werte für Differenzen anzeigen** berechnet die *p*-Werte für die Differenzen der Kleinste-Quadrate-Mittelwerte. Die folgenden Optionen sind verfügbar: □ **Ohne** berechnet keine *p*-Werte. □ **Standard** - berechnet die paarweisen Differenzen ohne Korrektur. □ **Alle paarweisfen Differenzen** - berechnet alle paarweisen Differenzen. □ **Erste Ausprägung als Kontrolle** - berechnet die Differenzen mit einer Kontrolle, die standardmäßig die erste Ausprägung aller angegebenen Kleinste-Quadrate-Mittelwert-Effekte ist. □ **Test, ob kleiner als Kontrolle** - Prüft, ob die nicht als Kontrolle verwendeten Ausprägungen kleiner sind als die Kontrolle. Eine nicht als Kontrolle verwendete Ausprägung gilt als signifikant kleiner, wenn die dazugehörige obere Konfidenzgrenze für die nicht als Kontrolle verwendete Ausprägung minus der Kontrolle kleiner als null ist und die dazugehörigen unteren Konfidenzgrenzen (die auf minus unendlich gesetzt werden) ignoriert werden. □ **Test, ob größer als Kontrolle** - Prüft, ob die nicht als Kontrolle verwendeten Ausprägungen größer sind als die Kontrolle. Eine nicht als Kontrolle verwendete Ausprägung gilt als signifikant größer, wenn die dazugehörige untere Konfidenzgrenze für die nicht als Kontrolle verwendete Ausprägung minus der Kontrolle größer als null ist und die dazugehörigen oberen Konfidenzgrenzen (die auf minus unendlich gesetzt werden) ignoriert werden.

Optionsname	Beschreibung
	■ **Korrekturmethode für Vergleich** fordert eine mehrfache Vergleichskorrektur der *p*-Werte und Konfidenzgrenzen für die Differenzen der Kleinste-Quadrate-Mittelwerte an.
Konfidenzgrenzen	Wählen Sie **Wahr** für die Option **Konfidenzgrenzen anzeigen**, um die Konfidenzgrenzen für einzelne Kleinste-Quadrat-Mittelwerte anzuzeigen. Wenn Sie für die Option **p-Werte für Differenzen anzeigen** einen Wert angeben, werden auch die Konfidenzgrenzen für die Differenzen zwischen Mittelwerten angezeigt.

Festlegen von Optionen für Tests auf Mittelwerte für arithmetische Mittelwerte

Diesse Optionen sind im Feld **Optionen für Tests auf Mittelwerte** verfügbar.

Optionsname	Beschreibung
Zu verwendende Klasseneffekte	Wählen Sie die Klasseneffekte, die in der Analyse verwendet werden sollen. Standardmäßig werden keine Klasseneffekte verwendet.
Vergleiche	Geben Sie die Vergleichsmethode an, die in der Analyse verwendet werden soll: ■ **Standard** - gibt an, dass keine Vergleichsmethode durchgeführt wird. ■ **t-Tests nach Bonferroni** - führt *t*-Tests nach Bonferroni von Differenzen zwischen Mittelwerten für alle Haupteffektmittelwerte durch. ■ **Multiple Spannweitentests nach Duncan** - führt multiple Spannweitentests nach Duncan für alle Haupteffektmittelwerte durch ■ **Multiple Vergleichstests nach Gabriel** - führt Gabriels multiple Vergleichsprozedur für alle Haupteffektmittelwerte durch ■ **Paarweise t-Tests** - führt paarweise *t*-Tests, entsprechend Fishers Test auf kleinste signifikante Differenzen für gleiche Zellengrößen, für alle Haupteffektmittelwerte durch ■ **Multipler Spannweitentest nach Ryan-Einot-Gabriel-Welsh** - führt den multiplen Spannweitentest nach Ryan-Einot-Gabriel-Welsh für alle Haupteffektmittelwerte durch

Optionsname	Beschreibung
	■ **Scheffes multiple Vergleichsprozedur** - führt Scheffes multiple Vergleichsprozedur für alle Haupteffektmittelwerte durch ■ **Sidaks korrigierter paarweiser t-Test** - führt paarweise *t*-Tests für Differenzen zwischen Mittelwerten durch, deren Ausprägungen gemäß Sidaks Ungleichheit für alle Haupteffektmittelwerte korrigiert wurden ■ **Multipler Spannweitentest nach Student-Newman-Keuls** - führt den multiplen Spannweitentest nach Student-Newman-Keuls für alle Haupteffektmittelwerte durch ■ **Tukeys studentisierter Spannweitentest** - führt Tukeys studentisierten Spannweitentest für alle Haupteffektmittelwerte durch ■ **K-Ratio-t-Test nach Waller-Duncan** - führt den *k*-Ratio-*t*-Test nach Waller-Duncan für alle Hautpeffektmittelwerte durch.
Mittleres Fehlerquadrat	Wählen Sie den Fehlereffekt, der in den multiplen Vergleichen verwendet werden soll. Sie können auch den Typ des mittleren Quadrats für den Fehlereffekt angeben.
Mittelwertsoptionen	Geben Sie folgende Optionen für die Mittelwerte an: ■ **Mittelwerte anzeigen für** - gibt an, ob die Mittelwerte für alle Modellvariablen oder nur für die abhängigen Variablen angezeigt werden. Standardmäßig zeigt die Anwendungsroutine die Mittelwerte für alle kontinuierlichen Variablen an. ■ **Nichtsignifikante Teilmengen joinen** - führt die mit der Vergleichsmethode berechneten Mittelwerte in absteigender Reihenfolge auf und gibt nichtsignifikante Teilmengen nach Liniensegmenten neben den entsprechenden Mittelwerten an. ■ **Mittelwerte absteigend sortieren** - sortiert die Mittelwerte in absteigender Reihenfolge. Wenn Sie **Wahr** für die Option **Mittelwerte absteigend sortieren** angeben und auch **Wahr** für eine der Optionen im Bereich **Konfidenzintervalle** wählen, werden die Mittelwerte nicht sortiert.
Konfidenzintervalle	Geben Sie an, ob Sie Konfidenzintervalle bei der Analyse einschließen möchten. Diese Optionen sind verfügbar: ■ **Für die Mittelwerte anzeigen** - stellt Ergebnisse der Vergleichsmethode als Intervalle für die Mittelwerte aller Ausprägungen der angegebenen Variablen dar. ■ **Für alle paarweisen Differenzen anzeigen** - stellt Ergebnisse der Vergleichsmethode als Konfidenzintervalle für alle paarweisen Differenzen zwischen Mittelwerten dar.

Optionsname	Beschreibung
Varianzhomogenität	Geben Sie an, ob ein Test auf Varianzhomogenität für die durch den Mittelwerteffekt definierten Gruppen durchgeführt werden soll. Diese Optionen sind verfügbar: ■ **Ohne** - gibt an, dass kein Test durchgeführt wird. Hierbei handelt es sich um die Standardeinstellung. ■ **Bartlett** - gibt den Test nach Bartlett, eine Modifikation des Normaltheorie-Likelihood-Ratio-Tests, an. ■ **Brown und Forsythe** - gibt Brown und Forsythes Variante des Tests nach Levene an. ■ **Levene (Quadratresiduen)** - verwendet Quadratresiduen im Test nach Levene. ■ **Levene (Absolutresiduen)** - verwendet Absolutresiduen im Test nach Levene. ■ **O'Brien** - gibt den Test nach O'Brien an, der im Wesentlichen eine Modifikation des Tests nach Levene mit Quadaratresiduen darstellt.

Lineare Modelle: Erzeugen von Plots

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Standardmäßig werden alle für die aktuelle Datenauswahl geeigneten Plots in der Ausgabe berücksichtigt. Sie können jedoch durch Auswahl der Option **Benutzerdefinierte Plot-Listen** festlegen, welche Plots Sie in die Ausgabe aufnehmen möchten. Sie können folgende Optionen wählen:

Optionsname	Beschreibung
Diagnose-Plots	erstellt ein Panel von Zusammenfassungsdiagnosen für die anzuzeigende Anpassung. Das Panel zeigt Streuungsdiagramme von Residuen, absoluten Residuen, studentisierten Residuen und beobachteter Abhängiger durch vorhergesagte Werte; studentisierter Residuen durch Leverage; Cook's *D* durch Beobachtung; einem Q-Q-Plot von Residuen; einem Residuum-Histogramm; und einem Residuum-Anpassungs-Verteilungs-Plot an.
Residuen-Plots	erstellt Streuungsdiagramme der Residuen gegen anzuzeigende kontinuierliche Kovariate.

Optionsname	Beschreibung
Box-Plot	erstellt einen Plot für den Modelleffekt in einer einfachen Analyse eines Varianzmodells. Um diesen Plot erstellen zu können, darf Ihr Modell nur eine Klassifizierungsvariable enthalten. Wenn Sie keine Klassifizierungsvariable zugewiesen haben, wird kein Box-Plot erstellt. Wenn Sie eine relative Gewichtungsvariable zugewiesen haben, werden diese Gewichte beim Erstellen des Box-Plots ignoriert.
Interaktions-Plot	erstellt einen Plot für eine zweifache Analyse des Varianzmodells. Um diesen Plot erstellen zu können, darf Ihr Modell nur zwei Klassifizierungsvariablen enthalten. Wenn Sie keine Klassifizierungsvariablen zugewiesen haben, wird kein Interaktions-Plot erstellt.

Lineare Modelle: Festlegen von Vorhersageoptionen

Klicken Sie im Auswahlfenster auf **Vorhersagen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Daten für Vorhersage	identifiziert die Datenquellen, die Sie zur Vorhersage der Ausgabedaten verwenden möchten. Wenn Sie das Kontrollkästchen **Weitere Daten** aktivieren sind diese Daten eingeschlossen. Um den Speicherort für Ihre zusätzlichen Daten zu bestimmen, klicken Sie auf **Durchsuchen**.
Ausgabedaten speichern	gibt an, ob die Vorhersagen und Diagnosestatistiken in einer Ausgabedatei gespeichert werden sollen: **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Optionsname	Beschreibung
Weitere Statistiken	gibt an, ob weitere Statistiken gespeichert werden sollen. ■ **Residuen** speichert das Likelihood-Residuum, das zum Identifizieren schlecht angepasster Beobachtungen verwendet wird, das standardisierte Pearson-Residuum (Chi), das zum Identifizieren von Beobachtungen dient, die im Modell nur unwesentlich berücksichtigt werden, und das standardisierte Abweichungsresiduum, das zum Identifizieren schlecht angepasster Beobachtungen dient. ■ **Vorhersagegrenzen** speichert die obere und untere Konfidenzgrenze und die Konfidenzgrenze für den vorhergesagten Wert.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

45

Lineare Regression

Allgemeines über die Anwendungsroutine 'Lineare Regression' 335
Lineare Regression: Zuweisen von Variablen zu Analyserollen 336
Lineare Regression: Festlegen von Modelloptionen 337
Lineare Regression: Festlegen von Statistikoptionen 340
Lineare Regression: Erzeugen von Plots 341
Lineare Regression: Festlegen von Vorhersageoptionen 343
Festlegen von Titeln und Fußnoten 344
Anzeigen von Eigenschaften 345

Allgemeines über die Anwendungsroutine 'Lineare Regression'

Bei der linearen Regressionsanalyse wird versucht, Ihren Daten mithilfe der Methode kleinster Quadrate eine lineare Funktion zuzuweisen. Unter Verwendung der Anwendungsroutine **Lineare Regression** können Sie lineare Regressionsanalysen für mehrere abhängige und unabhängige Variablen durchführen.

Beispielsweise könnten Sie mittels der Regressionsanalyse herausfinden, wie gut Sie das Gewicht eines Kindes vorhersagen können, wenn Sie seine Größe kennen. Angenommen, eine SAS-Datei enthält die Höhen- und Gewichtsmaße von 19 Kindern. Wenn Sie das Gewicht als abhängige Variable und die Größe als unabhängige Variable verwenden, können Sie anhand dieser Daten eine Regressionsanalyse durchführen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	REG
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Lineare Regression: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Abhängige Variablen	gibt die Variablen an, die als abhängige (Response) Variablen verwendet werden. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Erklärende Variablen	gibt die Variablen an, die als Regressorvariablen (erklärende Variablen) verwendet werden. Wenn Sie keine Regressorvariablen angeben, wird jede abhängige Variable an ein Nur-Mittelwert-Modell (Mean-Only) angepasst. Das bedeutet, dass jede abhängige Variable an ein Modell mit nur einem konstanten Term angepasst wird.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Sie können die Analyse nicht nach einer Variablen gruppieren, die Sie als abhängige Variable ausgewählt haben.

Optionsname	Beschreibung
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt die Variable mit den Werten an, die relative Gewichtungen für eine gewichtete Kleinste-Quadrate-Anpassung darstellen. Die Werte der relativen Gewichtungsvariablen müssen positiv sein. Sie können dieser Rolle maximal eine Variable zuweisen.

Lineare Regression: Festlegen von Modelloptionen

Klicken Sie im Auswahlfenster auf **Modell**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Modellauswahlmethode	Wählen Sie aus der Dropdown-Liste eine Modellauswahlmethode aus. Wenn Sie keine Methode auswählen, wird das volle Modell verwendet. Je nach Modellauswahlmethode sind weitere Optionen für diese Methode verfügbar. **Volles Modell (keine Auswahl)** Diese Methode ist der Standardwert und bietet keine Modellauswahlfunktion. Dies ist das Modell, das beim Zuweisen der abhängigen und quantitativen (erklärenden) Variablen erstellt wird. **Aufwärtsverfahren** Das Vorwärtsverfahren beginnt ohne Variablen im Modell. Diese Methode berechnet für jede erklärende Variable *F*-Statistiken, die den Beitrag der Variable wiedergeben, wenn diese ins Modell einbezogen wird. Die *p*-Werte für diese *F*-Statistiken werden mit dem Signifikanzniveau verglichen, das bei Einschließen einer Variablen in das Modell angegeben wird. Standardmäßig beträgt dieser Wert `0,5`. Um das Signifikanzniveau zu ändern, geben Sie den Wert in das Textfeld **Für Aufnahme ins Modell** ein. Wenn keine *F*-Statistik einen größeren Wert als diesen aufweist, hält das Vorwärtsverfahren an. Andernfalls fügt das Vorwärtsverfahren die Variable mit der größten *F*-Statistik zum Modell hinzu. Das Vorwärtsverfahren berechnet dann erneut *F*-Statistiken für die Variablen, die nicht im Modell verwendet werden und das Bewertungsverfahren wird wiederholt. Die Variablen werden also solange nacheinander zum Modell hinzugefügt, bis keine verbleibende

Optionsname	Beschreibung
	Variable mehr eine signifikante *F*-Statistik erzeugt. Wenn eine Variable dem Modell hinzugefügt wurde, verbleibt sie dort. **Rückwärtsverfahren** Das Rückwärtsverfahren berechnet zuerst *F*-Statistiken für ein Modell, einschließlich aller erklärenden Variablen. Die Variablen werden anschließend solange nacheinander aus dem Modell gelöscht, bis alle im Modell verbleibenden Variablen F-Statistiken erzeugen. Das Signifikanzniveau wird im Textfeld **Für Verbleiben im Modell** angegeben. Standardmäßig beträgt dieser Wert `0,10`. Bei jedem Schritt wird die am wenigsten zum Modell beitragende Variable gelöscht. **Schrittweise Auswahl** Die schrittweise Selektionsmethode beruht auf der Vorwärtsselektion. Bei der schrittweisen Auswahl verbleiben die bereits enthaltenen Variablen jedoch nicht notwendigerweise im Modell. Wie auch bei der Vorwärtsselektion werden Variablen nacheinander zum Modell hinzugefügt, zudem muss die *F*-Statistik signifikant sein, um eine Variable hinzuzufügen. Das Signifikanzniveau wird im Textfeld **Für Aufnahme ins Modell** angegeben. Nachdem eine Variable hinzugefügt wurde, werden alle bereits im Modell enthaltenen Variablen überprüft. Weist eine Variable keine signifikante *F*-Statistik auf, wird sie aus dem Modell gelöscht. Das Signifikanzniveau wird im Textfeld **Für Verbleiben im Modell** angegeben. Erst nach dieser Prüfung und nach Ausführung der notwendigen Löschvorgänge, kann eine andere Variable dem Modell hinzugefügt werden. Die schrittweise Auswahl ist beendet, wenn eine der folgenden Bedingungen erfüllt ist: ■ keine Variable außerhalb des Modells besitzt eine signifikante *F*-Statistik und jede Variable im Modell ist signifikant für das angegebene Signifikanzmodell für das Verbleiben im Modell ■ die Variable, die zum Modell hinzugefügt werden soll, ist die zuvor aus dem Modell gelöschte Variable
Modellauswahlmethode (Forts.)	**Maximale R-Quadrat-Verbesserung** Die Methode "Maximale R-Quadrat-Verbesserung" legt sich nicht auf ein einzelnes Modell fest. Es wird hingegen versucht, das beste Modell mit einer Variablen, mit zwei Variablen usw. zu finden. Es ist nicht gewährleistet, dass das jeweils beste Modell mit dem größten Wert größten Wert für R^2 ermittelt werden kann. Diese Methode beginnt mit der Suche nach dem Modell mit einer Variablen, welches das größte R^2 erzeugt. Danach wird eine weitere Variable, und zwar die Variable mit der größten Erhöhung von R^2 hinzugefügt. Wenn das Modell mit zwei Variablen vorliegt, wird jede im Modell enthaltene Variable mit jeder nicht im Modell enthaltenen Variablen verglichen. Diese Methode ermittelt bei jedem Vergleich, ob durch das Entfernen und Ersetzen einer Variablen durch eine andere Variable R^2 steigt. Nachdem alle Austauschmöglichkeiten verglichen wurden, tauscht diese Methode die Variable aus, die die größte Erhöhung von R^2 erzeugt. Die Vergleiche beginnen erneut. Der Vorgang wird solange fortgesetzt, bis diese Methode keine weiteren Variablen zum Austauschen findet, durch die R^2 steigt. Das resultierende Modell mit zwei Variablen gilt deshalb als das "beste" Modell mit zwei Variablen, das die Methode finden kann. Anschließend wird eine weitere Variable dem Modell hinzugefügt, und der Prozess des Vergleichens

Optionsname	Beschreibung
	und Austauschens wird wiederholt, um das "beste" Modell mit drei Variablen zu finden, usw.. Der Unterschied zwischen der schrittweisen Auswahl und der maximalen R^2-Auswahlmethode liegt darin, dass bei der maximalen R^2-Methode jeder Austauschvorgang vorher evaluiert wird. Bei der schrittweisen Auswahlmethode ist es möglich, die "schlechteste" Variable zu entfernen, ohne zu berücksichtigen, was das Hinzufügen der "besten" verbleibenden Variablen bewirken kann. **Minimale R-Quadrat-Verbesserung** Die Methode "Minimale R-Quadrat-Verbesserung" ähnelt der Methode "Maximale R-Quadrat-Verbesserung" mit der Ausnahme, dass Variablen so ausgetauscht werden, dass die geringste Erhöhung von R^2 erreicht wird. Bei einer bestimmten Anzahl von Variablen im Modell, erzeugen die Maximale R-Quadrat- und die Minimale R-Quadrat-Methode das gleiche "beste" Modell. Die Minimale R-Quadrat-Methode prüft jedoch mehr Modelle von jeder Größe. Bei der Methode "R-Quadrat-Auswahl" werden Teilmengen mit erklärenden Variablen gesucht, die eine abhängige Variable durch lineare Regression in der angegebenen Stichprobe am besten vorhersagt. Sie können die größte und kleinste Anzahl der in einer Teilmenge angezeigten erklärenden Variablen und die Anzahl der Teilmengen jeder auszuwählenden Größe angeben. Die R-Quadrat-Methode kann alle möglichen Regressionen mit Teilmengen effizient durchführen und die Modelle in absteigender Reihenfolge von R^2 in jeder Teilmengengröße anzeigen. Für den Vergleich von Teilmengen verschiedener Größen, stehen weitere Statistiken zur Verfügung. Diese Statistiken, sowie auch geschätzte Regressionskoeffizienten, können angezeigt oder als Ausgabe an eine SAS-Datei geschickt werden. Die Methode "R-Quadrat-Auswahl" unterscheidet sich von den anderen Auswahlmethoden darin, dass sie für jede Anzahl der berücksichtigten Variablen immer das Modell mit dem größten R^2 ermittelt. **Angepasste R-Quadrat-Auswahl** Die Methode "Angepasste R-Quadrat-Auswahl" ähnelt der Methode "R-Quadrat-Auswahl" mit der Ausnahme, dass die angepasste R^2-Statistik als Kriterium zum Auswählen von Modellen verwendet wird. Außerdem sucht die Methode das Modell mit dem höchsten angepassten R^2 innerhalb der Bereichsgrößen. **Mallows Cp-Auswahl** Die Methode "Mallows Cp-Auswahl" ähnelt der Methode "Angepasste R-Quadrat-Auswahl" mit der Ausnahme, dass Mallows Cp-Statistik als Kriterium für die Modellauswahl verwendet wird. Die Modelle werden in aufsteigender Reihenfolge von C_p aufgeführt.
Signifikanzniveaus	Beim Vorwärtsverfahren geben Sie ein Signifikanzniveau an, um eine erklärende Variable in das Modell aufzunehmen. Beim Rückwärtsverfahren geben Sie ein Signifikanzniveau an, damit eine quantitative Variable im Modell verbleibt. Bei der Methode "Schrittweise Auswahl" geben Sie beide Signifikanzniveaus an.
Modellanpassungsstatistiken	Bei Auswahl von R-Quadrat, Angepasstes R-Quadrat oder Mallows Cp verwenden Sie die Kontrollkästchen, um die Modellanpassungsstatistiken auszuwählen, die in die Ergebnisse einbezogen werden sollen.

Optionsname	Beschreibung
Für das Modell erzwungene Effekte	Bei allen Methoden, außer der Methode "Volles Modell", verwenden Sie die Kontrollkästchen, um Effekte in das Modell aufzunehmen. Ordnen Sie die Effekte mit den Pfeiltasten neu an. Um einen Effekt zu löschen, deaktivieren Sie das Kontrollkästchen.
Inkl. Konstante	Deaktivieren Sie bei allen Methoden das Kontrollkästchen, um die Effekte der Konstante aus der Analyse zu entfernen. Die Konstante wird standardmäßig eingeschlossen.

Lineare Regression: Festlegen von Statistikoptionen

Klicken Sie im Auswahlfenster auf **Statistiken**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Details zu Schätzern	legt die Schätzeroptionen fest. Folgende Optionen können ausgewählt werden. ■ standardisierte Regressionskoeffizienten. Ein standardisierter Regressionskoeffizient wird durch Dividieren eines Parameterschätzers durch das Verhältnis der Stichproben-Standardabweichung der abhängigen Variablen zur Stichproben-Standardabweichung des Regressors berechnet. ■ Die sequentielle Summe der Quadrate (Typ 1) und die partielle Summe der Quadrate (Typ 2) zusammen mit Parameterschätzern für jeden Term im Modell. ■ Die Korrelationsmatrix der Schätzer. ■ Die Kovarianzmatrix der Schätzer. ■ Konfidenzgrenzen für Parameterschätzer. Der Standardwert ist ein Konfidenzniveau von 95%.
Korrelationen	gibt an, ob quadratische partielle und quadratische semi-partielle Korrelationskoeffizienten eingeschlossen werden sollen. ■ Die partiellen Quadratkorrelationskoeffizienten werden als SS/(SS+SSE) berechnet, wobei SSE die Fehlerquadratsumme ist. ■ Die quadrierten semipartiellen Korrelationskoeffizienten werden als SS/SST berechnet, wobei SST die korrigierte Summe SS ist.

Optionsname	Beschreibung
	Wenn Sie die Konstante nicht im Modell berücksichtigen möchten, wird die unkorrigierte Summe SS im Nenner verwendet. Jeder Korrelationstyp verwendet zur Berechnung der Korrelationskoeffizienten sowohl Typ 1 als auch Typ 2 der Summe der Quadrate.
Diagnosen	gibt die durchzuführenden diagnostischen Analysen an. ■ Kollinearitätsanalyse. Diese Option fordert eine detaillierte Kollinearitätsanalyse der Regressoren an. Hierzu gehören Eigenwerte, Bedingungsindizes und die Dekomposition der Varianzen von Schätzern mit Bezug auf die einzelnen Eigenwerte. ■ Kollinearitätsanalyse ohne Konstante. Diese Option fordert dieselbe Analyse an wie die Option "Kollinearitätsanalyse", nur dass die Konstante in der Analyse nicht berücksichtigt wird. ■ Toleranzwerte für Schätzer. Toleranz einer Variablen wird definiert als 1-R2, wobei R2 aus der Regression der Variablen auf alle anderen Regressoren im Modell ermittelt wird. Varianzinflationswerte. Die Varianzinflation ist der Kehrwert der Toleranz. ■ Test auf Ungleichheit der Varianzen. Diese Option prüft, ob die ersten und zweiten Momente des Modells richtig angegeben wurden. ■ Asymptotische Kovarianzmatrix. Diese Option zeigt die geschätzte asymptotische Kovarianzmatrix der Schätzer unter der Hypothese der Ungleichheit der Varianzen an. ■ Durbin-Watson-Statistik. Die Durbin-Watson-Statistik zeigt, ob die Fehler eine Autokorrelation erster Ordnung aufweisen. (Dieser Test eignet sich nur für Zeitreihendaten.) Die Stichproben-Autokorrelation der Residuen wird ebenfalls erstellt.

Lineare Regression: Erzeugen von Plots

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Standardmäßig werden alle für die aktuelle Datenauswahl geeigneten Plots in der Ausgabe berücksichtigt. Sie können jedoch durch Auswahl der Option **Benutzerdefinierte Plot-Listen** festlegen, welche Plots Sie in die Ausgabe aufnehmen möchten. Folgende Optionen können ausgewählt werden:

Plot-Name	Beschreibung
Histogramm-Plot der Residuen	erstellt ein Histogramm der Anpassungsresiduen.
Residuen-gegen-Vorhergesagte-Plot	erstellt einen Plot der Residuen gegen vorhergesagte Werte. Ein Residuum ist die Differenz zwischen dem beobachteten und dem vorhergesagten Wert.
Studentisierte-Residuen-gegen-Vorhergesagte	erstellt einen Plot der studentisiertes Residuen gegen vorhergesagte Werte.
Beobachtet-gegen-Vorhergesagt-Plot	erstellt einen Plot der beobachteten gegen die vorhergesagten Werte.
Cook's D-Statistik-Plot	erstellt einen Plot der Cook's D-Statistik gegen die Beobachtungszahl.
Studentisierte-Residuen-gegen-Leverage-Plot	erstellt einen Plot der studentisierten Residuen gegen den Leverage. Ein studentisiertes Residuum ist der Wert des Residuums dividiert durch den Standardfehler bei gelöschter aktueller Beobachtung. Der Leverage-Wert für eine Beobachtung ist der Wert des entsprechenden diagonalen Elements in der *H*-Matrix.
Normalquantil-Plot der Residuen	erstellt einen Q-Q-Plot.
Residualanpassungs-Plot	erstellt vergleichende Plots der Quantilen der zentrierten Anpassung und der Residuen.
Box-Plot der Residuen	erstellt einen Box-Plot der Residuen.
Diagnose-Plots	erstellt Plots der Anpassungsdiagnose.
DFFITS-Plots	erstellt Plots von DFFITS (Differenz der Anpassung, standardisiert). Die DFFITS-Statistik stellt den Standardeinfluss der Beobachtung auf den vorhergesagten Wert dar. Ein großer Wert weist auf einen großen Einfluss der Beobachtung hin.

Plot-Name	Beschreibung
DFBETAS-Plots	stellt den standardisierten Unterschied im Parameterschätzer aufgrund der Löschung der Beobachtung dar. Es ist hilfreich, den Effekt einer einzelnen Beobachtung für jeden einzelnen Parameterschätzer des angepassten Modells zu bestimmen.
Residuen-Plots	erstellt Plots der Residuen gegen die Regressoren.
Streuungsdiagramm mit Regressionslinie	erstellt einen Plot der Regressionslinie, überlagert auf einem Streuungsdiagramm der Daten.

Lineare Regression: Festlegen von Vorhersageoptionen

Klicken Sie im Auswahlfenster auf **Vorhersagen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Daten für Vorhersage	identifiziert die Datenquellen, die Sie zur Vorhersage der Ausgabedaten verwenden möchten. Wenn Sie das Kontrollkästchen **Weitere Daten** aktivieren, werden diese Daten eingeschlossen.
Weitere Statistiken	gibt die zusätzlichen Statistiken an, die eingeschlossen werden sollen. ■ Die Auswahl von **Residuen** schließt die gewöhnlichen, standardisierten und studentisierten Residuen in die Ausgabe ein. Ein gewöhnliches Residuum ist die Differenz zwischen dem beobachteten und dem vorhergesagten Wert. Ein standardisiertes Residuum ist der Wert des Residuums dividiert durch den Standardfehler. Ein studentisiertes Residuum ist der Wert des Residuums dividiert durch den Standardfehler bei gelöschter aktueller Beobachtung. ■ Die Auswahl von **Vorhersagegrenzen** schließt die obere und untere Grenze eines 100*(1-alpha)%-Konfidenzintervalls für jede Vorhersage und den erwarteten Mittelwert der abhängigen Variable in die Ausgabe ein.

Optionsname	Beschreibung
Ausgabedaten speichern	legt fest, ob Prognosen und Diagnosestatistiken dauerhaft in einer Ausgabedatei gespeichert werden sollnen. Folgende Diagnosestatistiken sind verfügbar: ■ Cooks *D*-Einfluss-Statistik ■ der Standardeinfluss der Beobachtung auf den vorhergesagten Wert (die DFFITS-Statistik) ■ der Leverage, ■ der Standardfehler des einzelnen vorhergesagten Wertes ■ der Standardfehler des mittleren vorhergesagten Wertes ■ der Standardfehler des Residuums **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Anzeigen von Ausgabe und Plots	zeigt die Ergebnisse an. Um die Ergebnisse zu unterdrücken, deaktivieren Sie das Kontrollkästchen. Wenn Sie dieses Kontrollkästchen deaktivieren, sind die Optionen im Fenster **Plots** nicht verfügbar. Die Ergebnisse werden standardmäßig angezeigt. Um die Vorhersagen in die Ergebnisse einzuschließen, aktivieren Sie das Kontrollkästchen **Vorhersagen anzeigen**. Diese Option ist nur verfügbar, wenn im Bereich **Daten für Vorhersage** eine Datenquelle ausgewählt und das Kontrollkästchen **Ausgabe und Plots anzeigen** aktiviert ist.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

46

Listenbericht

Allgemeines über die Anwendungsroutine 'Listenbericht' **347**
Listenbericht: Zuweisen von Variablen zu Analyserollen **348**
Listenbericht: Festlegen von Listenoptionen **349**
Festlegen von Titeln und Fußnoten **351**
Anzeigen von Eigenschaften **351**

Allgemeines über die Anwendungsroutine 'Listenbericht'

Die Anwendungsroutine **Listenbericht** gibt die Beobachtungen in einer SAS-Datei aus und verwendet dabei alle oder einige der Variablen. Sie können eine Vielzahl von Berichten erstellen, von einfachen Listenberichten bis zu sehr angepassten Berichten, welche die Daten gruppieren und für numerische Variablen Summen und Zwischensummen berechnen.

Beispielsweise können Sie die Anwendungsroutine **Listenbericht** zum Erstellen eines Berichts verwenden, der die Ausgaben und Einnahmen der einzelnen Verkaufsregionen zusammenfasst, diese Werte mit den Gesamtausgaben und dem Gesamtumsatz des Unternehmens vergleicht, die Anzahl der Beobachtungen in den einzelnen Untergruppen und im gesamten Bericht anzeigt und über einen angepassten Titel verfügt, der den Namen der Region enthält.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	PRINT
Erforderliche SAS-Produkte	Base SAS

Name der Anforderung	Prozedur- und Produktnamen
Empfohlene zusätzliche SAS-Produkte	Ohne

Listenbericht: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Variablen auflisten	druckt die angegebenen Variablen in der Reihenfolge, in der sie angegeben werden. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Analyse gruppieren nach	gibt an, wie die Tabelle sortiert wird. Es wird für jeden Einzelwert (oder jede BY-Gruppe) in der Variablen oder Kombination aus Variablen eine Liste erstellt.
Seitengruppierung nach	druckt bei jeder Wertänderung der angegebenen Variablen, oder wenn die nächste BY-Gruppe beginnt, eine neue Seite. Die Variable, die Sie dieser Rolle zuweisen, muss auch eine Variable in der Rolle **Analyse gruppieren nach** sein. Sie können dieser Rolle maximal eine Variable zuweisen.
Summe von	druckt die Summen der ausgewählten Variablen am unteren Rand des Listenberichts.
Zwischensumme von	druckt bei jeder Wertänderung der angegebenen Variablen, oder wenn die nächste BY-Gruppe beginnt, eine Zwischensumme. Die Variable, die Sie dieser Rolle zuweisen, muss auch eine Variable in der Rolle **Analyse**

Optionsname	Beschreibung
	gruppieren nach sein. Sie können dieser Rolle maximal eine Variable zuweisen.
ID-Etikett	gibt an, wie die Zeilen identifiziert werden sollen, anstatt Beobachtungsnummern zu verwenden.

Listenbericht: Festlegen von Listenoptionen

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Aufzulistende Zeilen	gibt die Zeilen an, die aufgelistet werden sollen. Folgende Optionen können ausgewählt werden: ■ **Alle Zeilen** druckt alle Zeilen. ■ **Erste n Zeilen** druckt die Anzahl der Zeilen, die Sie im Textfeld **Menge (n)** angegeben haben. Wenn Sie beispielsweise 50 für *n* angeben, enthalten die Ausgabedaten die Zeilen 1 bis 50 aus der Eingabedatei. ■ **Erste n Prozent der Zeilen**: Druckt den Prozentwert der Zeilen, die Sie im Textfeld **Menge (n)** angegeben haben. Wenn Ihre Eingabedatei beispielsweise 100 Zeilen enthält und Sie *n* auf 25% setzen, enthalten die Ausgabedaten die Zeilen 1 bis 25 aus der Eingabedatei. ■ **Jede n. Zeile** druckt jede *n*. Zeile in der Datei. Sie geben den Wert *n* im Textfeld **Menge (n)** an. Wenn Sie beispielsweise *n* auf 10 setzen, enthalten die Ausgabedaten jede 10. Zeile der Eingabedatei.
Zeilennummer drucken	schließt eine Spalte in die Ausgabe ein, die die Zeilennummer für jede Beobachtung auflistet. Sie können für diese Spalte im Textfeld **Spaltenüberschrift** ein Etikett angeben. Der Standardname dieser Spalte ist `Zeilennummer`.
Variablenetiketten als Spaltenüberschriften verwenden	verwendet als Spaltenüberschrift das Variablenetikett anstelle des Variablennamens.

Optionsname	Beschreibung
Anzahl der Zeilen drucken	listet die Anzahl der Zeilen in der Tabelle am Ende der Ausgabe oder die Anzahl der Zeilen in jeder BY-Gruppe am Ende der Ausgabe jeder BY-Gruppe auf.
Werte vor dem Summieren runden	rundet jeden numerischen Wert auf die Anzahl der Dezimalstellen in seinem Format oder auf zwei Dezimalstellen, wenn kein Format angegeben wurde. Wenn diese Option angegeben ist, führt die Anwendungsroutine Listenbericht die Rundung vor dem Summieren der Variablen durch.
Seite in Abschnitte unterteilen	unterteilt jede Seite in Abschnitte, wenn es mehr Spalten gibt als auf die Seite passen. Wenn Sie die Option **Seite in Abschnitte unterteilen** deaktivieren, platziert die Anwendungsroutine so viele Beobachtungen wie möglich auf jede Seite. Wenn es mehr Spalten gibt als auf eine Seite passen, werden die nicht mehr passenden Spalten auf den nächsten Seiten platziert. Je nach Anzahl der Zeilen und Spalten in der Tabelle werden durch diese Option weniger Druckseiten erzeugt. **Hinweis:** Diese Option betrifft nur Textausgabe. Diese Option wirkt sich nicht auf Ergebnisse im SAS Report-, HTML-, RTF- oder PDF-Format aus.
Überschriftenrichtung	druckt Spaltenüberschriften horizontal oder vertikal. Wählen Sie **Standard** und lassen Sie SAS die optimale Anordnung für jede Spalte bestimmen.
Spaltenbreite	legt fest, wie die Anwendungsroutine Aufgabe **Listenbericht** die Spaltenbreiten bestimmt: ■ **Standard**: Die Anwendungsroutine bestimmt die Spaltenbreite auf einer Pro-Seite-Basis. ■ **Voll**: Die Anwendungsroutine verwendet Formatbreite (oder Standardbreite, wenn kein Format angegeben wurde) für alle Seiten. ■ **Minimum**: Die Anwendungsroutine verwendet die kleinstmögliche Spaltenbreite auf einer Pro-Seite-Basis. ■ **Einheitlich**: Liest die gesamte Tabelle, um die geeignete Spaltenbreite vor dem Erzeugen von Ausgabe zu bestimmen. Wenn diese Option nicht ausgewählt ist, können unterschiedliche Seiten unterschiedliche Breiten für die gleiche Spalte aufweisen. ■ **Einheitlich nach**: Die Anwendungsroutine druckt alle Seiten einheitlich innerhalb jeder BY-Gruppe.

Optionsname	Beschreibung
	Hinweis: Diese Option betrifft nur Textausgabe. Diese Option wirkt sich nicht auf Ergebnisse im SAS Report-, HTML-, RTF- oder PDF-Format aus.

Etiketten trennen bei	wenn die Variablenetiketten eines der Trennzeichen (*, !, @, #, $, %, ^, &, or +) enthalten, werden die Etiketten beim Trennzeichen umgebrochen. Wenn ein Variablenetikett z.B. "Dies ist*ein Etikett" lautet und das Zeichen * als Trennzeichen ausgewählt ist, dann lautet die Spaltenüberschrift:
	`Dies ist`
	`ein Etikett`
	Es ist nicht nötig, die Optionen **Variablenetikett als Spaltenüberschrift verwenden** und **Etiketten trennen bei** gemeinsam auszuwählen. **Etiketten trennen bei** impliziert, dass Sie Variablenetiketten verwenden möchten.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

47

Listenberichtassistent

Allgemeines über den Listenberichtassistenten **353**
Listenberichtassistent: Auswählen von Daten **354**
Listenberichtassistent: Definieren der Liste **354**
Listenberichtassistent: Anzeigen von Summen **355**
Listenberichtassistent: Festlegen von Titeln und Fußnoten **356**

Allgemeines über den Listenberichtassistenten

Mit dem **Listenberichtassistenten** können Sie Detail- oder Zusammenfassungsberichte erstellen. Ein Detailbericht enthält eine Zeile für jede Beobachtung, die für den Bericht ausgewählt wurde. Ein Zusammenfassungsbericht konsolidiert Daten, sodass jede Zeile mehrere Beobachtungen enthält. Sowohl Detail- als auch Zusammenfassungsberichte können Summen enthalten, die numerische Daten für eine Reihe von Zeilen oder alle Zeilen zusammenfassen. Sie können auch separate Tabellen oder Seitenumbrüche basierend auf den Werten einer Variablen erstellen.

So können Sie beispielsweise mit dem **Listenberichtassistenten** einen Bericht erstellen, in dem die Umsätze für Mineralwasser und Kaffee in Lebensmittelgeschäften jeder Region aufgeführt werden. Die Manager in den einzelnen Regionen werden in einer separaten Tabelle gruppiert. In jeder Tabelle werden die einzelnen Produktumsätze und die Gesamtumsätze für jeden Manager erfasst.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	REPORT
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	Ohne

Listenberichtassistent: Auswählen von Daten

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Klicken Sie auf **Weiter**, um die Liste zu definieren.

Listenberichtassistent: Definieren der Liste

Wenn die Datenquelle weniger als 20 Spalten enthält, berücksichtigt der **Listenberichtsassistent** alle Spalten in der Datenquelle. Diese Spalten werden im **Vorschaufenster** angezeigt. Wenn Ihre Daten mehr als 20 Spalten enthalten, werden im Listenbericht keine Spalten angezeigt und das **Vorschaufenster** ist leer.

Sie müssen die Spalten angeben, die dem Bericht hinzugefügt werden sollen. Sie können alle Spalten in der Datenquelle, eine bestimmte Spalte oder die Häufigkeit hinzufügen.

Klicken Sie auf **Weiter**, um die Summen anzuzeigen.

Listenberichtassistent: Anzeigen von Summen

Sie können die Zwischensummen oder die Gesamtsummen für alle numerischen Variablen anzeigen, bei denen es sich weder um Klassifizierungs- noch um übergreifende Variable handelt.

So geben Sie Summen an:

1. Bevor Sie die Summen angeben können, müssen Sie die Liste definieren.
2. Aktivieren Sie im Fenster **Summen auswählen** die Kontrollkästchen für die Variablen, für die Sie die Summe errechnen möchten. Um die Summen für alle Variablen anzuzeigen, aktivieren Sie das Kontrollkästchen **Alle auswählen**.
3. Zur Angabe des Summentyps, der Position des Gesamtwertes und des Etiketts für die Zeile, die die Summe enthält, klicken Sie auf **Bearbeiten**. Das Dialogfenster **Summen bearbeiten** wird angezeigt.

   Folgende Optionen können ausgewählt werden:

   - **Summentyp** - Gibt den Summentyp an, der in der Tabelle angezeigt wird. Sie können die Gesammtsummen und die Zwischensummen anzeigen lassen. Sie müssen mindestens einen anzuzeigenden Summentyp auswählen. Wenn Sie alle Summen in der Tabelle anzeigen möchten, aktivieren Sie das Kontrollkästchen **Alle auswählen**.
   - **Position** - Bestimmt die Position, an der die Summenzeile in der Tabelle angezeigt wird. Sie können die Summe vor oder nach den Zeilen bzw. Beobachtungen anzeigen lassen, die für die Berechnung der Summe verwendet wurden.
   - **Etikett Summen** - Gibt das Etikett an, das für die Summenzeilen verwendet wird. Das Standardetikett ist Summe.
   - **Etikett Zw.-Summen** - Gibt das Etikett an, das für die Zwischensummenzeilen verwendet wird. Das Standardetikett ist Zwischensumme. Sie können angeben, ob das Etikett vor oder nach dem Text in dieser Zeile angezeigt wird. Standardmäßig wird kein Leerzeichen zwischen dem Etikett und eventuell vorangehendem oder folgendem Text eingefügt. Um ein Leerzeichen einzufügen, müssen Sie es vor oder nach dem Etikett im Feld **Etikett Zw.-Summen** eingeben.

4. Klicken Sie auf **Weiter**, um Titel und Fußnoten anzugeben.

Listenberichtassistent: Festlegen von Titeln und Fußnoten

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

1. Bevor Sie die Titel und Fußnoten zuweisen, müssen Sie die Summen für alle numerischen Variablen anzeigen.
2. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld. Sie können in Titeln und Fußnoten Makrovariablen verwenden.
3. Klicken Sie auf **Fertig stellen**, um die Anwendungsroutine auszuführen.

48

Logistische Regression

Allgemeines über die Anwendungsroutine 'Logistische Regression' ... **357**

Logistische Regression: Zuweisen von Variablen zu Analyserollen ... **358**

Logistische Regression: Erstellen eines Modells ... **360**
- Angeben der Ausprägung ... 360
- Erstellen eines Modells ... 361
- Angeben einer Modellauswahlmethode ... 362

Logistische Regression: Festlegen von Modelloptionen ... **364**

Logistische Regression: Erzeugen von Plots ... **366**

Logistische Regression: Festlegen von Vorhersageoptionen ... **368**

Festlegen von Titeln und Fußnoten ... **369**

Anzeigen von Eigenschaften ... **369**

Allgemeines über die Anwendungsroutine 'Logistische Regression'

Binäre abhängige Werte, wie Erfolg und Versagen, und ordinale abhängige Werte, wie klein, mittel und groß, kommen in vielen Forschungsgebieten vor. Die logistische Regressionsanalyse untersucht den Zusammenhang zwischen diesen diskreten abhängigen Werten und einer Reihe von erklärenden Variablen.

Eventuell möchten Sie mittels der logistischen Regression die Auswirkungen einer klinischen Studie bestimmen. Angenommen, Sie verwenden die Anwendungsroutine Logistische Regression, um die Ergebnisse einer Studie zu den analgetischen Auswirkungen der Behandlung älterer Patienten mit Neuralgie zu analysieren. Es werden zwei Testbehandlungen und eine Placebo-Behandlung verglichen. Die abhängige Variable ist, ob der Patient über Schmerzen klagt oder nicht. Die Forscher notieren Alter und Geschlecht der Patienten sowie die Dauer der

Beschwerden vor Beginn der Behandlung. Anhand der Ausgabe der Anwendungsroutine **Logistische Regression** können Sie bestimmen, ob die schmerzlindernde Wirkung dieser beiden Testbehandlungen unterschiedlich ist.

Hinweis: Beim Öffnen der Anwendungsroutine **Logistische Regression** kann es ggf. zu einer kleinen zeitlichen Verzögerung kommen, da SAS Enterprise Guide im Hintergrund ermittelt, welche Variablen binäre Werte aufweisen. Nach Abschluss der Analyse bietet die Anwendungsroutine Informationen über die Variable, die Sie der Rolle **Abhängige Variable** zuweisen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	LOGISTIC
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Logistische Regression: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Abhängige Variable	gibt die Variable an, die als abhängige (Response) Variable verwendet wird. Sie können dieser Rolle nur eine Variable zuweisen.
Quantitative Variablen	gibt die Variablen an, die als unabhängige (erklärende) Variablen verwendet werden. Sie können auch die Änderungseinheiten für die quantitativen Variablen angeben. Um beispielsweise geschätzte benutzerdefinierte Odds Ratios in den Ergebnissen zu generieren, geben Sie `-2` in das Feld **Einheiten** ein, so dass ein

Optionsname	Beschreibung
	Odds Ratio die Änderung in den Odds darstellt, wenn die Variable um zwei Einheiten verringert wird.
Klassifizierungsvariablen	gibt die Klassifizierungsvariablen an, die in der Analyse verwendet werden sollen. Hierbei kann es sich um alphanumerische oder numerische Variablen handeln. Bei numerischen Variablen müssen Einzelwerte statt kontinuierlicher Werte vorliegen. Um die Parametrisierungsmethode für jede Variable auszuwählen, wählen Sie die Klassifizierungsvariable aus. Wählen Sie im Bereich **Codierstil für *Variablenname*** aus, ob als Parametrisierungsmethode für die Variable die Effektcodierung (**Effekte**) oder Referenzzellencodierung (**Verweis**) verwendet werden soll.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Sie können die Analyse nicht nach einer Variablen gruppieren, die Sie als abhängige Variable ausgewählt haben.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt eine Variable an, die als relative Gewichtung für die entsprechenden Werte der Variablen verwendet wird, die Sie als abhängige Variable ausgewählt haben. Negative oder fehlende Werte werden aus der Analyse ausgeschlossen. Um die Gewichtung der Variablen zu normalisieren, aktivieren Sie das Kontrollkästchen **Normalisieren**.

Logistische Regression: Erstellen eines Modells

Angeben der Ausprägung

Klicken Sie im Auswahlfenster unter der Überschrift **Modell** auf **Ausprägung**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Typ abhängiger Ausprägungen	Welche Typen abhängiger Ausprägungen verfügbar sind, hängt von der Anzahl der Ausprägungen in der abhängigen Variable ab. ■ Wenn Ihre abhängige Variable nur zwei Ausprägungen hat, ist der Typ abhängiger Ausprägungen binär. ■ Wenn Ihre abhängige Variable mehr als zwei Ausprägungen hat, können Sie als Response-Typ Geordnet oder Ungeordnet auswählen.
Modelltyp	Sie können aus folgenden Modelltypen auswählen: ■ **Logit** zeigt die Ausgabe aus der Log-Odds-Funktion (Standard) an. Das binäre Logit-Modell wird verwendet, wenn zwei abhängige Kategorien vorhanden sind. Das kumulative Logit-Modell wird verwendet, wenn mehr als zwei abhängige Kategorien vorhanden sind. ■ **Probit** zeigt die Ausgabe aus der inversen Integralfunktion für die Standard-Normalwahrscheinlichkeit an. Das binäre Probit-Modell wird verwendet, wenn zwei abhängige Kategorien vorhanden sind. Das kumulative Probit-Modell wird verwendet, wenn mehr als zwei abhängige Kategorien vorhanden sind. ■ **Komplementäres Log-Log** zeigt die Ausgabe aus der komplementären Log-Log-Funktion an. Das binäre komplementäre Log-Log-Modell wird verwendet, wenn zwei abhängige Kategorien vorhanden sind. Das kumulatve komplementäre Log-Log-Modell wird verwendet, wenn mehr als zwei abhängige Kategorien vorhanden sind. ■ **Log-Log** zeigt die Ausgabe aus der Log-Log-Funktion an. Das binäre Log-Log-Modell wird verwendet, wenn zwei abhängige Kategorien vorhanden sind. Das kumulatve Log-

Optionsname	Beschreibung
	Log-Modell wird verwendet, wenn mehr als zwei abhängige Kategorien vorhanden sind. ■ **glogit** zeigt die Ausgabe der generalisierten Logit-Funktion an. Im generalisierten Logit-Modell wird jeder Kategorie ohne Verweis eine Verweiskategorie gegenübergestellt. **Hinweis:** Die Option ist nur dann verfügbar, wenn Sie **Ungeordnet** als Typ der abhängigen Ausprägung angegeben haben.
Abhängige Ausprägungen	Im Feld "Abhängige Ausprägungen" werden in aufsteigender Reihenfolge die formatierten Ebenen für die abhängige Variable aufgeführt.
Modell an Niveau anpassen	Sie können angeben, wie das Modell an ein Niveau angepasst wird. ■ Wenn Sie als Typ abhängiger Ausprägungen **Binär** angegeben haben, wählen Sie in der Dropdown-Liste Modell an Niveau anpassen die Ereigniskategorie für das binäre Response-Modell. ■ Wenn Sie als Typ abhängiger Ausprägung **Geordnet** angegeben haben, können Sie mit den Werten in der Dropdown-Liste **Modell an Niveau anpassen** die Reihenfolge der Response-Kategorien festlegen. Wenn Sie den ersten geordneten Wert auswählen, werden die Response-Kategorien in der Reihenfolge sortiert, in der sie im Feld **Abhängige Ausprägungen** aufgeführt sind. Wenn Sie den letzten geordneten Wert auswählen, werden die Response-Kategorien in der umgekehrten Reihenfolge sortiert.
Verweisebene	Die Option ist nur dann verfügbar, wenn Sie **Ungeordnet** als Typ der abhängigen Ausprägung angegeben haben. Im generalisierten Logit-Modell können Sie die Verweiskategorie in der Dropdown-Liste **Verweisebene** angeben. Im generalisierten Logit-Modell stellt jedes Logbit einer Kategorie ohne Verweis eine Verweiskategorie gegenüber.

Erstellen eines Modells

Sie können Modelle in mehreren SAS-Anwendungsroutinen erstellen. Weitere Informationen, siehe Anhang 3, "Erstellen eines Modells" auf Seite 765.

Angeben einer Modellauswahlmethode

Klicken Sie im Auswahlfenster unter der Überschrift 'Modell' auf **Auswählen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Modellauswahlmethode	Wählen Sie aus der Dropdown-Liste eine Modellauswahlmethode aus. Wenn Sie keine Methode auswählen, wird das volle Modell verwendet. Je nach Modellauswahlmethode sind weitere Optionen für diese Methode verfügbar. Folgende Methoden stehen zur Verfügung: ■ **Volles Modell (keine Auswahl)** ist der Standard und ermöglicht keine Auswahl von Effekten. Dies ist das Modell, das beim Zuweisen der abhängigen und quantitativen (erklärenden) Variablen erstellt wird. ■ **Vorwärtsverfahren** startet ohne Effekte im Modell. Diese Methode berechnet für jede erklärende Variable *F*-Statistiken, die den Beitrag der Variable wiedergeben, wenn diese ins Modell einbezogen wird. Die *p*-Werte für diese *F*-Statistiken werden mit dem Signifikanzniveau verglichen, das für den Einschluss einer Variable in das Modell angegeben wird. Standardmäßig beträgt dieser Wert `0,05`. Um das Signifikanzniveau zu ändern, geben Sie den Wert in das Textfeld **Für Aufnahme ins Modell** ein. Wenn keine *F*-Statistik einen größeren Wert als diesen aufweist, hält das Vorwärtsverfahren an. Andernfalls fügt das Vorwärtsverfahren die Variable mit der größten *F*-Statistik zum Modell hinzu. Das Vorwärtsverfahren berechnet dann erneut *F*-Statistiken für die Variablen, die nicht im Modell verwendet werden und das Bewertungsverfahren wird wiederholt. Die Variablen werden also solange nacheinander zum Modell hinzugefügt, bis keine verbleibende Variable mehr eine signifikante *F*-Statistik erzeugt. Eine Variable, die zum Modell hinzugefügt wurde, verbleibt im Modell. ■ Das **Rückwärtsverfahren** berechnet zuerst *F*-Statistiken für ein Modell, einschließlich aller erklärenden Variablen. Die Variablen werden anschließend solange nacheinander aus dem Modell gelöscht, bis alle im Modell verbleibenden Variablen *F*-Statistiken erzeugen. Das Signifikanzniveau wird im Textfeld Für Verbleiben im Modell angegeben. Standardmäßig beträgt dieser Wert `0,05`. Bei jedem Schritt wird die am wenigsten zum Modell beitragende Variable gelöscht.

Optionsname	Beschreibung
	■ **Schrittweise Auswahl** ist eine Variante des Vorwärtsverfahrens. Bei der schrittweisen Auswahl verbleiben die bereits enthaltenen Variablen jedoch nicht notwendigerweise im Modell. Wie auch bei der Vorwärtsselektion werden Variablen nacheinander zum Modell hinzugefügt, wenn eine signifikante *F*-Statistik vorliegt. Das Signifikanzniveau wird im Textfeld Für Aufnahme ins Modell angegeben. Nachdem eine Variable hinzugefügt wurde, werden alle bereits im Modell enthaltenen Variablen überprüft. Weist eine Variable keine signifikante *F*-Statistik auf, wird sie aus dem Modell gelöscht. Das Signifikanzniveau wird im Textfeld **Für Verbleiben im Modell** angegeben. Erst nach dieser Prüfung und nach Ausführung der notwendigen Löschvorgänge, kann eine andere Variable dem Modell hinzugefügt werden. Die schrittweise Auswahl ist beendet, wenn eine der folgenden Bedingungen erfüllt ist: □ keine Variable außerhalb des Modells besitzt eine signifikante *F*-Statistik und jede Variable im Modell ist hinsichtlich des für das Verbleiben im Modell Signifikanzniveaus signifikant. □ die Variable, die zum Modell hinzugefügt werden soll, ist die zuvor aus dem Modell gelöschte Variable. ■ **Beste Teilmenge** verwendet einen Algorithmus, mit dem eine bestimmte Anzahl von Modellen mit der höchsten Chi-Quadrat-Statistik für alle möglichen Modellgrößen gesucht wird. Diese reichen von 1, 2, 3 Effektmodellen usw. bis hin zu dem Einzelmodell, das alle quantitativen Variablen enthält.
Signifikanzniveaus	■ Beim Vorwärtsverfahren geben Sie ein Signifikanzniveau an, um eine quantitative (erklärende) Variable in das Modell aufzunehmen. ■ Beim Rückwärtsverfahren geben Sie ein Signifikanzniveau an, damit eine quantitative Variable im Modell verbleibt. ■ Bei der Methode **Schrittweise Auswahl** geben Sie beide Signifikanzniveaus an.
Für das Modell erzwungene Effekte	Wählen Sie den Effekt aus, den Sie zum Modell hinzufügen möchten. Die Variable wird in das Feld **Für das Modell erzwungene Effekte** eingefügt. Sie können diese Variablen umordnen, indem Sie die Variable markieren und mit den Pfeiltasten verschieben. **Hinweis:** Diese Option ist nicht verfügbar, wenn Sie die Modellauswahlmethode **Volles Modell (keine Auswahl)** wählen.

Logistische Regression: Festlegen von Modelloptionen

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Details zu Schätzern	gibt an, ob die Korrelationsmatrix der Parameterschätzungen, die Kovarianzmatrix der Parameterschätzungen oder beides angezeigt werden soll.
Modellanpassungsbewertung	gibt folgende Modellanpassungsstatistiken an: ■ **Einfluss-Statistiken** zeigt Diagnosemaße zum Identifizieren beeinflussender Beobachtungen im Fall eines binären Response-Modells an. ■ **Hosmer-und-Lemeshow-Test auf Güte der Anpassung** zeigt den Hosmer-und-Lemeshow-Test auf Güte der Anpassung für den Fall eines binären Response-Modells an. ■ **Abweichung und Pearsonsche Güte der Anpassung** zeigt die Pearsonsche Chi-Quadrat-Statistik, die Abweichung, deren Freiheitsgrade, das Verhältnis jeder Statistik dividiert durch ihre Freiheitsgrade und den entsprechenden *p*-Wert an. ■ **Verallgemeinertes R-Quadrat** zeigt einen verallgemeinerten Bestimmtheitskoeffizienten für das angepasste Modell an. **Hinweis:** Die Optionen **Einfluss-Statistiken**, **Hosmer-/Lesmeshow-Test auf Anpassungsgüte** und **Abweichung und Pearsonsche Güte der Anpassung** sind nur dann verfügbar, wenn Sie der Rolle **Abhängige Variable** eine binäre Variable zuweisen.
Klassifizierungstabelle anzeigen	klassifiziert die binären abhängigen Eingabebeobachtungen in Abhängigkeit davon, ob die vorhergesagten Ereigniswahrscheinlichkeiten oberhalb oder unterhalb eines Cutpoint-Werts im Bereich (0,1) liegen. Eine Beobachtung wird als Ereignis vorhergesagt, wenn die vorhergesagte Ereigniswahrscheinlichkeit über dem Cutpoint-Wert liegt. In der Standardtabelle wird die Klassifizierung für einen Wahrscheinlichkeitsbereich von der kleinsten geschätzten Wahrscheinlichkeit (abgerundet auf den nächsten Wert von 0,02) bis zur höchsten geschätzten Wahrscheinlichkeit (aufgerundet auf den nächsten Wert von 0,02) in Inkrementen von 0,02 angezeigt.

Optionsname	Beschreibung
	Um andere als die in der Standardliste enthaltenen Cutpoints anzugeben, geben Sie mindestens einen Wert in das Textfeld **Kritische Wahrscheinlichkeitswerte (Cutpoints)** ein. Jeder Wert muss zwischen 0 und 1 einschließlich liegen. **Hinweis:** Diese Option ist verfügbar, wenn die Variable, die Sie der Rolle **Abhängige Variable** zugewiesen haben, binär ist.
Konfidenzgrenzen	gibt die Konfidenzintervalle für die Parameter im Bereich **Parameter** und das Odds Ratio im Bereich **Bedingte Odd Ratios** an: ■ **Wald** Berechnet Konfidenzintervalle auf der Grundlage individueller Wald-Tests. ■ **Profile Likelihood** berechnet die Konfidenzintervalle basierend auf der Profile Likelihood. Die Erstellung dieses Intervalls wird von der asymptotischen Chi-Quadrat-Verteilung des verallgemeinerten Likelihood-Ratio-Tests abgeleitet. Sie können auch einen Wert aus der Dropdown-Liste **Konfidenzniveau** wählen, um das Signifikanzniveau für die Regressionsparameter oder die Odd Ratios festzulegen.
Modellanpassungsmethoden	gibt an, wie das Modell angepasst werden soll. Standardmäßig ermittelt die Anwendungsroutine automatisch die beste Anpassungsmethode für das Modell. In der Dropdown-Liste **Anpassungstechnik** können Sie **Fisher-Scoring** oder **Newton-Raphson** als Optimierungstechnik für die Schätzung der Regressionsparameter auswählen. **Hinweis:** Wenn Sie **Ungeordnet** als Typ abhängiger Ausprägungen ausgewählt haben, ist die Optimierungstechnik **Fisher-Scoring** nicht verfügbar. Aktivieren Sie das Dialogfeld **Penalisierte Likelihood nach Firth**, um eine penalisierte Maximum-Likelihood-Schätzung nach Firth durchzuführen, die die systematischen Fehler in den Parameterschätzern aussondert. Diese Methode ist bei Trennbarkeit sinnvoll, was häufig vorkommt, wenn das Ereignis selten ist, und stellt eine Alternative zur Durchführung einer exakten logistischen Regression dar. Hinweis **Hinweis:** Das Kontrollkästchen **Penalisierte Likelihood nach Firth** ist nur verfügbar, wenn der Rolle **Abhängige Variable** eine binäre Variable zugewiesen wurde.

Logistische Regression: Erzeugen von Plots

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Standardmäßig werden alle für die aktuelle Datenauswahl geeigneten Plots in der Ausgabe berücksichtigt. Sie können jedoch durch Auswahl der Option **Benutzerdefinierte Plot-Listen** festlegen, welche Plots Sie in die Ausgabe aufnehmen möchten. Folgende Optionen können ausgewählt werden:

Plot-Name	Beschreibung
Odds Ratio-Plots	generiert Plots der Profil-Likelihood-Konfidenzgrenzen für Odds-Ratios und die Waldschen Konfidenzgrenzen für Odds-Ratios. Diese Option ist nur verfügbar, wenn Sie **Bedingte Odd Ratios** im Fenster **Modell > Optionen** gewählt haben.
ROC-Plots	generiert einen Plot der Receiver Operating Characteristic (ROC). Die Fläche unter der ROC-Kurve wird durch die Statistik *c* in der Tabelle zur Assoziation vorhergesagter Wahrscheinlichkeiten und beobachteter Abhängiger bestimmt.
Einfluss-Plots	zeigt die folgenden Plots an: ■ Index-Plots des Pearsonschen Residuums (Chi-Quadrat), das zum Identifizieren von Beobachtungen dient, die vom Modell nur unwesentlich berücksichtigt werden ■ das Abweichungsresiduum zum Identifizieren schlecht angepasster Beobachtungen ■ Leverage-Werte ■ die Verschiebungsdiagnose der Konfidenzintervalle, die den Einfluss individueller Beobachtungen auf die Regressionsschätzer misst ■ die Verschiebungsdiagnose der Konfidenzintervalle, die die generellen Änderungen in den globalen Regressionsschätzern aufgrund des Löschens der individuellen Beobachtung misst ■ die Änderung in der Chi-Quadrat-/Güte-der-Anpassung-Statistik aufgrund der Löschung einer individuellen Beobachtung ■ die Änderung der Abweichung aufgrund der Löschung der individuellen Beobachtung

Plot-Name	Beschreibung
DFBETAS-Plots	stellt den standardisierten Unterschied im Parameterschätzer aufgrund der Löschung der Beobachtung dar. Es ist hilfreich, den Effekt einer einzelnen Beobachtung für jeden einzelnen Parameterschätzer des angepassten Modells zu bestimmen.
Phat-Plots	zeigt diese Plots an: ■ die Änderung in der Chi-Quadrat-/Güte-der-Anpassung-Statistik aufgrund der Löschung einer individuellen Beobachtung ■ die Änderung der Abweichung aufgrund der Löschung der individuellen Beobachtung ■ die Verschiebung der Konfidenzintervalle ■ die Leverage gegen die vorhergesagte Ereigniswahrscheinlichkeit
Leverage-Plots	zeigt diese Plots an: ■ die Änderung in der Chi-Quadrat-/Güte-der-Anpassung-Statistik aufgrund der Löschung einer individuellen Beobachtung ■ die Änderung der Abweichung aufgrund der Löschung der individuellen Beobachtung ■ die Verschiebung der Konfidenzintervalle ■ die vorhergesagte Wahrscheinlichkeit gegen die Leverage
DPC-Plots	zeigt diese Plots an: ■ die Änderung der Chi-Quadrat-/Güte-der-Anpassung-Statistik aufgrund der Löschung einer individuellen Beobachtung gegen die vorhergesagte Ereigniswahrscheinlichkeit ■ die Änderung der Abweichung, die auf das Löschen der individuellen Beobachtung zurückzuführen ist, gegen die vorhergesagte Ereigniswahrscheinlichkeit In diesen Plots sind die Datenmarkierungen entsprechend dem Wert des Konfidenzintervallverschiebung farblich gekennzeichnet.
Effekt-Plots	zeigt Plots der vorhergesagten Wahrscheinlichkeiten für einen Wert der abhängigen Variablen gegen die Modelleffekte an.

Hinweis: ROC-Plots, Einflussplots, DFBETAS-Plots, Leverage-Plots, Phat-Plots und DPC-Plots sind nur dann verfügbar, wenn Sie der Rolle Abhängige Variable eine Variable zugewiesen haben.

Logistische Regression: Festlegen von Vorhersageoptionen

Klicken Sie im Auswahlfenster auf **Vorhersagen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Daten für Vorhersage	identifiziert die Datenquellen, die Sie zur Vorhersage der Ausgabedaten verwenden möchten. Wenn Sie **Weitere Daten** auswählen, werden diese Daten eingeschlossen.
Ausgabedaten speichern	legt fest, ob Prognosen und Diagnosestatistiken dauerhaft in einer Ausgabedatei gespeichert werden sollnen. ■ **Vorhersagen** speichert die vorhergesagte Wahrscheinlichkeit aller abhängigen Ausprägungen. Für eine abhängige Variable Y mit drei Ausprägungen, 1, 2 und 3, lauten die individuellen Wahrscheinlichkeiten Pr(Y=1), Pr(Y=2) und Pr(Y=3). ■ **Diagnosestatistiken** speichert DFBETAS-Diagnosestatistik, die Verschiebungsdiagnose der Konfidenzintervalle (C und CBAR), die *H*-Matrix und die Schätzung des linearen Prädiktors mit Standardfehlerschätzungen. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Weitere Statistiken	■ Wählen Sie **Residuen**, um folgende Statistiken zu speichern: □ Abweichungsresiduum □ Pearsonsches Residuum (Chi-Quadrat) □ die Änderung der Abweichung aufgrund der Löschung der individuellen Beobachtung □ die Änderung in der Chi-Quadrat-/Güte-der-Anpassung-Statistik aufgrund der Löschung einer individuellen Beobachtung ■ Die Auswahl von **Vorhersagegrenzen** schließt die obere und untere Grenze eines Konfidenzintervalls für jede Vorhersage

Optionsname	Beschreibung
	und den erwarteten Mittelwert der abhängigen Variable in die Ausgabe ein.
Anzeigen von Ausgabe und Plots	zeigt die Ergebnisse an. Um die Ergebnisse zu unterdrücken, deaktivieren Sie das Kontrollkästchen. Wenn Sie dieses Kontrollkästchen deaktivieren, sind die Optionen im Fenster **Plots** nicht verfügbar. Die Ergebnisse werden standardmäßig angezeigt. Um die Vorhersagen in die Ergebnisse einzuschließen, aktivieren Sie das Kontrollkästchen **Vorhersagen anzeigen**. Diese Option ist nur verfügbar, wenn im Bereich **Daten für Vorhersage** eine Datenquelle ausgewählt und das Kontrollkästchen **Ausgabe und Plots anzeigen** aktiviert ist.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die

Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

49

Kartendiagramm

Allgemeines über die Anwendungsroutine 'Kartendiagramm' ... **371**

Kartendiagramm: Auswählen des Kartentyps ... **372**

Kartendiagramm: Zuweisen von Variablen zu Analyserollen ... **373**

Kartendiagramm: Festlegen von Darstellungsoptionen ... **374**
- Festlegen von Kartenoptionen ... 374
- Festlegen des Layouts ... 375
- Anpassen der Farben ... 376
- Weitere Darstellungsoptionen ... 376

Festlegen von Titeln und Fußnoten ... **376**

Kartendiagramm: Ergebnisvorschau ... **377**

Anzeigen von Eigenschaften ... **377**

Allgemeines über die Anwendungsroutine 'Kartendiagramm'

Die Anwendungsroutine **Kartendiagramm** erzeugt eine zweidimensionale (Draufsicht) oder dreidimensionale (Block und Prisma) Farbkarte, welche die verschiedenen Werte einer abhängigen Variablen für verschiedene geografische Bereiche wie z.B. Landkreise, Bundesländer und Länder anzeigt.

Dieses Beispiel verdeutlicht die Unterschiede zwischen den beiden Kartentypen. Sie möchten eine Karte der USA erstellen, aus der für jeden Staat die Anzahl der Sondermülldeponien ersichtlich ist.

- In einer Chloroplethenkarte (Draufsicht) wird die Anzahl der Deponien in einem Staat (der abhängigen Variable) durch unterschiedliche Rottöne dargestellt. Staaten, in denen sich eine große Anzahl Sondermülldeponien befindet, werden dunkelrot markiert. Staaten, in denen sich keine Sondermülldeponie befindet, werden weiß dargestellt.

- In einer Blockkarte wird die Anzahl der Deponien in einem Staat durch die Höhe und die Farbe des Blocks in der Mitte eines Kartenbereichs dargestellt.
- In einer Prismakarte wird die Anzahl der Deponien in einem Staat durch die Höhe und die Farbe der Polyeder (erhobene Polygone) für einen Kartenbereich dargestellt.

Weitere Informationen, siehe “Auswählen des Grafikausgabeformats” auf Seite 6.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GMAP
Erforderliche SAS-Produkte	Base SAS, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	Ohne

Kartendiagramm: Auswählen des Kartentyps

Klicken Sie im Auswahlfenster auf **Kartendiagramm**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
2D-Draufsichtskarte	Erstellt eine zweidimensionale Choroplethenkarte, in der die Werte der angegebenen abhängigen Variable durch unterschiedliche Farben dargestellt werden. Um die Daten Ihrer Kartendatei nur kurz anzuzeigen, aktivieren Sie das Kontrollkästchen **Leere Karte erstellen**. Dabei ist es nicht erforderlich, eine Datei mit abhängigen Ausprägungen anzugeben. Als Ergebnis erhalten Sie eine Choroplethengrafik.
Erhebungskarte	Erstellt dreidimensionale Blockkarten, auf denen die Größenordnungsklassen der angegebenen abhängigen Variablen durch Blöcke mit unterschiedlichen Höhen und Farben angegeben werden.
3D Prismakarte	Erstellt eine dreidimensionale prismaförmige Karte, in der die Größenordnungsklassen der angegebenen abhängigen

Optionsname	Beschreibung
	Variablen durch Polyeder (erhobene Polygone) mit unterschiedlichen Höhen und Farben dargestellt werden.

Kartendiagramm: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Für die Anwendungsroutine Kartendiagramme sind zwei Eingabedateien erforderlich:

- die Datenquelle mit den abhängigen Daten, in der die Informationen enthalten sind, die in der Karte überlagert werden. Zum Beispiels kann es sich hierbei um Umsatzzahlen oder demographische Daten handeln. Standardmäßig wird die Datenquelle als abhängige Datenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben.
- die Datenquelle mit den Kartendaten, in der die X- und Y-Koordinaten für die Karte enthalten sind. Diese Karten-Datenquelle müssen Sie nach dem Öffnen der Anwendungsroutine angeben.

Die Namen der einzelnen Datenquellen werden in den Feldern **Quelle für Kartendaten** und **Abhängige - Datenquelle** angezeigt. Wenn Sie eine leere Karte erstellen, sind die Datenquelle für die Karte und für die abhängigen Daten identisch.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
ID	Indentifiziert die Variablen in der Karte und in den Dateien mit abhängigen Ausprägungen, durch die der Kartenbereich definiert wird. Jede Spalte, die Sie der Rolle "ID-Variable" zuweisen, muss sowohl in der Karte als auch in den Dateien enthalten sein. ID-Variablen können numerisch oder alphanumerisch sein und sollten denselben Namen, Typ und dieselbe Länge in Karte und Dateien aufweisen.
Ausprägung	Legt die Variable in der Datei mit abhängigen Ausprägungen fest, in der die abhängigen Werte enthalten sind, die in der Karte dargestellt werden sollen. Die Art und Weise der Darstellung der Werte der abhängigen Variable wird durch den Kartentyp bestimmt. Sie können nur eine abhängige Variable zuweisen.

Optionsname	Beschreibung
	■ In Choroplethkarten werden die Werte der abhängigen Variablen in den Kartenbereichen durch verschiedene Farben dargestellt. ■ In Blockkarten werden die Werte der abhängigen Variablen durch unterschiedliche Höhen, Muster oder Farben dargestellt. ■ In Prismakarten werdend ie Werte der abhängigen Variablen durch erhobene Kartenbereiche dargestellt.
Diagramme gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Diagramme, die erstellt werden. Für jeden eindeutigen Wert wird eine separate Grafik erzeugt.

Kartendiagramm: Festlegen von Darstellungsoptionen

Festlegen von Kartenoptionen

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Zuordnung**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Alle Kartenbereiche aus der Kartendatei einbeziehen	Legt fest, ob in der Ausgabe geographische Unterteilungen enthalten sein sollen, die keine Daten enthalten.
Eine separate abhängige Ausprägung für alle verschiedenen Werte der abhängigen Variable generieren	Legt fest, ob für jeden eindeutigen Wert der abhängigen Variable eine eigene Farbe verwendet werden soll.
Fehlende Werte als gültig in der	Legt fest, ob fehlende Werte als gültige Etiketten zu der abhängigen Variable hinzugefügt werden sollen.

Optionsname	Beschreibung
abhängigen Variable akzeptieren	
Abhängige Werte als prozentualen Anteil am Gesamten erfassen	Gibt an, ob alle abhängigen Werte (bzw. deren Statistiken) erfasst und alle Regionen grafisch als Prozentsatz des Gesamten dargestellt werden sollen. Wählen sie die Option **Darzustellende Statistiken festlegen**, wenn Sie zur Ermittlung dieses Prozentsatzes eine Statistik für die abhängigen Werte verwenden möchten. Wenn sie dies nicht möchten, wird nur der abhängige Wert der ersten Beobachtung jeder Region gezählt.
Darzustellende Statistiken festlegen	Legt die Statistiken fest, die für die abhängige Variable berechnet werden. Standardmäßig wird der abhängige Wert von nur der ersten Beobachtung in jeder Region gezählt. Wenn in der abhängigen Variable alphanumerische Werte enthalten sind, können Sie wählen, ob die erste Beobachtung oder die Häufigkeit dargestellt werden soll. Sind in der abhängigen Variable numerische Werte enthalten, können Sie die angeben, ob die erste Beobachtung, die Summe, die Häufigkeit oder der Mittelwert der abhängigen Variable dargestellt werden soll.
Anzahl der darzustellenden abhängigen Ausprägungen festlegen	Legt die Anzahl der Farben fest, die zur Darstellung der Datenwerte verwendet werden.
Dichte der verwendeten Kartenbeobachtungen festlegen	Legt den Abstand der Kartenpunkte untereinander fest. Wenn viele Kartenpunkte eng beinander liegen (hohe Dichte), ist es möglich, eine gewisse Anzahl an Punkten zu eliminieren, ohne dabei die Güte der Karte zu verschlechtern. Sie können hier eine Ganzzahl von 0 bis 6 angeben. Wenn Sie keinen Wert angeben, werden alle Beobachtungen aus der Kartendatei angezeigt. **Hinweis:** Diese Option ist nur verfügbar, wenn in der Kartendatei die Variable DENSITY enthalten ist.

Festlegen des Layouts

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Layout**, um auf diese Optionen zuzugreifen.

Für die folgenden Kartenbereiche können Sie in allen Kartentypen Farben verwenden:

- Als Auffüllung für die geographischen Bereiche, die keinen Wert aufweisen ("leere Werte")
- Als Umriss für die geographischen Bereiche, die keinen Wert aufweisen
- Als Umriss für die geographischen Bereiche, die einen Wert aufweisen

Darüber hinaus können Sie die Breite des Umrisses festlegen.

Wenn Sie eine Blockkarte erstellen, können Sie die Form des Blocks in der Karte, die Farbe für die Umrisse des Blocks sowie die Größe (Breite) der Blöcke festlegen.

Anpassen der Farben

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Farben**, um auf diese Optionen zuzugreifen.

In der Anwendungsroutine Kartendiagramm wird ein in SAS integriertes Farbschema verwendet. Sie können jedoch auch eigene, benutzerdefinierte Farben einsetzen. Aktivieren Sie dazu das Kontrollkästchen **Benutzerdefinierte Farben verwenden** und wählen Sie die gewünschten Farben aus. Die Farben werden in der Reihenfolge verwendet, in der sie in der Liste mit den benutzerdefinierten Farben aufgeführt sind.

Weitere Darstellungsoptionen

Einige Darstellungsoptionen gelten für alle Grafik-Anwendungsroutinen. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- "Festlegen von Legendenoptionen" auf Seite 749
- "Anpassen des Diagrammbereichs" auf Seite 750

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Kartendiagramm: Ergebnisvorschau

Klicken Sie im Auswahlfenster auf **Vorschau**, um eine Voransicht der Ergebnisse aufzurufen.

Überprüfen Sie die Ergebnisse in der Vorschau und nehmen Sie ggf. notwendige Änderungen vor, bevor Sie die Ergebnisse erstellen. Da Kartendateien sehr groß werden können, ist die Ergebnisvorschau ein geeignetes Mittel, um Zeit einzusparen. Wenn Sie Änderungen an der Anwendungsroutine vorgenommen haben, klicken Sie auf **Aktualisieren**, um eine neue Ergebnisvorschau zu erstellen. Tritt beim Erstellen der Ergebnisse ein Fehler auf, wird das SAS-Log angezeigt.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

50

X-quer- und R-Regelkarte

Allgemeines über die Anwendungsroutine 'X-quer- und R-Regelkarte' **379**
X-quer- und R-Regelkarte: Zuweisen von Variablen zu Analyserollen **380**
Regelkarten: Festlegen von Regelgrenzen **381**
Regelkarte: Auswählen der durchzuführenden Tests **383**
Regelkarten: Festlegen von Optionen für die Plotdarstellung **384**
Regelkarten: Speichern von Analyseergebnissen **384**
Festlegen von Titeln und Fußnoten **385**
Anzeigen von Eigenschaften **386**

Allgemeines über die Anwendungsroutine 'X-quer- und R-Regelkarte'

Die Shewhart-Regelkarte ist ein Grafik- und Analysetool, mit dem bestimmt werden kann, ob ein Prozess innerhalb der statistischen Kontrolle liegt. Die Anwendungsroutine **X-quer- und R-Regelkarte** erstellt X-quer- und R-Regelkarten für die Untergruppenmittelwerte und -spannweiten. Diese Regelkarten eignen sich zum Analysieren der zentralen Tendenz und der Variabilität eines Prozesses.

Sie können diese Anwendungsroutine beispielsweise zum Analysieren Ihrer Qualitätskontrolle verwenden. Angenommen, in der Produktion von Kristallscheiben werden für Chargen mit je fünf Kristallscheiben stichprobenweise Durchmessermessungen (in Millimeter) vorgenommen. Die Messungen für 25 Chargen werden in einer SAS-Datei gespeichert, die zum Erstellen der X-quer- und R-Regelkarten dienen. Jeder Punkt auf der X-quer-Regelkarte steht für den Durchschnitt (Mittelwert) der Messungen einer bestimmten Charge. Jeder Punkt auf der R-Regelkarte steht für die Spannweite der Messungen einer bestimmten

Charge. Wenn alle Punkte innerhalb der Regelgrenzen liegen, können Sie daraus schließen, dass sich der Prozess innerhalb der statistischen Kontrolle befindet.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SHEWHART
Erforderliche SAS-Produkte	Base SAS, SAS/QC, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

X-quer- und R-Regelkarte: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Prozessmaß	gibt die zu analysierenden Prozessvariablen an. Diese Variablen enthalten die Einzelmaße. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Untergruppen-Identifier	gibt die Variable an, welche die Untergruppen in den Daten identifiziert. Die Werte der Variablen geben an, wie die Beobachtungen in der Eingabetabelle in rationale Untergruppen gegliedert werden. Dies sind die üblichen Werte: ■ die Indizes, welche die Reihenfolge angeben, in der die Untergruppenstichproben gesammelt wurden ■ die Daten oder Zeiten, zu denen die Untergruppenstichproben gesammelt wurden

Optionsname	Beschreibung
	■ die Etiketten, welche die Untergruppenstichproben eindeutig identifizieren. Ein Maß-Identifier ist erforderlich. Sie können dieser Rolle nur genau eine Variable zuweisen. Um die Daten nach den identifizierten Untergruppen zu sortieren, aktivieren Sie das Kontrollkästchen **Nach Untergruppe sortieren**.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.
Blockvariablen	gibt bis zu zwei Variablen an, welche die Daten in Blöcke aufeinanderfolgender Untergruppen gruppieren. Diese Blöcke werden in einer Legende beschriftet. Jede Blockvariable liefert eine Etikettenausprägung in der Legende.

Regelkarten: Festlegen von Regelgrenzen

Klicken Sie im Auswahlfenster auf **Regelgrenzen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Sigma-Grenzen	gibt die Breite der Regelgrenzen als Mehrfaches des Standardfehlers (Sigma) der beschreibenden Untergruppenstatistik an, die im Diagramm dargestellt wird. Die Breite muss positiv sein. Der Standardwert für das Mehrfache ist 3.

Optionsname	Beschreibung
Berechnungsmethode auswählen	gibt die Berechnungsmethode für die Regelgrenzen an. Folgende Optionen können ausgewählt werden: ■ **Regelgrenzen aus den aktiven Daten berechnen** ■ **Regelgrenzen aus der ausgewählten Datei berechnen** Klicken Sie auf **Durchsuchen**, um die Datei anzugeben, die die Regelgrenzen bzw. die Parameter zu deren Berechnung enthält. Wählen Sie die gewünschte Datei aus und klicken Sie auf **Öffnen**. **Hinweis:** Bei der ausgewählten Datei muss es sich um eine LIMITS= Datei handeln. Weitere Informationen über die in einer LIMITS= Datei erforderlichen Variablen finden Sie in der Prozedurendokumentation zur erstellten Regelkarte. Wenn Sie die beschreibenden Statistiken für alle Untergruppen darstellen möchten, aktivieren Sie das Kontrollkästchen **Beschreibende Statistiken für alle Untergruppen darstellen, unabhängig davon, ob die Stichprobengröße der Untergruppe der nominalen Stichprobengröße für Regelgrenzen entspricht**. Sie können diese Statistiken beispielsweise darstellen, wenn fast alle Untergruppen eine gemeinsame Stichprobengröße haben und wenn Sie die festen (anstelle der variablen) Regelgrenzen, die der nominalen Stichprobengröße *n* entsprechen, anzeigen möchten. Wenn Sie diese Option auswählen und sich Ihre Stichprobengrößen für die Untergruppen deutlich unterscheiden, ist die Interpretation der Regelgrenzen nur für die Untergruppen sinnvoll, deren Stichprobengröße *n* entspricht. Um spezielle Marker für die Datenpunkte hinzuzufügen, bei denen die Stichprobengröße ungleich `n` ist, aktivieren Sie das Kontrollkästchen **Spezielle Marker für Punkte hinzufügen, die Stichprobengrößen entsprechen, welche ungleich nominalen Stichprobengrößen für feste Regelgrenzen sind**. ■ **Regelgrenzen angeben** Geben Sie in den Feldern **Oben**, **Mitte** und **Unten** die obere Regelgrenze, die Mittellinie bzw. die untere Regelgrenze an. Wenn die Stichprobengrößen Ihrer Untergruppe variieren, müssen Sie eine nominale Stichprobengröße angeben. **Hinweis:** Wenn Sie der Rolle **Analyse gruppieren nach** Variablen zuweisen, können Sie keine Werte für Regelgrenzen angeben.

Regelkarte: Auswählen der durchzuführenden Tests

Klicken Sie im Auswahlfenster auf **Tests**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Tests wenden nicht die Regelgrenzen an, die keine Drei-Sigma-Grenzen sind oder die mit der Untergruppenstichprobengröße variieren.

Optionsname	Beschreibung
Test auswählen	Fordern Sie über die Kontrollkästchen einen oder mehrere Tests auf besondere Ursachen an, die auch als Ausführungstests, Mustertests und Western-Electric-Regeln bekannt sind. Diese Tests entdecken bestimmte nichtzufällige Muster in den Punkten, die in der primären Regelkarte dargestellt werden. Das Vorkommen eines nichtzufälligen Musters wird als Signal bezeichnet. Die Tests können auf das Vorhandensein einer besonderen Variationsursache hinweisen. Wenn Sie die Tests auswählen, erscheint eine Beschreibung des jeweiligen Tests im Feld **Beschreibung**.
Etikett	Geben Sie in den Textfeldern die Etiketten für Punkte an, bei denen die Tests auf besondere Ursachen positiv sind. Die Etiketten dürfen maximal 16 Zeichen lang sein. Jedes Etikett erscheint bei allen Punkten im Test, an denen ein Signal vorliegt.
Text zum Identifizieren der Testsignale	■ Wählen Sie über die Dropdown-Liste eine Farbe für die angegebenen Etiketten aus. ■ Markieren Sie das Kontrollkästchen **Zonenlinien anzeigen**, um Linien einzuschließen, welche die Zonen A, B und C in einem primären Diagramm kennzeichnen. Diese Zonen werden gebildet, indem die Intervalle zwischen den Regelgrenzen in sechs gleichmäßig verteilte Unterintervalle unterteilt werden. ■ Verwenden Sie das Kontrollkästchen **3 Sigma-Grenzen überschreiben**, um nach speziellen Ursachen zu suchen, wenn Sie eine andere Sigma-Regelgrenze als das Standardvielfache von 3 angeben. Um die Sigma-Grenze zu ändern, klicken Sie im Auswahlfenster auf **Regelgrenzen**.

Optionsname	Beschreibung
	■ Über das Kontrollkästchen **Test auf überlappende Punktmuster anwenden** wenden Sie Tests auf besondere Ursachen auf überlappende Punktmuster an.

Regelkarten: Festlegen von Optionen für die Plotdarstellung

Einige der Darstellungsoptionen sind für unterschiedliche Regelkarten-Anwendungsroutinen gleich. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Festlegen von Achsenoptionen” auf Seite 759
- “Hinzufügen von Referenzlinien” auf Seite 760
- “Festlegen weiterer Plotoptionen” auf Seite 760
- “Festlegen von Blockvariablenoptionen” auf Seite 762

Regelkarten: Speichern von Analyseergebnissen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

Um die Analyseergebnisse zu speichern, können Sie Ausgabedateien mit diversen Statistiken speichern. Markieren Sie die Kontrollkästchen der Ausgabedaten, die Sie permanent speichern möchten.

Optionsname	Beschreibung
Datei mit Untergruppenstatistiken und Regelgrenzen	erstellt eine Ausgabedatei, welche die im Diagramm dargestellten Informationen enthält. Dies beinhaltet die Werte der Untergruppenvariablen und ihre entsprechenden beschreibenden Statistiken und Regelgrenzen. **Hinweis:** Diese Option erstellt eine Datei OUTTABLE=. Sie können diese Datei mit der Option TABLE= aus der

Optionsname	Beschreibung
	SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Untergruppenstatistiken	erstellt eine Ausgabedatei mit den beschreibenden Statistiken der Untergruppe. **Hinweis:** Diese Option erstellt eine Datei mit OUTHISTORY=. Sie können diese Datei mit der Option HISTORY= aus der SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Regelgrenzen	erstellt eine Ausgabedatei, welche die Regelgrenzen speichert. Sie können diese Datei verwenden, wenn Sie die Berechnungsmethode für die Regelgrenzen später in dieser Anwendungsroutine auswählen.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

   Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

51

X-quer- und S-Regelkarte

Allgemeines über die 'X-quer- und S-Regelkarte' 387
X-quer- und S-Regelkarte: Zuweisen von Variablen zu Analyserollen 388
Regelkarten: Festlegen von Regelgrenzen 389
Regelkarte: Auswählen der durchzuführenden Tests 391
Regelkarten: Festlegen von Optionen für die Plotdarstellung 392
Regelkarten: Speichern von Analyseergebnissen 392
Festlegen von Titeln und Fußnoten 393
Anzeigen von Eigenschaften 394

Allgemeines über die 'X-quer- und S-Regelkarte'

Die Shewhart-Regelkarte ist ein Grafik- und Analysetool, mit dem bestimmt werden kann, ob ein Prozess innerhalb der statistischen Kontrolle liegt. Die Anwendungsroutine **X-quer- und S-Regelkarte** erstellt X-quer- und S-Regelkarten für die Untergruppenmittelwerte und -abweichungen. Diese Regelkarten eignen sich zum Analysieren der zentralen Tendenz und der Variabilität eines Prozesses.

Mithilfe dieser Anwendungsroutine können Sie die Verteilung der Ausgabe ermitteln und feststellen, ob ein Prozess innerhalb der statistischen Kontrolle liegt. Angenommen, ein Erdölunternehmen verwendet eine Turbine zum Umwandeln von Wasser in Dampf, der in den Boden gepumpt wird, um das Öl dünnflüssiger zu machen, sodass es leichter extrahiert werden kann. Dieser Prozess ist täglich 20 Mal erforderlich, und die Leistungsmenge (in Kilowatt), die zum Erhitzen des Wassers bis zur gewünschten Temperatur erforderlich ist, wird aufgezeichnet. Jeder Punkt auf der X-quer-Regelkarte stellt den Mittelwert der Messungen eines bestimmten Tages dar. Jeder Punkt auf der S-Regelkarte stellt die Standardabweichung der Messungen eines bestimmten Tages dar. Wenn alle Punkte innerhalb der Regelgrenzen liegen (standardmäßig ist dies 3 Sigma), kann

daraus geschlossen werden, dass der Prozess innerhalb der statistischen Kontrolle liegt.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SHEWHART
Erforderliche SAS-Produkte	Base SAS, SAS/QC, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

X-quer- und S-Regelkarte: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Prozessmaß	gibt die zu analysierenden Prozessvariablen an. Diese Variablen enthalten die Einzelmaße. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Untergruppen-Identifier	gibt die Variable an, welche die Untergruppen in den Daten identifiziert. Die Werte der Variablen geben an, wie die Beobachtungen in der Eingabetabelle in rationale Untergruppen gegliedert werden. Dies sind die üblichen Werte: ■ die Indizes, welche die Reihenfolge angeben, in der die Untergruppenstichproben gesammelt wurden ■ die Daten oder Zeiten, zu denen die Untergruppenstichproben gesammelt wurden

Optionsname	Beschreibung
	■ die Etiketten, welche die Untergruppenstichproben eindeutig identifizieren. Ein Maß-Identifier ist erforderlich. Sie können dieser Rolle nur genau eine Variable zuweisen. Um die Daten nach den identifizierten Untergruppen zu sortieren, aktivieren Sie das Kontrollkästchen **Nach Untergruppe sortieren**.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.
Blockvariablen	gibt bis zu zwei Variablen an, welche die Daten in Blöcke aufeinanderfolgender Untergruppen gruppieren. Diese Blöcke werden in einer Legende beschriftet. Jede Blockvariable liefert eine Etikettenausprägung in der Legende.

Regelkarten: Festlegen von Regelgrenzen

Klicken Sie im Auswahlfenster auf **Regelgrenzen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Sigma-Grenzen	gibt die Breite der Regelgrenzen als Mehrfaches des Standardfehlers (Sigma) der beschreibenden Untergruppenstatistik an, die im Diagramm dargestellt wird. Die Breite muss positiv sein. Der Standardwert für das Mehrfache ist 3.
Berechnungsmethode auswählen	gibt die Berechnungsmethode für die Regelgrenzen an. Folgende Optionen können ausgewählt werden:

Optionsname	Beschreibung
	■ **Regelgrenzen aus den aktiven Daten berechnen** ■ **Regelgrenzen aus der ausgewählten Datei berechnen** Klicken Sie auf **Durchsuchen**, um die Datei anzugeben, die die Regelgrenzen bzw. die Parameter zu deren Berechnung enthält. Wählen Sie die gewünschte Datei aus und klicken Sie auf **Öffnen**. **Hinweis:** Bei der ausgewählten Datei muss es sich um eine LIMITS= Datei handeln. Weitere Informationen über die in einer LIMITS= Datei erforderlichen Variablen finden Sie in der Prozedurendokumentation zur erstellten Regelkarte. Wenn Sie die beschreibenden Statistiken für alle Untergruppen darstellen möchten, aktivieren Sie das Kontrollkästchen **Beschreibende Statistiken für alle Untergruppen darstellen, unabhängig davon, ob die Stichprobengröße der Untergruppe der nominalen Stichprobengröße für Regelgrenzen entspricht**. Sie können diese Statistiken beispielsweise darstellen, wenn fast alle Untergruppen eine gemeinsame Stichprobengröße haben und wenn Sie die festen (anstelle der variablen) Regelgrenzen, die der nominalen Stichprobengröße *n* entsprechen, anzeigen möchten. Wenn Sie diese Option auswählen und sich Ihre Stichprobengrößen für die Untergruppen deutlich unterscheiden, ist die Interpretation der Regelgrenzen nur für die Untergruppen sinnvoll, deren Stichprobengröße *n* entspricht. Um spezielle Marker für die Datenpunkte hinzuzufügen, bei denen die Stichprobengröße ungleich `n` ist, aktivieren Sie das Kontrollkästchen **Spezielle Marker für Punkte hinzufügen, die Stichprobengrößen entsprechen, welche ungleich nominalen Stichprobengrößen für feste Regelgrenzen sind**. ■ **Regelgrenzen angeben** Geben Sie in den Feldern **Oben**, **Mitte** und **Unten** die obere Regelgrenze, die Mittellinie bzw. die untere Regelgrenze an. Wenn die Stichprobengrößen Ihrer Untergruppe variieren, müssen Sie eine nominale Stichprobengröße angeben. **Hinweis:** Wenn Sie der Rolle **Analyse gruppieren nach** Variablen zuweisen, können Sie keine Werte für Regelgrenzen angeben.

Regelkarte: Auswählen der durchzuführenden Tests

Klicken Sie im Auswahlfenster auf **Tests**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Tests wenden nicht die Regelgrenzen an, die keine Drei-Sigma-Grenzen sind oder die mit der Untergruppenstichprobengröße variieren.

Optionsname	Beschreibung
Test auswählen	Fordern Sie über die Kontrollkästchen einen oder mehrere Tests auf besondere Ursachen an, die auch als Ausführungstests, Mustertests und Western-Electric-Regeln bekannt sind. Diese Tests entdecken bestimmte nichtzufällige Muster in den Punkten, die in der primären Regelkarte dargestellt werden. Das Vorkommen eines nichtzufälligen Musters wird als Signal bezeichnet. Die Tests können auf das Vorhandensein einer besonderen Variationsursache hinweisen. Wenn Sie die Tests auswählen, erscheint eine Beschreibung des jeweiligen Tests im Feld **Beschreibung**.
Etikett	Geben Sie in den Textfeldern die Etiketten für Punkte an, bei denen die Tests auf besondere Ursachen positiv sind. Die Etiketten dürfen maximal 16 Zeichen lang sein. Jedes Etikett erscheint bei allen Punkten im Test, an denen ein Signal vorliegt.
Text zum Identifizieren der Testsignale	■ Wählen Sie über die Dropdown-Liste eine Farbe für die angegebenen Etiketten aus. ■ Markieren Sie das Kontrollkästchen **Zonenlinien anzeigen**, um Linien einzuschließen, welche die Zonen A, B und C in einem primären Diagramm kennzeichnen. Diese Zonen werden gebildet, indem die Intervalle zwischen den Regelgrenzen in sechs gleichmäßig verteilte Unterintervalle unterteilt werden. ■ Verwenden Sie das Kontrollkästchen **3 Sigma-Grenzen überschreiben**, um nach speziellen Ursachen zu suchen, wenn Sie eine andere Sigma-Regelgrenze als das Standardvielfache von 3 angeben. Um die Sigma-Grenze zu ändern, klicken Sie im Auswahlfenster auf **Regelgrenzen**.

Optionsname	Beschreibung
	■ Über das Kontrollkästchen **Test auf überlappende Punktmuster anwenden** wenden Sie Tests auf besondere Ursachen auf überlappende Punktmuster an.

Regelkarten: Festlegen von Optionen für die Plotdarstellung

Einige der Darstellungsoptionen sind für unterschiedliche Regelkarten-Anwendungsroutinen gleich. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Festlegen von Achsenoptionen” auf Seite 759
- “Hinzufügen von Referenzlinien” auf Seite 760
- “Festlegen weiterer Plotoptionen” auf Seite 760
- “Festlegen von Blockvariablenoptionen” auf Seite 762

Regelkarten: Speichern von Analyseergebnissen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

Um die Analyseergebnisse zu speichern, können Sie Ausgabedateien mit diversen Statistiken speichern. Markieren Sie die Kontrollkästchen der Ausgabedaten, die Sie permanent speichern möchten.

Optionsname	Beschreibung
Datei mit Untergruppenstatistiken und Regelgrenzen	erstellt eine Ausgabedatei, welche die im Diagramm dargestellten Informationen enthält. Dies beinhaltet die Werte der Untergruppenvariablen und ihre entsprechenden beschreibenden Statistiken und Regelgrenzen. **Hinweis:** Diese Option erstellt eine Datei OUTTABLE=. Sie können diese Datei mit der Option TABLE= aus der

Optionsname	Beschreibung
	SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Untergruppenstatistiken	erstellt eine Ausgabedatei mit den beschreibenden Statistiken der Untergruppe. **Hinweis:** Diese Option erstellt eine Datei mit OUTHISTORY=. Sie können diese Datei mit der Option HISTORY= aus der SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Regelgrenzen	erstellt eine Ausgabedatei, welche die Regelgrenzen speichert. Sie können diese Datei verwenden, wenn Sie die Berechnungsmethode für die Regelgrenzen später in dieser Anwendungsroutine auswählen.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

   Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

52

Gemischte Modelle

Allgemeines über die Anwendungsroutine 'Gemischte Modelle' ... 395
Gemischte Modelle: Zuweisen von Variablen zu Analyserollen ... 396
Gemischte Modelle: Bilden von Modellen mit festen Effekten ... 397
Gemischte Modelle: Festlegen von Optionen für Modelle mit festen Effekten ... 397
Gemischte Modelle: Angeben zufälliger Effekte ... 399
Gemischte Modelle: Angeben wiederholter Effekte ... 401
Gemischte Modelle: Angeben der Mittelwerte der kleinsten Quadrate ... 403
Gemischte Modelle: Erzeugen von Plots ... 405
Gemischte Modelle: Festlegen von Vorhersageoptionen ... 406
Festlegen von Titeln und Fußnoten ... 407
Anzeigen von Eigenschaften ... 407

Allgemeines über die Anwendungsroutine 'Gemischte Modelle'

Ein gemischt lineares Modell ist eine Verallgemeinerung des standardlinearen Modells. Verallgemeinerung bedeutet, dass es mit den Daten möglich ist, Korrelationen und nicht konstante Variabilitäten aufzuzeigen. Das gemischt lineare Modell bietet Ihnen die Flexibilität, neben Mittelwerten auch Varianzen und Kovarianzen Ihrer Daten zu modellieren.

Sie können diese Anwendungsroutine einsetzen, um eine Interaktion zwischen zwei Haupteffekten zu modellieren. Angenommen, Sie haben die Größe von 18 Personen gemessen, die nach Familie und Geschlecht klassifiziert wurden. Wenn Sie die Höhe als abhängige Variable, Geschlecht als festen Klassifizierungseffekt und Familie als zufälligen Klassifizierungseffekt verwenden, können Sie mithilfe der Anwendungsroutine **Gemischte Modelle** ein gemischtes lineares Modell mit diesen Daten verknüpfen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	MIXED
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Gemischte Modelle: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Abhängige Variable	gibt die Variable an, die als abhängige (Response) Variable verwendet wird. Sie müssen dieser Rolle genau eine Variable zuweisen.
Quantitative Variablen	gibt die Variablen an, die als kontinuierliche unabhängige (erklärende) Variablen verwendet werden sollen.
Klassifizierungsvariablen	gibt die Variablen an, die als diskrete unabhängige Effekte verwendet werden. Variablen, die Sie dieser Rolle zuweisen, können numerisch oder alphanumerisch sein, sollten aber eine begrenzte Anzahl an diskreten Werten aufweisen.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen.

Optionsname	Beschreibung
	Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits als abhängige Variable oder als unabhängige Variable ausgewählt ist.
Relative Gewichtung	gibt die Variable an, die verwendet wird, um die entsprechenden Werte der Variablen zu gewichten, die Sie als abhängige Variable ausgewählt haben. Sie können dieser Rolle maximal eine Variable zuweisen.
ID-Etikett	gibt die zusätzlichen Variablen an, die in die Tabelle mit den vorhergesagten Werten eingeschlossen werden.

Gemischte Modelle: Bilden von Modellen mit festen Effekten

Klicken Sie im Auswahlfenster auf **Modell mit festen Effekten**, um auf diese Optionen zuzugreifen.

Sie können Modelle in mehreren SAS-Anwendungsroutinen erstellen. Weitere Informationen, siehe Anhang 3, “Erstellen eines Modells” auf Seite 765.

Gemischte Modelle: Festlegen von Optionen für Modelle mit festen Effekten

Klicken Sie im Auswahlfenster auf **Modell mit festen Effekten - Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Tests für Konstante anzeigen	Wählen Sie diese Option, um die Ergebnisse aus Hypothesentests anzuzeigen, die an der Konstanten durchgeführt wurden.
Hypothesentesttyp	Bei jedem Hypothesentest werden eine oder mehrere schätzbaren Funktionen angewendet. Schätzbare Funktionen sind lineare Kombinationen der Modellparameter. Typ 3 ist der Standard-Hypothesentest. Sie können Hypothesentests hinzufügen oder entfernen, indem Sie das entsprechende Kontrollkästchen aktivieren bzw. deaktivieren.
Parameterschätzer anzeigen	Der Wert des Parameterschätzers ist abhängig von der gewählten Schätzmethode.
Konfidenzgrenzen für Parameterschätzer	Ein Konfidenzintervall ist eine Ober- und Untergrenze, die einen Bereich angemessener Werte für den Parameter angeben, der mit einem zuvor angegebenen Konfidenzniveau geschätzt wird. Sie können einen Wert für das Konfidenzniveau angeben oder einen Wert aus der Dropdown-Liste auswählen. Der Standardwert ist ein Konfidenzniveau von 95%.
Schätzmethode	Eine Schätzmethode ist eine Technik, die verwendet wird, um die Elemente der Varianz- und Kovarianzmatrixanteile des Modells zu schätzen. Die Standardmethode ist **Residuen-Maximum-Likelihood**. Sie können aus der Dropdown-Liste eine andere Schätzmethode auswählen.
Freiheitsgradmethode	Freiheitsgrade sind eine Statistik, die sich auf die Anzahl der unabhängigen Informationsanteile in einer Summe bezieht. Sie können eine der folgenden Methoden aus der Dropdown-Liste auswählen. Die Standardmethode ist **Standard**. Wenn Sie diese Option aktivieren, wählt die Anwendungsroutine **Gemischte Modelle** die Standardmethode in Abhängigkeit von Ihrer Modellspezifikation und von der Tatsache aus, ob Ihr Modell zufällige und wiederholte Effekte enthält. Die folgende Liste zeigt, wie die Standardmethode in Abhängigkeit von existierenden zufälligen oder wiederholten Effekten ausgewählt wird: ■ Wenn Ihr Modell zufällige Effekte enthält, ist die Containmentmethode die Standardmethode. ■ Wenn Ihr Modell nur wiederholte Effekte enthält, ist die Methode Between- und Within-Subject-Anteile die Standardmethode. ■ Wenn Ihr Modell sowohl zufällige als auch wiederholge Effekte enthält, ist die Standardmethode **Containmentmethode**.

Optionsname	Beschreibung
	■ Enthält Ihr Modell weder einen zufälligen noch einen wiederholten Effekt, ist Residuum die Standarmethode.
Schätzbare Funktionen	Aktivieren Sie die Kontrollkästchen **Typ I**, **Typ II** und **Typ III**, um die Matrixkoeffizienten der schätzbaren Funktionen von Typ I - Typ IV für jedes Modell anzuzeigen. Aktivieren Sie jeweils das Kontrollkästchen neben der Typbezeichnung, um die entsprechenden Matrixkoeffizienten einer bestimmten schätzbaren Funktion anzuzeigen. **Hinweis:** Sie können auch die schätzbaren Funktionen für die Konstante anzeigen, indem Sie das Kontrollkästchen **Tests für Konstante anzeigen** aktivieren.
Tuning-Parameter verwenden	**Hinweis:** Die Tuning-Parameter eignen sich zur Lösung schlecht konditionierter Probleme. Um die Empfindlichkeit (Sensitivität) bei der Bildung von Typ III-Funktionen zu ändern, geben Sie den neuen Wert in das Feld **Empfindlichkeit für Typ-III-Funktion** ein. Alle Elemente in der schätzbaren Funktionsbasis, deren absoluter Wert kleiner als der angegebene Wert ist, werden auf 0 gesetzt. Der Standardwert ist `1E-08`. Um die Empfindlichkeit (Sensitivität) bei der Bestimmung einer Matrix zu ändern, geben Sie den gewünschten Wert in das Feld **Singularitätsprüfwert** ein. Wenn ein diagonales Pivotelement kleiner ist als die korrigierte Quadratsumme für die Variable, multipliziert mit dem angegebenen Wert, dann wird die verknüpfte Spalte für die vorherigen Spalten als linear abhängig erklärt, der Verknüpfte Parameter wird auf 0 gesetzt. Der Standardwert ist `1E-07`.

Gemischte Modelle: Angeben zufälliger Effekte

Klicken Sie im Auswahlfenster auf **Zufällige Effekte**, um auf diese Optionen zuzugreifen.

1 Klicken Sie auf **Hinzufügen**, um dem Feld **Zu schätzende Zufallseffekte** einen Effekt hinzuzufügen.

2 Für jeden Effekt in der Liste geben Sie im Feld **Zufällige Effekte und Optionen** Optionen an. Diese Optionen sind verfügbar:

Optionsname	Beschreibung
Zufällige Effekte	Geben Sie die zufälligen Effekte in Ihrem gemischten Modell an. Standardmäßig werden keine zufälligen Effekte verwendet. Sie können auch eine Konstante als zufälligen Effekt in das Modell einschließen. Verwenden Sie **Effekte erstellen**, um zufällige Effekte für das gemischte Modell anzugeben. Um einen Effekt anzugeben, müssen Sie der Rolle **Quantitative Variablen** oder **Klassifizierungsvariablen** mindestens eine Variable zuweisen. Sie können Kombinationen der Variablen auswählen, um Kreuz-, Verschachtelungs-, Faktor- oder Polynomeffekte zu erzeugen. Nachdem Sie die zufälligen Effekte angegeben haben, klicken Sie auf **OK**, um die Änderungen zu speichern und zur Anwendungsroutine **Gemischte Modelle** zurückzukehren. Informationen zu den für ein Modellerstellung erforderlichen Schritten, siehe "Gemischte Modelle: Bilden von Modellen mit festen Effekten" auf Seite 397.
Konfidenzgrenzen	Geben Sie an, ob die *t*-Typ-Konfidenzgrenzen für alle zufälligen Effektschätzer angezeigt werden sollen.
Modellobjekte	Geben Sie die Objekt-ID an, die in die Analyse einbezogen werden soll. Die Objekte müssen voneinander vollkommen unabhängig sein. Diese Option erzeugt in der *G*-Matrix eine blockdiagonale Struktur mit identischen Blöcken. Die *Z*-Matrix wird modifiziert, um diese Blockdiagonalität zu berücksichtigen. Verwenden Sie **Effekte erstellen**, um die Objekt-ID für das gemischte Modell anzugeben. Um einen Effekt anzugeben, müssen Sie der Rolle **Klassifizierungsvariablen** mindestens eine Variable zuweisen. Sie können Kombinationen der Variablen auswählen, um Kreuz-, Verschachtelungs-, Faktor- oder Polynomeffekte zu erzeugen. Der angegebene Effekt wird im Feld **Subjekt-ID** angezeigt. Um diesen Effekt zu entfernen, klicken Sie auf **Effekte entfernen**. Klicken Sie auf **OK**, um diese Objekt-ID zu speichern und zur Anwendungsroutine Gemischte Modelle zurückzukehren.
Gruppeneffekt	geben Sie den Gruppen-Identifier an, der in die Analyse einbezogen werden soll. Mit dieser Option können Sie einen Effekt definieren, der die Heterogenität in der Kovarianzstruktur der *G*-Matrix angibt. Alle Beobachtungen mit der gleichen Ausprägung des Gruppeneffekts haben die gleichen Kovarianzparameter. Jede neue Ausprägung des Gruppeneffekts erzeugt eine neue Reihe von Kovarianzparametern, die die gleiche Struktur wie die ursprüngliche Gruppe aufweisen.

Optionsname	Beschreibung
	Verwenden Sie **Effekte erstellen**, um den Gruppen-Identifier für das gemischte Modell anzugeben. Um einen Effekt anzugeben, müssen Sie der Rolle **Klassifizierungsvariablen** mindestens eine Variable zuweisen. Sie können Kombinationen der Variablen auswählen, um Kreuz-, Verschachtelungs-, Faktor- oder Polynomeffekte zu erzeugen. Der angegebene Effekt wird im Feld **Gruppen-Identifier** angezeigt. Um diesen Effekt zu entfernen, klicken Sie auf **Effekte entfernen**. Klicken Sie auf **OK**, um diesen Gruppen-Identifier zu speichern und zur Anwendungsroutine Gemischte Modelle zurückzukehren.
Kovarianzstruktur	Geben Sie die Kovarianzstruktur der *G*-Matrix an.
Weitere Optionen	Geben Sie an, ob die *G*-Matrix in der Analyse angezeigt werden soll und welche Optionen für die *G*-Matrix verwendet werden sollen. Diese Optionen sind verfügbar: ■ **G-Matrix anzeigen** zeigt die geschätzte *G*-Matrix an. ■ **Cholesky-Wurzel der G-Matrix anzeigen** zeigt die Cholesky-Lösung (untere Dreiecksmatrix) der geschätzten *G*-Matrix an. ■ **Inverse Cholesky-Wurzel der G-Matrix anzeigen** zeigt die inverse Cholesky-Wurzel der geschätzten *G*-Matrix an. ■ **Inverse der G-Matrix anzeigen** zeigt das Inverse der geschätzten *G*-Matrix an. ■ **Korrelationsmatrix anzeigen** zeigt die Korrelationsmatrix an, die der geschätzten *G*-Matrix entspricht. ■ **Schätzwerte einschließen** zeigt die Lösung für die zufälligen Effektparameter an.

Gemischte Modelle: Angeben wiederholter Effekte

Klicken Sie im Auswahlfenster auf **Wiederholte Effekte**, um auf diese Optionen zuzugreifen.

Für jeden Effekt in der Liste geben Sie im Feld **Wiederholte Effekte und Optionen** Optionen an. Diese Optionen sind verfügbar:

Optionsname	Beschreibung
Zu verwendende Effekte	Wählen Sie den Within-Subjects-Effekt, der in der Analyse verwendet werden soll. Die Within-Subjects-Effekte sind die Variablen, die Sie der Rolle **Klassifizierungsvariablen** zugewiesen haben. Sie können ein Within-Subjects-Effekt angeben, wenn Ihre Daten nicht nach dem geeigneten wiederholten Maß innerhalb eines Subjekts sortiert worden ist. Wenn z. B. Time der wiederholte Effekt ist und die Daten nicht innerhalb jedes Subjekts nach Zeit sortiert sind, müssen Sie Time als Within-Subjects-Effekt festlegen, damit das Modell die Daten richtig interpretieren kann. Standardmäßig werden keine Within-Subjects-Effekte verwendet.
Modellobjekte	Geben Sie die Objekt-ID an, die in die Analyse einbezogen werden soll. Bei vielen wiederholten Messungsmodellen sind keine wiederholten Effekte erforderlich. Wenn Sie eine Subjekt-ID angeben, wird mit der wiederholten Anweisung die *R*-Matrix im gemischten Modell angegeben. Die Subjekt-ID legt die Blöcke in der *R*-Matrix fest. Die Objekte müssen voneinander vollkommen unabhängig sein. Verwenden Sie **Effekte erstellen**, um die Objekt-ID für das gemischte Modell anzugeben. Um einen Effekt anzugeben, müssen Sie der Rolle **Klassifizierungsvariablen** mindestens eine Variable zuweisen. Sie können Kombinationen der Variablen auswählen, um Kreuz-, Verschachtelungs-, Faktor- oder Polynomeffekte zu erzeugen. Der angegebene Effekt wird im Feld **Subjekt-ID** angezeigt. Um diesen Effekt zu entfernen, klicken Sie auf **Effekte entfernen**. Klicken Sie auf **OK**, um diese Objekt-ID zu speichern und zur Anwendungsroutine Gemischte Modelle zurückzukehren.
Kovarianzstruktur	Geben Sie die Kovarianzstruktur der *R*-Matrix an.
Gruppeneffekt	Geben Sie den Gruppen-Identifier an, der in die Analyse einbezogen werden soll. Mit dieser Option können Sie einen Effekt definieren, der die Heterogenität in der Kovarianzstruktur der *G*-Matrix angibt. Alle Beobachtungen mit der gleichen Ausprägung des Gruppeneffekts haben die gleichen Kovarianzparameter. Jede neue Ausprägung des Gruppeneffekts erzeugt eine neue Reihe von Kovarianzparametern, die die gleiche Struktur wie die ursprüngliche Gruppe aufweisen. Verwenden Sie **Effekte erstellen**, um den Gruppen-Identifier für das gemischte Modell anzugeben. Um einen Effekt anzugeben, müssen Sie der Rolle **Klassifizierungsvariablen** mindestens eine Variable zuweisen. Sie können Kombinationen der Variablen auswählen, um Kreuz-, Verschachtelungs-, Faktor- oder Polynomeffekte zu erzeugen. Der angegebene Effekt wird im Feld **Gruppen-Identifier** angezeigt. Um diesen Effekt zu entfernen, klicken Sie auf

Optionsname	Beschreibung
	Effekte entfernen. Klicken Sie auf **OK**, um diesen Gruppen-Identifier zu speichern und zur Anwendungsroutine Gemischte Modelle zurückzukehren.
R-Matrixoptionen	Geben Sie an, ob die *R*-Matrix in der Analyse angezeigt werden soll, und welche *R*-Matrixoptionen Sie verwenden möchten. Diese Optionen sind verfügbar: ■ **R-Matrix anzeigen** zeigt die geschätzte *R*-Matrix an. ■ **Korrelationsmatrix anzeigen** zeigt die Korrelationsmatrix an, die der geschätzten *R*-Matrix entspricht. ■ **Inverse der R-Matrix anzeigen** zeigt das Inverse der geschätzten *R*-Matrix an. ■ **Cholesky-Wurzel der R-Matrix anzeigen** zeigt die Cholesky-Lösung (untere Dreiecksmatrix) der geschätzten *R*-Matrix an.

Gemischte Modelle: Angeben der Mittelwerte der kleinsten Quadrate

Klicken Sie im Auswahlfenster auf **Post-Hoc-Tests der kleinsten Quadrate**, um auf diese Optionen zuzugreifen.

1 Klicken Sie auf **Hinzufügen**, um dem Feld **Zu schätzende Effekte** einen Effekt hinzuzufügen.

2 Für jeden Effekt in der Liste geben Sie im Feld **Prüfung und Optionen des Mittelwerts der kleinsten Quadrate** Optionen an. Diese Optionen sind verfügbar:

Optionsname	Beschreibung
Zu verwendende Effekte	Wählen Sie die Klasseneffekte, die in der Analyse verwendet werden sollen. Die Klasseneffekte sind die Variablen, die Sie der Rolle **Klassifizierungsvariablen** zugewiesen haben. Standardmäßig werden keine Klasseneffekte verwendet.
Vergleiche	Geben Sie die Vergleichsmethode an, die in der Analyse verwendet werden soll. ■ **p-Werte für Differenzen anzeigen** - zeigt die Differenzen der Mittelwerte der kleinsten Quadrate an. Diese Optionen sind verfügbar:

Optionsname	Beschreibung
	◻ **Ohne** zeigt keine Differenzen an. ◻ **Standard** zeigt alle paarweisen Differenzen an. ◻ **Alle paarweisen Differenzen** zeigt alle paarweisen Differenzen an. Hierbei handelt es sich um die Standardeinstellung. ◻ **Erste Ausprägung als Kontrolle** - zeigt die Differenzen mit einer Kontrolle an, die standardmäßig die erste Ausprägung aller angegebenen Kleinste-Quadrate-Mittelwert-Effekte ist. Mit diesen Differenzentypen sind zweiseitige Tests und Konfidenzgrenzen verknüpft. ◻ **Test, ob kleiner als Kontrolle** - prüft, ob die nicht als Kontrolle verwendeten Ausprägungen signifikant kleiner sind als die Kontrolle. Sie können diesen Test für einseitige Ergebnisse verwenden. ◻ **Test, ob größer als Kontrolle** - prüft, ob die nicht als Kontrolle verwendeten Ausprägungen signifikant größer sind als die Kontrolle. Sie können diesen Test für einseitige Ergebnisse verwenden. ■ **Korrekturmethode für Vergleich** ermöglicht mehrfache Vergleichskorrekturen der *p*-Werte und Konfidenzgrenzen für die Differenzen der Kleinste-Quadrate-Mittelwerte.
Konfidenzgrenzen	Geben Sie an, ob die *t*-Typ-Konfidenzgrenzen für die Kleinste-Quadrate-Mittelwerte eingeschlossen werden sollen.
Kovariate	Geben Sie an, ob Sie die Kovariate in die Berechnung einschließen möchten, und ob die *L*-Matrixkoeffizienten angezeigt werden sollen. ■ **Kovariate** - Geben Sie an, wie die Kovariate berechnet werden sollen. Diese Optionen sind verfügbar: ◻ **Mittelwerte** - setzt zur Berechnung der Standard-Kleinste-Quadrate-Mittelwerte alle kovariaten Effekte mit ihren Mittelwerten gleich. ◻ **Mittelwerte und korrigierte Kreuzprodukte** - Setzt die kovariaten Effekte mit ihren Mittelwerten gleich und schließt eine Korrektur der Kreuzprodukte der Kovariate ein. ◻ **Pro Variable** - ermöglicht die Angabe willkürlicher Werte für die Kovariaten. In der Ausgabe weisen zusätzliche Spalten auf die Werte der Kovariaten hin. ■ **Angezeigte L-Matrixkoeffizienten** - zeigt die *L*-Matrixkoeffizienten für alle Kleinste-Quadrate-Mittelwerteffekte an.

Gemischte Modelle: Erzeugen von Plots

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Standardmäßig werden alle für die aktuelle Datenauswahl geeigneten Plots in der Ausgabe berücksichtigt. Sie können jedoch durch Auswahl der Option **Benutzerdefinierte Plot-Listen** festlegen, welche Plots Sie in die Ausgabe aufnehmen möchten. Sie können folgende Optionen wählen:

Optionsname	Beschreibung
Box-Plot	Erstellt Box-Plots für die Effekte in Ihrem Modell, die ausschließlich aus Klassifizierungseffekten bestehen. Es ist möglich, dass diese Effekte mehr als eine Klassifizierungsvariable einbeziehen (Interaktions- und verschachtelte Effekte). Es können jedoch keine kontinuierlichen Variablen einbezogen sein. Standardmäßig werden die Plots basierend auf den Rohresiduen für die qualifizierenden Effekte im Modell für feste Effekte, für zufällige Effekte und für wiederholte Effekte erstellt.
Abstand-Plot	erstellt einen Plot der Likelihood- oder beschränkten Likelihood-Distanz.
Einflussschätzer-Plots	erstellt ein Panel der gelöschten Schätzer in einer Influenzanalyse.
Einfluss-Statistiken-Panel	erstellt Panels der Einflussstatistiken. Bei iterativen Einfluss-Statistiken wird im Panel die Cook'sche *D*- und die CovRatio-Statistik für feste Effekte und Kovarianzparametern angezeigt. Dadurch wird das Darstellen der Auswirkung auf Schätzungen und Präzision für beide Schätzungstypen möglich. Bei nicht-iterativen Analysen werden nur Statistiken für feste Effekte dargestellt.
Gewöhnliches-Residuum-Plot	erstellt ein Panel der gewöhnlichen Residuen. Standardmäßig werden konditionale Residuen erzeugt.
Studentisiertes-Residuum-Plot	erstellt ein Panel der studentisierten Residuen. Standardmäßig werden konditionale Residuen erzeugt.
Pearson-Residuum-Plot	erstellt ein Panel des Pearsonschen Residuums. Standardmäßig werden konditionale Residuen erzeugt.

Optionsname	Beschreibung
PRESS-Residuum-Plot	erstellt einen Plot der PRESS-Residuen oder PRESS-Statistiken. Basierend auf einer "Leave-One-Out-" oder "Leave-Set-Out"-Prognose des marginalen Mittels.
Skaliertes-Residuum-Plot	erstellt ein Panel, das aus Residuengrafiken besteht, basierend auf den skalierten Residuen, den Response- und marginalen Residuen, die durch die umgekehrte Cholesky-Wurzel der marginalen Varianz-Kovarianzmatrix skaliert wurden.

Gemischte Modelle: Festlegen von Vorhersageoptionen

Klicken Sie im Auswahlfenster auf **Vorhersagen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Daten für Vorhersage	identifiziert die Datenquellen, die Sie zur Vorhersage der Ausgabedaten verwenden möchten. Wenn Sie das Kontrollkästchen **Weitere Daten** aktivieren, können Sie den Speicherort für Ihre zusätzlichen Daten bestimmen, indem Sie auf **Durchsuchen** klicken.
Ausgabedaten speichern	gibt an, ob die Vorhersagen und vorhergesagten Mittelwerte in einer Ausgabedatei gespeichert werden sollen: **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Vorhersagen anzeigen	zeigt die Vorhersagen in den Ergebnissen an.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

53

Modell-Scoring

Allgemeines über den Modell-Scoring-Assistenten 409
Modell-Scoring: Auswählen der Daten 410
Modell-Scoring: Auswählen eines Scoring-Modells 410
Modell-Scoring: Zuweisen von Datenvariablen zu erforderlichen Modelleingabewerten 411
Modell-Scoring: Auswählen der Ausgabedatenspalten 411
Modell-Scoring: Speichern der Ausgabedaten 412
Modell-Scoring: Bestätigen Ihrer Auswahl 412
Modell-Scoring: Modell-Scoring: Interpretieren der Ergebnisse 413

Allgemeines über den Modell-Scoring-Assistenten

Der **Model Scoring**-Assistent ermöglicht es Ihnen, Scoring-Code von einem vorhandenen SAS Enterprise Miner-Modell auf eine Datei anzuwenden. Das Erstellen von Rangfolgen bzw. Scoring von Dateien ist eines der Hauptziele im Data Mining. Unter Verwendung eines Vorhersagemodells können Sie Scoring mit Dateien durchführen, in denen beispielsweise Listen von Darlehensantragstellern oder Teilnehmer von Direktmarketing-Aktionen usw. enthalten sind. Sie können mit dieser Anwendungsroutine auch eine oder mehrere Segmentvariablen für die Eingabedatei erstellen.

Bevor Sie vom **Modell-Scoring**-Assistenten auf ein Modell zugreifen können, muss zuerst das Modell, das den Scoring-Code erzeugt, erstellt und in SAS Enterprise Miner validiert werden. Der vom Modell erzeugte Scoring-Code muss auch im SAS Metadata Repository gespeichert werden.

Sie können diesen Assistenten verwenden, um Datentabellen mit Kundenattributen auszuwerten. Basierend auf dieser Analyse können Sie eine Mailing-Liste von Kunden erstellen, die höchstwahrscheinlich einen Kauf tätigen werden. Sie könnten

den Score-Code eines anderen Modells verwenden, um basierend auf einem oder mehreren Attribut-Scorewerten jeden Kunden einem Marktsegment zuzuweisen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	Ohne
Erforderliche SAS-Produkte	SAS Enterprise Miner
Empfohlene zusätzliche SAS-Produkte	Ohne

Modell-Scoring: Auswählen der Daten

Der **Modell-Scoring**-Assistent verwendet standardmäßig die aktiven Daten im Projekt.

Klicken Sie auf **Weiter**, um ein Scoring-Modell auszuwählen.

Modell-Scoring: Auswählen eines Scoring-Modells

1 Bevor Sie die ein Scoring-Modell auswählen können, müssen Sie die Daten auswählen.

2 Klicken Sie auf **Durchsuchen**, um ein Scoring-Modell auszuwählen. Das Fenster **Model Scoring** wird geöffnet. Wählen Sie auf dem Reiter **Durchsuchen** das Modell aus, das Sie verwenden möchten.

   Klicken Sie nach Auswahl des Modells auf **OK**. Der Verzeichnispfad für das Modell wird im Feld **Scoring-Modell** und die Eigenschaften werden im Feld **Modellattribute** angezeigt. Wenn Sie den vom Modell generierten SAS-Code anzeigen möchten, klicken Sie auf **Quelle**.

3 Klicken Sie auf **Weiter**, um die Datenvariablen den erforderlichen Eingabewerten des Modells zuzuweisen.

Modell-Scoring: Zuweisen von Datenvariablen zu erforderlichen Modelleingabewerten

1 Bevor Sie die Datenvariablen den erforderlichen Modelleingabewerten zuweisen können, müssen Sie ein Scoring-Modell auswählen.

2 Wählen Sie für jeden erforderlichen Modelleingabewert die Datenvariable aus den Quelldaten aus. Standardmäßig werden Modelleingabewerte und Datenvariablen mit dem gleichen Namen und vom gleichen Typ einander zugeordnet. Um die einem Modelleingabewert zugeordnete Datenvariable zu ändern, markieren Sie die Variable und klicken Sie auf **Bearbeiten**. Wählen Sie aus der angezeigten Dropdown-Liste die gewünschte Datenvariable aus.

3 Klicken Sie auf **Weiter**, um die Ausgabespalten anzugeben.

Hinweis: Die Schaltfläche **Weiter** ist erst verfügbar, wenn die Datenvariablen allen erforderlichen Eingabewerten zugewiesen wurden.

Modell-Scoring: Auswählen der Ausgabedatenspalten

1 Bevor Sie die Ausgabedatenspalten auswählen können, müssen Sie die Datenvariablen den erforderlichen Eingabewerten zuweisen.

2 Wählen Sie auf den Reitern **Datenvariablen** und **Modell-Ausgabe** die Elemente aus, die Sie in die Ausgabe einschließen möchten. Auf dem Reiter **Datenvariablen** sind alle Variablen der Eingabedatei enthalten. Auf dem Reiter **Modell-Ausgabe** sind die Variablen **Date**, **Datetime** und **ModelKey** enthalten. Wählen Sie auf jedem Reiter die Elemente aus, die Sie in die Ausgabe einschließen möchten, und ziehen Sie diese in den Bereich **Ausgewählte Spalten**. Die im Bereich **Ausgewählte Spalten** aufgeführten Variablen und Modell-Ausgaben werden in der Ausgabe angezeigt. Legen Sie mit dem Nach-oben- bzw. Nach-unten-Pfeil die Reihenfolge der Variablen und der Ausgaben fest.

3 Klicken Sie auf **Weiter**, um die Ausgabedaten zu speichern.

Modell-Scoring: Speichern der Ausgabedaten

Bevor Sie die Ausgabedaten speichern können, müssen Sie die Ausgabespalten auswählen.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Klicken Sie auf **Weiter**, um Ihre Auswahl zu bestätigen.

Die Ausgabedatei enthält die ursprünglichen Daten und die Scorewerte, wie z.B. vorhergesagte Werte, Residuen und Klassifizierungsergebnisse. Weitere Informationen zu den Variablen in der Ausgabedatei finden Sie unter Interpretieren der Ergebnisse.

Modell-Scoring: Bestätigen Ihrer Auswahl

Bevor Sie Ihre Auswahl bestätigen können, müssen Sie die Ausgabedatenspalten auswählen.

Im letzten Schritt wird Ihre Auswahl zusammengefasst. Wenn Sie eine Auswahl ändern möchten, klicken Sie auf **Zurück**. Klicken Sie auf **Fertig stellen**, um die Anwendungsroutine auszuführen.

Modell-Scoring: Modell-Scoring: Interpretieren der Ergebnisse

Die Ausgabedatei enthält die ursprünglichen Daten und die Scorewerte, wie z.B. vorhergesagte Werte, Residuen und Klassifizierungsergebnisse. Die Namen der Scorewert-Variablen werden durch Hinzufügen eines Präfixes zum Namen der Zielvariablen, der Eingabevariablen oder der Eingabedatei gebildet.

Die folgende Tabelle listet jedes dieser Präfixe auf, enthält eine Kurzbeschreibung des Präfix, gibt an, ob eine Zielvariable erforderlich ist und ob diese Variable in der Entscheidungsmatrix enthalten ist. Weitere Informationen zur Interpretation der Ergebnisse finden Sie in der Hilfe zu SAS Enterprise Miner.

Präfix	Beschreibung	Ziel erforderlich?	Entscheidungsmatrix?
AOV_	Identifiziert die Anzahl der Klassen der intervallskalierten Variablen, die in 16 gleich große Gruppen unterteilt ist	Nein	Nein
B_	Klassenvariable	Ja	Nein
BL_	Bestmöglicher Verlust aller Entscheidungen, – *B(i)*	Ja	Ja
BP_	Bestmöglicher Gewinn aller Entscheidungen, *B(i)*	Ja	Ja
CL_	Verlust wird aus dem Zielwert errechnet, – *C(i)*	Ja	Ja
CP_	Gewinn wird aus dem Zielwert errechnet, *C(i)*	Ja	Ja
D_	Etikett der durch das Modell gewählten Entscheidung	Nein	Ja

Präfix	Beschreibung	Ziel erforderlich?	Entscheidungsmatrix?
EL_	Erwarteter Verlust für die vom Modell gewählte Entscheidung, *–E(i)*	Nein	Ja
EM_CCF	Durchschnittlicher Kreditkostenfaktor	Nein	Nein
EM_CLASSIFICATION	Normalisierte Kategorie, in die der Fall klassifiziert ist	Nein	Ja
EM_DECISION	Etikett der durch das Modell gewählten Entscheidung	Nein	Ja
EM_EVENTPROBABILITY	Posteriori-Wahrscheinlichkeit des Zielereignisses	Nein	Nein
EM_EXPOSURE	Durchschnittlicher Wert (Exposure)	Nein	Nein
EM_LGD	Durchschnittlicher Verlust aufgrund des Standardwerts	Nein	Nein
EM_PD	Durchschnittlicher vorhergesagter Wert	Nein	Nein
EM_PREDICTION	Vorhergesagter Wert	Nein	Nein
EM_PROBABILITY	Maximale Posteriori-Wahrscheinlichkeit, die der Variablen EM_CLASSIFICATION zugewiesenen Ausprägung entspricht	Nein	Nein
EM_PROFITLOSS	Erwarteter Gewinn oder Verlust	Nein	Ja
EM_SEGMENT	Segmentvariable	Nein	Nein
EP_	Erwarteter Gewinn für die vom Modell	Nein	Ja

Präfix	Beschreibung	Ziel erforderlich?	Entscheidungsmatrix?
	gewählte Entscheidung, *E(i)*		
F_	Normalisierte Kategorie, aus der der Fall stammt	Ja	Nein
G_	Identifikationsnummer der Gruppe, zu der die Beobachtung gehört	Nein	Nein
GRP_	Identifikationsnummer der Gruppe, zu der die Beobachtung gehört	Nein	Nein
I_	Normalisierte Kategorie, in die der Fall klassifiziert ist	Nein	Nein
IC_	Investitionskosten, *IC(i)*	Nein	Ja
IMP_	Variable mit Ersetzung fehlender Werte	Nein	Nein
LBL	Textetikett für die in Tabellen und Plots zu verwendende Variable	Nein	Nein
M_	Indikatorvariable, die fehlende, berücksichtigte Beobachtungen identifiziert, oder Indikatorvariable, die gefilterte Beobachtungen identifiziert	—	Nein
_N	Nachbarvariable; Listen der Nachbarbeobachtungen	Ja	Nein
NODE	Numerische ID für jedes Blatt einer Ausgabedatei in einem Entscheidungsbaum	Ja	Nein

Präfix	Beschreibung	Ziel erforderlich?	Entscheidungsmatrix?
P_	Vorhergesate Werte und Posteriori-Wahrscheinlichkeiten	Nein	Nein
R_	Einfache Residuen; Ziel minus Vorhersage	Ja	Nein
REPL_	Ersetzte Variable	Ja	Nein
ROI_	Return on investment, *ROI(i)*	Ja	Ja
SCORECARD_BIN	Jeder Beobachtung zugewiesenes Intervall	Nein	Nein
SCORECARD_POINTS	Gesamte Scorecard-Punkte für jedes Individuum	Nein	Nein
SEGMENT	Segmentvariable	Nein	Nein
SOM_DIMENSION1	Identifiziert Zeilen in einer Self-organizing Map (SOM)	Nein	Nein
SOM_DIMENSION2	Identifiziert Zeilen in einer SOM	Nein	Nein
SOM_ID	SOM-Variable	Nein	Nein
SOM_SEGMENT	Identifiziert mit einer SOM generierte Cluster	Nein	Nein
U_	Nicht formatierte Kategorie, in die der Fall klassifiziert ist	Nein	Nein
WOE_	Identifikationsnummer der Gruppe, zu der die Beobachtung gehört	Nein	Nein

Gescorte Ausgabedateien enthalten ebenfalls eine Variable mit dem Namen _WARN_, die auf Probleme bei der Berechnung vorhergesagter Werte oder bei der Entscheidungsfindung hinweist. _WARN_ ist eine alphanumerische Variable, die entweder leer ist (ein Hinweis darauf, dass keine Probleme aufgetreten sind) oder mindestens einen der folgenden alphanumerischen Codes enthält:

Code	Bedeutung
C	Fehlende Kostenvariable
M	Fehlende Input-Variablen
P	Ungültige Posteriori-Wahrscheinlichkeit (z. B. <0 oder >1)
U	Nicht erkannte Input-Kategorie

54

Nichtlineare Regression

Allgemeines über die Anwendungsroutine 'Nichtlineare Regression' ... **419**

Nichtlineare Regression: Zuweisen von Variablen zu Analyserollen ... **420**

Nichtlineare Regression: Festlegen von Modelloptionen ... **421**

Nichtlineare Regression: Berechnungsoptionen festlegen ... **421**
Iterationsoptionen festlegen ... 421
Angeben der Methode für Schrittweitensuche ... 422

Nichtlineare Regression: Erzeugen von Plots ... **423**

Nichtlineare Regression: Festlegen von Ergebnisoptionen ... **424**

Festlegen von Titeln und Fußnoten ... **425**

Anzeigen von Eigenschaften ... **425**

Allgemeines über die Anwendungsroutine 'Nichtlineare Regression'

Die Anwendungsroutine **Nichtlineare Regression** erzeugt Kleinste-Quadrate- oder gewichtete Kleinste-Quadrate-Schätzer der Parameter eines nichtlinearen Modells. Nichtlineare Modelle sind schwieriger anzugeben und zu schätzen als lineare Modelle. Statt lediglich Regressorvariablen aufzulisten, müssen Sie den Regressionsausdruck auswählen, Parameternamen deklarieren und anfängliche Parameterwerte angeben. Einige Modelle sind schwierig anzupassen, und es ist nicht gewährleistet, dass die Prozedur das Modell erfolgreich anpassen kann.

Sie können diese Anwendungsroutine verwenden, wenn Sie ein nichtlineares Modell für Ihr Experiment erstellen möchten. Sie eignet sich beispielsweise zum Untersuchen der Beziehung zwischen Konzentration und Geschwindigkeit eines bestimmten Enzym-Substrat-Paares. Sie zeichnen die Reaktionsrate (Geschwindigkeit) auf, die bei verschiedenen Substratkonzentrationen beobachtet

wurde, und speichern diese Werte in einer SAS-Datei. Die Anwendungsroutine Nichtlineare Regression ermöglicht Ihnen die Berechnung der Schätzer jedes Parameters, der zugehörigen asymptotischen Standardfehler und des asymptotischen Konfidenzintervalls.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	NLIN
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	Ohne

Nichtlineare Regression: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Abhängige Variable	gibt die Variable an, die als abhängige (Response) Variable verwendet wird. Sie müssen dieser Rolle genau eine Variable zuweisen.
Erklärende Variable	gibt die Variablen an, die als unabhängige (erklärende) Variablen verwendet werden. Sie müssen dieser Rolle genau eine Variable zuweisen.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer anderen Rolle zugewiesen ist.

Nichtlineare Regression: Festlegen von Modelloptionen

Klicken Sie im Auswahlfenster auf **Modell**, um auf diese Optionen zuzugreifen.

Wählen Sie einen nichtlinearen Modellausdruck aus der Dropdown-Liste aus. Dieser Ausdruck wird verwendet, um vorhergesagte Werte zu bewerten.

Geben Sie im Feld **Parameter für ausgewähltes Modell** die Parametertypen a, b und c entweder als Einzelwert oder als Wertebereich an. Wenn Sie Einzelwert als Parametertyp angeben, wird dem Parameter der Wert im Feld **Einzelwert** zugeordnet. Wenn Sie einen Wertebereich als Parametertyp angeben, können Sie auch den ersten und letzten Wert sowie Wertzunahmen angeben.

Um Beschränkungen oder Grenzen für die Parameter zu definieren, klicken Sie auf **Parametergrenzen** und geben im Fenster **Parametergrenzen** die Grenzen an.

Hinweis: Im Fenster **Parametergrenzen** wird nur die Syntax der eingegebenen Parametergrenzen validiert. Ob die Grenzen mathematisch sinnvoll sind, wird nicht überprüft. Beispielsweise können Sie die Grenze 2 < b < 1 eingeben, auch wenn dadurch das Modell nicht konvergiert.

Nichtlineare Regression: Berechnungsoptionen festlegen

Iterationsoptionen festlegen

Klicken Sie im Auswahlfenster auf **Berechnung**, um auf diese Optionen zuzugreifen.

Plot-Name	Beschreibung
Methode	Wählen Sie eine Methode aus der Dropdown-Liste **Methode** aus. Die Optionen **Gauss-Newton**, **Marquardt**, **Newton** und **Steilster Abstieg (Gradient)** sind verfügbar.

Plot-Name	Beschreibung
Maximale Iterationen	Geben Sie die Anzahl der Iterationen an, die die Anwendungsroutine durchführt, bevor sie aufhört, nach Konvergenzen zu streben. Dieser Wert muss eine positive Ganzzahl sein. Standardmäßig beträgt maximale Anzahl der Iterationen `100`.
Bei Parameterschätzung eine Moore-Penrose-Inverse verwenden	Wählen Sie dieses Kontrollkästchen aus, um die Moore-Penrose-Inverse bei der Parameterschätzung zu verwenden.
Mittleren Quadratfehler ersetzen	Wählen Sie das Kontrollkästchen aus, und geben Sie anschließend einen Wert ein, um den mittleren Quadratfehler zu ersetzen und die Standardfehler der Schätzer im Textfeld zu berechnen.
Minimierungstuning	Geben Sie ein Konvergenz- oder Singularitätskriterium für das Modell in den Textfeldern an. Für alle iterativen Methoden wird standardmäßig das relative Offset-Konvergenzmaß von Bates und Watts verwendet, um die Konvergenz zu bestimmen. Der Standard ist `1.0E-5`. Das Singularitätskriterium entspricht dem absoluten Betrag des kleinsten Pivot-Werts, der beim Umkehren des Hessian-Werts oder bei der Hessian-Approximation zulässig ist. Der Standard ist `1.0E-8`.

Angeben der Methode für Schrittweitensuche

Klicken Sie im Auswahlfenster auf **Berechnung**, um auf diese Optionen zuzugreifen.

Plot-Name	Beschreibung
Methode für Schrittweitensuche	Wählen Sie eine Suchmethode. Die folgenden Optionen sind verfügbar: ■ **Keine Schrittweitensuche** hebt die Einschränkung auf, dass der Zielwert bei jeder Iteration kleiner werden muss. Schritthalbierungen werden weiterhin verwendet, um die Parameterschätzer innerhalb angegebener Grenzen zu halten und sicherzustellen, dass die Anzahl der auswertbaren Beobachtungen nicht weniger wird. Diese Option eignet sich hervorragend für iterativ neu gewichtete Probleme kleiner Quadrate. Hierbei handelt es sich um die Standardeinstellung.

Plot-Name	Beschreibung
	■ **Schritthalbierungen** sucht nach der kleinsten Quadratfehlersumme (SSE). ■ **Kubische Interpolation** schätzt die Schrittgröße mittels einer kubischen Interpolation. Wenn die geschätzte Schrittgröße nicht zu einer Verkleinerung der Quadratfehlersumme (SSE) führt, wird die Schritthalbierung verwendet. ■ **Goldener Schnitt** ermittelt die Schrittgröße anhand des goldenen Schnitts. Der Parameter Tau bestimmt die Länge des zu durchsuchenden Anfangsintervalls. Der Parameter Rho gibt an, wie fein die Suche ausgeführt werden muss.
Maximale Anzahl der Schritthalbierungen	Geben Sie die maximale Anzahl der durchzuführenden Schritthalbierungen an. Dieser Wert muss eine positive Ganzzahl sein.

Nichtlineare Regression: Erzeugen von Plots

Diese Optionen werden im Auswahlfenster unter der Überschrift **Plots** angezeigt. Wählen Sie den Plottyp, den Sie generieren möchten. Sie können alle Plottypen erstellen.

Hinweis: Im Feld **Zusammenfassung angeforderter Plots** werden alle angeforderten Plots aufgeführt.

Plot-Name	Beschreibung
Vorhergesagt	Sie können Streuungsdiagramme beobachteter Werte gegen vorhergesagte Werte für jede abhängige Variable sowie Streuungsdiagramme beobachteter Werte jeder abhängigen Variablen gegen jede erklärende (unabhängige) Variable generieren. Bei Plots Beobachtet vs. unabhängig können Sie außerdem Konfidenzgrenzen oder Vorhersagegrenzen für alle Vorhersageplots bestimmen. ■ Wenn Sie Konfidenzgrenzen angeben, zeigt der Plot die Konfidenzintervalle für den Mittelwert.

Plot-Name	Beschreibung
	■ Wenn Sie Vorhersagegrenzen angeben, zeigt der Plot die Vorhersageintervalle für die einzelnen Responses. Bei beiden Plots ist alpha ,05.
Residuum	Sie können gewöhnliche, standardisierte und studentisierte Residuen jeder abhängigen Variablen gegenüber vorhergesagten Werten jeder abhängigen Variablen oder Werten jeder erklärenden (unabhängigen) Variablen darstellen. Die Residuen werden mithilfe folgender Formeln berechnet: ■ Ein gewöhnliches Residuum ist die Differenz zwischen dem beobachteten und dem vorhergesagten Wert. ■ Ein standardisiertes Residuum ist der Wert des Residuums dividiert durch den Standardfehler. ■ Ein studentisiertes Residuum ist der Wert des Residuums dividiert durch den Standardfehler bei gelöschter aktueller Beobachtung.
Einfluss	Sie können den Leverage-Wert gegenüber den vorhergesagten Werten jeder abhängigen Variablen oder Werten jeder erklärenden (unabhängigen) Variablen darstellen. Der Leverage-Wert für eine Beobachtung ist der Wert des entsprechenden diagonalen Elements in der *H*-Matrix.

Nichtlineare Regression: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Sie können Statistiken, Parameterschätzer oder beides permanent speichern. Wenn Sie Parameterschätzer speichern, können Sie auch die endgültigen Schätzer speichern, wenn die iterative Grenze überschritten wird. Sie können die iterative Grenze in der Option Maximale Iterationen angeben.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Darüber hinaus können Sie den Iterationsverlauf, Hougaards Maß der Schiefe oder beides in den Ergebnissen speichern. Mit Hougaards Maß der Schiefe können Sie beurteilen, wie nahe ein Parameter an der Linearität ist oder ob er ein beträchtliches Maß an Nichtlinearität aufweist.

Hinweis: Diese Optionen sind nicht verfügbar, wenn Sie das Kontrollkästchen **Gesamte angezeigte Ausgaben unterdrücken** aktivieren.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

   Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine

- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

55

Nichtparametrische einfache ANOVA

Allgemeines über die Anwendungsroutine 'Nichtparametrische einfache ANOVA' 427
Nichtparametrische einfache ANOVA: Zuweisen von Variablen zu Analyserollen 428
Nichtparametrische einfache ANOVA: Auswählen der Analyse 429
Nichtparametrische einfache ANOVA: Berechnen exakter p-Werte 430
Nichtparametrische einfache ANOVA: Festlegen von Ergebnisoptionen 430
Festlegen von Titeln und Fußnoten 431
Anzeigen von Eigenschaften 432

Allgemeines über die Anwendungsroutine 'Nichtparametrische einfache ANOVA'

Die Anwendungsroutine **Nichtparametrische einfache ANOVA** besteht aus nichtparametrischen Tests auf Lageparameter und Skalierungsunterschiede über eine einfache Klassifizierung. Die Anwendungsroutine bietet auch eine Standard-Varianzanalyse auf den Rohdaten und Statistiken, basierend auf der empirischen Verteilungsfunktion.

Sie können diese Anwendungsroutine zum Testen der Nullhypothese gegen eine alternative Hypothese verwenden. Angenommen, 59 weibliche Patienten mit rheumatischer Arthritis haben an einem klinischen Versuch teilgenommen und wurden in zwei Gruppen aufgeteilt: Aktiv und Placebo. Der Reaktionsstatus (hervorragend=5, gut=4, mittelmäßig=3, relativ schlecht=2, schlecht=1) jedes Patienten wird aufgezeichnet. Mithilfe der Anwendungsroutine **Nichtparametrische einfache ANOVA** können Sie die Nullhypothese, dass es keinen Unterschied im

Patientenreaktionsstatus zwischen den beiden Behandlungsgruppen gibt, gegen eine alternative Hypothese testen, dass sich die beiden Patientenreaktionsstatus zwischen den Behandlungsgruppen unterscheiden.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	NPAR1WAY
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Nichtparametrische einfache ANOVA: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Abhängige Variablen	gibt die Variablen an, die als abhängige (Response) Variablen verwendet werden. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Unabhängige Variable	gibt die Variable an, die als unabhängige (erklärende) Variable verwendet wird. Sie müssen dieser Rolle genau eine Variable zuweisen. Für die Untergruppen werden separate Analysen durchgeführt, die von der Variablen in dieser Liste definiert werden.

Optionsname	Beschreibung
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits als abhängige Variable oder als unabhängige Variable ausgewählt ist.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.

Nichtparametrische einfache ANOVA: Auswählen der Analyse

Klicken Sie im Auswahlfenster auf **Analyse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Test-Scorewerte	Wählen Sie die durchzuführenden Analysen aus.
Empirische Verteilungsfunktion	Aktivieren Sie das Kontrollkästchen **Empirische Verteilungsfunktionsstatistiken berechnen (EDF)**, wenn Sie die Empirische Verteilungsfunktion durchführen möchten.
Fehlende Werte	Sie können festlegen, ob fehlende Werte zulässige Ausprägungen für Klassifizierungsvariablen bilden, indem Sie das entsprechende Kontrollkästchen markieren.
Kontinuitätskorrektur	Die Anwendungsroutine umfasst bei der Berechnung der standardisierten Teststatistiken für Wilcoxon- und Siegel-Tukey-Scorewerte eine Kontinuitätskorrektur. Wenn Sie für die Tests keine Kontinuitätskorrektur verwenden möchten,

Optionsname	Beschreibung
	wählen Sie das Kontrollkästchen **Kontinuitätskorrektur unterdrücken** aus.

Nichtparametrische einfache ANOVA: Berechnen exakter p-Werte

Klicken Sie im Auswahlfenster auf **Exakte p-Werte**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Exakte p-Werte	Standardmäßig berechnet die Anwendungsroutine asymptotische statt exakter *p*-Werte. Sie können über die entsprechenden Kontrollkästchen festlegen, dass exakte *p*-Werte für die verschiedenen Analysen berechnet werden sollen. **Hinweis:** Die Test-Scorewerte wurden markiert, um die Kontrollkästchen im Bereich **Exakte p-Werte** zu aktivieren. Um die Test-Scorewerte zu markieren, klicken Sie im Auswahlfenster auf **Analyse**. Klicken Sie auf **Zurücksetzen**, um die Berechnungen exakter p-Werte abzubrechen.
Berechnungszeitgrenze	Geben Sie im Bereich **Berechnungszeitgrenze** die Zeitgrenze zur Berechnung des exakten *p*-Werts an, oder geben Sie die Zeitgrenze als **Unbestimmt** an. Das Berechnen exakter *p*-Werte kann viel Zeit und Arbeitsspeicher in Anspruch nehmen.

Nichtparametrische einfache ANOVA: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Speichern von Ausgabedaten	speichert Ihre Ergebnisse teilweise oder vollständig in einer Ausgabetabelle. Aktivieren Sie die Kontrollkästchen der Statistiken, die in Ihrer Ausgabe enthalten sein sollen. **Hinweis:** Die Test-Scorewerte wurden markiert, um die Kontrollkästchen im Bereich **Einzuschließende Statistiken** zu aktivieren. Um die Test-Scorewerte zu markieren, klicken Sie im Auswahlfenster auf **Analyse**. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Berichte unterdrücken	Bewirkt, dass keine Ergebnisse angezeigt werden.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

56

np-Regelkarte

Allgemeines über die Anwendungsroutine 'np-Regelkarte' ... 433
np-Regelkarte: Zuweisen von Variablen zu Analyserollen ... 434
Regelkarten: Festlegen von Regelgrenzen ... 436
Regelkarte: Auswählen der durchzuführenden Tests ... 437
Regelkarten: Festlegen von Optionen für die Plotdarstellung ... 438
Regelkarten: Speichern von Analyseergebnissen ... 439
Festlegen von Titeln und Fußnoten ... 440
Anzeigen von Eigenschaften ... 440

Allgemeines über die Anwendungsroutine 'np-Regelkarte'

Die Shewhart-Regelkarte ist ein Grafik- und Analysetool, mit dem bestimmt werden kann, ob ein Prozess innerhalb der statistischen Kontrolle liegt. Die Anwendungsroutine **np-Regelkarte** erstellt np-Regelkarten für die Anzahl der Nichtkonformitäten (Fehler) in den Untergruppenstichproben.

Sie können diese Anwendungsroutine zum Überwachen der Anzahl der Defekte in einem Fertigungsprozess verwenden. Angenommen, ein Elektronikunternehmen stellt Schaltungen in Chargen zu je 500 Stück her und verwendet eine np-Regelkarte zur Überwachung der Anzahl der fehlerhaften Schaltungen. Es werden dreißig Chargen untersucht, wobei die Fehler in den einzelnen Chargen gezählt werden. Die Fehlerzählungen werden in einer SAS-Datei gespeichert, die zum Erstellen der np-Regelkarte verwendet wird. Jeder Punkt auf der np-Regelkarte stellt die Anzahl der nichtkonformen Elemente in einer bestimmten Untergruppe dar. Beispielswiese wird für die erste Charge der Wert 5 dargestellt. Wenn alle Punkte innerhalb der Regelgrenzen liegen (standardmäßig ist dies 3 Sigma), kann daraus geschlossen werden, dass der Prozess innerhalb der statistischen Kontrolle liegt.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SHEWHART
Erforderliche SAS-Produkte	Base SAS, SAS/QC, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

np-Regelkarte: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Anzahl der Nichtkonformitäten	gibt die Variablen an, welche die Anzahl der Nichtkonformitäten enthalten. Sie müssen dieser Rolle mindestens eine Variable zuweisen
Untergruppen-Identifier	gibt die Variable an, welche die Untergruppen in den Daten identifiziert. Die Werte der Variablen geben an, wie die Beobachtungen in der Eingabetabelle in rationale Untergruppen gegliedert werden. Dies sind die üblichen Werte: ■ die Indizes, welche die Reihenfolge angeben, in der die Untergruppenstichproben gesammelt wurden ■ die Daten oder Zeiten, zu denen die Untergruppenstichproben gesammelt wurden ■ die Etiketten, welche die Untergruppenstichproben eindeutig identifizieren.

Optionsname	Beschreibung
	Ein Untergruppen-Identifier ist erforderlich. Sie können dieser Rolle nur eine Variable zuweisen. Um die Daten nach den identifizierten Untergruppen zu sortieren, aktivieren Sie das Kontrollkästchen **Nach Untergruppe sortieren**.
Stichprobengröße der Untergruppe	gibt die Stichprobengröße der Untergruppe als Werte einer Variablen an. Die Stichprobengröße der Untergruppe ist erforderlich, und Sie können dieser Rolle nur eine numerische Variable zuweisen. Aktivieren Sie das Kontrollkästchen **Numerischen Wert für Untergruppe verwenden (statt Variable)**, wenn die Stichprobengröße der Untergruppe fest ist und Sie einen konstanten Wert für die Größe angeben müssen. Geben Sie den gewünschten Wert im Feld **Stichprobengröße der Untergruppe** ein. Der Wert für die Stichprobengröße der Untergruppe muss zwischen dem maximalen Wert der Prozessvariablen und 99999 (einschließlich) liegen. **Hinweis:** Diese Rolle ist nicht erforderlich, wenn Sie das Kontrollkästchen **Numerischen Wert für Untergruppe verwenden (statt Variable)** im Fenster **Optionen** aktiviert haben.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.
Blockvariablen	gibt bis zu zwei Variablen an, welche die Daten in Blöcke aufeinanderfolgender Untergruppen gruppieren. Diese Blöcke werden in einer Legende beschriftet. Jede Blockvariable liefert eine Etikettenausprägung in der Legende.

Regelkarten: Festlegen von Regelgrenzen

Klicken Sie im Auswahlfenster auf **Regelgrenzen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Sigma-Grenzen	gibt die Breite der Regelgrenzen als Mehrfaches des Standardfehlers (Sigma) der beschreibenden Untergruppenstatistik an, die im Diagramm dargestellt wird. Die Breite muss positiv sein. Der Standardwert für das Mehrfache ist 3.
Berechnungsmethode auswählen	gibt die Berechnungsmethode für die Regelgrenzen an. Folgende Optionen können ausgewählt werden: ■ **Regelgrenzen aus den aktiven Daten berechnen** ■ **Regelgrenzen aus der ausgewählten Datei berechnen** Klicken Sie auf **Durchsuchen**, um die Datei anzugeben, die die Regelgrenzen bzw. die Parameter zu deren Berechnung enthält. Wählen Sie die gewünschte Datei aus und klicken Sie auf **Öffnen**. **Hinweis:** Bei der ausgewählten Datei muss es sich um eine LIMITS= Datei handeln. Weitere Informationen über die in einer LIMITS= Datei erforderlichen Variablen finden Sie in der Prozedurendokumentation zur erstellten Regelkarte. Wenn Sie die beschreibenden Statistiken für alle Untergruppen darstellen möchten, aktivieren Sie das Kontrollkästchen **Beschreibende Statistiken für alle Untergruppen darstellen, unabhängig davon, ob die Stichprobengröße der Untergruppe der nominalen Stichprobengröße für Regelgrenzen entspricht**. Sie können diese Statistiken beispielsweise darstellen, wenn fast alle Untergruppen eine gemeinsame Stichprobengröße haben und wenn Sie die festen (anstelle der variablen) Regelgrenzen, die der nominalen Stichprobengröße *n* entsprechen, anzeigen möchten. Wenn Sie diese Option auswählen und sich Ihre Stichprobengrößen für die Untergruppen deutlich unterscheiden, ist die Interpretation der Regelgrenzen nur für die Untergruppen sinnvoll, deren Stichprobengröße *n* entspricht.

Optionsname	Beschreibung
	Um spezielle Marker für die Datenpunkte hinzuzufügen, bei denen die Stichprobengröße ungleich `n` ist, aktivieren Sie das Kontrollkästchen **Spezielle Marker für Punkte hinzufügen, die Stichprobengrößen entsprechen, welche ungleich nominalen Stichprobengrößen für feste Regelgrenzen sind**. ■ **Regelgrenzen angeben** Geben Sie in den Feldern **Oben**, **Mitte** und **Unten** die obere Regelgrenze, die Mittellinie bzw. die untere Regelgrenze an. Wenn die Stichprobengrößen Ihrer Untergruppe variieren, müssen Sie eine nominale Stichprobengröße angeben. **Hinweis:** Wenn Sie der Rolle **Analyse gruppieren nach** Variablen zuweisen, können Sie keine Werte für Regelgrenzen angeben.

Regelkarte: Auswählen der durchzuführenden Tests

Klicken Sie im Auswahlfenster auf **Tests**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Tests wenden nicht die Regelgrenzen an, die keine Drei-Sigma-Grenzen sind oder die mit der Untergruppenstichprobengröße variieren.

Optionsname	Beschreibung
Test auswählen	Fordern Sie über die Kontrollkästchen einen oder mehrere Tests auf besondere Ursachen an, die auch als Ausführungstests, Mustertests und Western-Electric-Regeln bekannt sind. Diese Tests entdecken bestimmte nichtzufällige Muster in den Punkten, die in der primären Regelkarte dargestellt werden. Das Vorkommen eines nichtzufälligen Musters wird als Signal bezeichnet. Die Tests können auf das Vorhandensein einer besonderen Variationsursache hinweisen. Wenn Sie die Tests auswählen, erscheint eine Beschreibung des jeweiligen Tests im Feld **Beschreibung**.

Optionsname	Beschreibung
Etikett	Geben Sie in den Textfeldern die Etiketten für Punkte an, bei denen die Tests auf besondere Ursachen positiv sind. Die Etiketten dürfen maximal 16 Zeichen lang sein. Jedes Etikett erscheint bei allen Punkten im Test, an denen ein Signal vorliegt.
Text zum Identifizieren der Testsignale	■ Wählen Sie über die Dropdown-Liste eine Farbe für die angegebenen Etiketten aus. ■ Markieren Sie das Kontrollkästchen **Zonenlinien anzeigen**, um Linien einzuschließen, welche die Zonen A, B und C in einem primären Diagramm kennzeichnen. Diese Zonen werden gebildet, indem die Intervalle zwischen den Regelgrenzen in sechs gleichmäßig verteilte Unterintervalle unterteilt werden. ■ Verwenden Sie das Kontrollkästchen **3 Sigma-Grenzen überschreiben**, um nach speziellen Ursachen zu suchen, wenn Sie eine andere Sigma-Regelgrenze als das Standardvielfache von 3 angeben. Um die Sigma-Grenze zu ändern, klicken Sie im Auswahlfenster auf **Regelgrenzen**. ■ Über das Kontrollkästchen **Test auf überlappende Punktmuster anwenden** wenden Sie Tests auf besondere Ursachen auf überlappende Punktmuster an.

Regelkarten: Festlegen von Optionen für die Plotdarstellung

Einige der Darstellungsoptionen sind für unterschiedliche Regelkarten-Anwendungsroutinen gleich. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

■ “Festlegen von Achsenoptionen” auf Seite 759
■ “Hinzufügen von Referenzlinien” auf Seite 760
■ “Festlegen weiterer Plotoptionen” auf Seite 760
■ “Festlegen von Blockvariablenoptionen” auf Seite 762

Regelkarten: Speichern von Analyseergebnissen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

Um die Analyseergebnisse zu speichern, können Sie Ausgabedateien mit diversen Statistiken speichern. Markieren Sie die Kontrollkästchen der Ausgabedaten, die Sie permanent speichern möchten.

Optionsname	Beschreibung
Datei mit Untergruppenstatistiken und Regelgrenzen	erstellt eine Ausgabedatei, welche die im Diagramm dargestellten Informationen enthält. Dies beinhaltet die Werte der Untergruppenvariablen und ihre entsprechenden beschreibenden Statistiken und Regelgrenzen. **Hinweis:** Diese Option erstellt eine Datei OUTTABLE=. Sie können diese Datei mit der Option TABLE= aus der SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Untergruppenstatistiken	erstellt eine Ausgabedatei mit den beschreibenden Statistiken der Untergruppe. **Hinweis:** Diese Option erstellt eine Datei mit OUTHISTORY=. Sie können diese Datei mit der Option HISTORY= aus der SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Regelgrenzen	erstellt eine Ausgabedatei, welche die Regelgrenzen speichert. Sie können diese Datei verwenden, wenn Sie die Berechnungsmethode für die Regelgrenzen später in dieser Anwendungsroutine auswählen.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

57

Einfache ANOVA

Allgemeines über die Anwendungsroutine 'Einfache ANOVA' ... 443
Einfache ANOVA: Zuweisen von Variablen zu Analyserollen ... 444
Einfache ANOVA: Auswählen von Tests ... 445
Einfache ANOVA: Vergleichen von Gruppenmittelwerten ... 446
Einfache ANOVA: Auswählen von Statistiken für quantitative Variablen ... 446
Einfache ANOVA: Erzeugen von Plots ... 447
Einfache ANOVA: Festlegen von Ergebnisoptionen ... 448
Festlegen von Titeln und Fußnoten ... 449
Anzeigen von Eigenschaften ... 450

Allgemeines über die Anwendungsroutine 'Einfache ANOVA'

Eine einfache Varianzanalyse (ANOVA) betrachtet einen Behandlungsfaktor mit zwei oder mehreren Behandlungsausprägungen. Das Ziel der Analyse besteht darin, auf Unterschiede bei den Mittelwerten der Ausprägungen zu testen, und diese Unterschiede zu quantifizieren. Wenn es zwei Behandlungsausprägungen gibt, ist diese Analyse äquivalent zu einem *t*-Test, bei dem zwei Gruppenmittelwerte verglichen werden.

Sie können die ANOVA-Anwendungsroutine zu folgenden Zwecken einsetzen:

- Erstellen einer Studie zur Auswirkung von Bakterien auf den Stickstoffgehalt von Rotem Klee. Als Faktor wird die Bakterienbelastung verwendet. Dieser ist in sechs Stufen unterteilt.
- Analyse eines vollständigen Zufallsblockdesigns. Angenommen, Sie möchten wissen, ob drei verschiedene Behandlungstypen unterschiedliche Auswirkungen auf den Ertrag und den Wert einer bestimmten Ernte haben. Da Sie glauben, dass die experimentellen Einheiten nicht homogen sind, führen Sie einen

Blockbildungsfaktor ein, der für die Homogenität der experimentellen Einheiten innerhalb der verschiedenen Blöcke sorgt. Die drei verschiedenen Behandlungstypen werden anschließend zufällig innerhalb der einzelnen Blöcke zugewiesen

- Vergleichen Sie die Lebensspannen dreier verschiedener Batteriemarken. Als Faktor dient die Marke. Dieser ist in drei Stufen unterteilt.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	ANOVA
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Einfache ANOVA: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Abhängige Variablen	gibt die Variablen an, die als abhängige (Response) Variablen verwendet werden. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Unabhängige Variable	gibt die Variable an, die als unabhängige (erklärende) Variable verwendet wird. Sie müssen dieser Rolle genau eine Variable zuweisen.

Optionsname	Beschreibung
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits als abhängige Variable oder als unabhängige Variable ausgewählt ist.

Einfache ANOVA: Auswählen von Tests

Klicken Sie im Auswahlfenster auf **Tests**, um auf diese Optionen zuzugreifen.

Sie können einen einfachen Varianzanalysetest verwenden, wenn die Variable, die der Rolle Unabhängige Variable zugewiesen ist, zwei oder mehr Einzelwerte aufweist. Sie können Ihr Modell unter der Annahme gleicher Varianzen testen.

Optionsname	Beschreibung
Varianz-gewichtete ANOVA nach Welch	testet die Gruppenmittelwerte mittels gewichteter Varianz. Diesen Test können Sie verwenden, wenn die Annahme gleicher Varianzen abgewiesen wurde.
Tests auf gleiche Varianz	wählen Sie einen dieser Tests: ■ **Test nach Bartlett** ist eine Modifikation des Normaltheorie-Likelihood-Ratio-Tests. Dieser Test berechnet genaue Typ-I-Fehlerraten bei normaler Datenverteilung. ■ **Brown-Forsythe-Test** ist eine Variation des Tests nach Levene. Gleiche Varianzen werden mittels der absoluten Abweichungen der Gruppenmediane bestimmt. Dieser Test eignet sich für die Bestimmung von Varianzdifferenzen, ist jedoch ressourcenintensiv, wenn Ihre Daten mehrere unterschiedlich große Gruppen enthalten. ■ **Test nach Levene** wird als Standardhomogenität der Varianztests angesehen. Dieser Test berechnet die Quadratresiduen, um gleiche Varianzen zu bestimmen.

Einfache ANOVA: Vergleichen von Gruppenmittelwerten

Klicken Sie im Auswahlfenster unter der Überschrift 'Mittelwerte' auf **Vergleich**, um auf diese Optionen zuzugreifen.

Wählen Sie die multiplen Vergleichstests für den Haupteffekt (also die unabhängige Variable), die in der Ausgabedatei enthalten sein sollen. Wählen Sie mit Hilfe der Dropdown-Liste **Konfidenzniveau** ein Konfidenzniveau für alle ausgewählten Tests aus. Der Standardwert ist ein Konfidenzniveau von 95%.

Diese Tests sind verfügbar:

- *t*-Test nach Bonferroni
- Tukeys studentisierter Spannweitentest (HSD)
- Duncan Multiple-Range-Test
- *t*-Test nach Dunnett
- Fishers Test auf kleinste signifikante Differenzen
- Multiple Vergleichstests nach Gabriel
- Multipler Spannweitentest nach Student-Newman-Keuls
- k-Ratio-t-Test nach Waller-Duncan
- Multiple Vergleichstest nach Scheffe
- Multipler Spannweitentest nach Ryan-Einot-Gabriel-Welsch

Einfache ANOVA: Auswählen von Statistiken für quantitative Variablen

Klicken Sie im Auswahlfenster unter der Überschrift 'Mittelwerte' auf **Aufspaltung**, um auf diese Optionen zuzugreifen.

Wählen Sie die beschreibenden Statistiken, die für die abhängigen Variablen auf jedem Niveau der unabhängigen Variablen berechnet und angezeigt werden müssen.

Optionsname	Beschreibung
Mittelwert	das arithmetische Mittel, wird durch Addition der Werte einer Stichprobenvariablen und Division dieser Summe durch die Anzahl der Beobachtungen berechnet.
Standardabweichung	ein statistisches Maß der Variabilität einer Gruppe von Datenwerten. Dieses Maß ist das am häufigsten verwendete Maß der Streuung einer Häufigkeitsverteilung. Es ist gleich der positiven Quadratwurzel der Varianz.
Standardfehler	die Standardabweichung des Stichprobenmittelwerts. Der Standardfehler wird als Verhältnis der Stichprobenstandardabweichung zur Quadratwurzel der Stichprobengröße definiert.
Varianz	ein statistisches Maß für die Streuung der Datenwerte. Dieses Maß ist ein Mittelwert der gesamten quadratischen Streuung zwischen den einzelnen Beobachtungen und dem Stichprobenmittelwert.
Anzahl Beobachtungen ohne fehlende Werte	die Anzahl der Beobachtungen ohne fehlende Werte.
Anzahl fehlender Beobachtungen	die Anzahl der Beobachtungen, für die kein Wert eingegeben wurde.
Minimum	der niedrigste Wert für eine Beobachtung.
Maximum	der höchste Wert für eine Beobachtung.

Einfache ANOVA: Erzeugen von Plots

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Markieren Sie die entsprechenden Kontrollkästchen, um ein Box-Plot, einen Mittelwert-Plot oder beides zu erzeugen.

Plot-Name	Beschreibung
Box-and-Whisker	Ein Box-Plot zeigt den Mittelwert, Quartile sowie minimale und maximale Beobachtungen für eine Gruppe an.

Plot-Name	Beschreibung
	Der Plot repräsentiert folgende Statistiken: ■ Der Endpunkt des oberen Whiskerplots stellt den Maximalwert dar. ■ Der oberste Teil des Boxplots stellt das dritte Quartil dar (75tes Perzentil). ■ Die Länge des Boxplots stellt den Quartilsabstand (d. h. den Abstand zwischen dem dritten und ersten Quartil) dar. ■ Der Punkt im Boxplot stellt den Mittelwert dar. ■ Die horizontale Linie im Boxplot stellt den Median dar (50tes Perzentil). ■ Der unterste Teil des Boxplots stellt das erste Quartil dar (25tes Perzentil). ■ Der Endpunkt des unteren Whiskerplots stellt den Minimalwert dar.
Mittelwerte	Ein Mittelwertplot zeigt den Mittelwert und den Standardfehler für die unabhängige Variable an. Jeder Punkt auf der Grafik ist von einem Varianzbalken umgeben. Wenn Sie **Mittelwerte** wählen, können Sie auch diese Optionen auswählen: ■ Auswählen, ob zum Darstellen der Balken der Standardfehler des Mittelwerts oder die Standardabweichung verwendet werden soll. Der Standardfehler des Mittelwerts ist die Standardabweichung des Mittelwerts. Die Standardabweichung ist die Quadratwurzel der Stichprobenvarianz. ■ Auswählen, ob die Balkenlänge über die gepoolte Varianz berechnet werden soll. Die gepoolte Varianz ist die kombinierte Varianz zweier unabhängiger Stichproben, bei denen die Grundgesamtheiten gleich sind. ■ Wählen Sie eine Höhe für die Balken: 1, 2 oder 3 Standardeinheiten. ■ Wählen Sie aus, ob die vertikale Achse bei 0 beginnen soll.

Einfache ANOVA: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Statistikausgabedaten speichern	speichert das Ergebnisse in einer Ausgabedatei. Die Ausgabedatei enthält Quadratsummen, Freiheitsgrade, *F*-Statistiken und Wahrscheinlichkeitsniveaus für jeden Effekt in einem Modell. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Anzeige für gesamte Ergebnisausgabe unterdrücken	Zeigt die Ergebnisse nicht an.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

58

Einfache Häufigkeiten

Allgemeines über die Anwendungsroutine 'Einfache Häufigkeiten' **451**
Einfache Häufigkeiten: Zuweisen von Variablen zu Analyserollen **452**
Einfache Häufigkeiten: Festlegen von Statistikoptionen **453**
Einfache Häufigkeiten: Erzeugen von Plots **454**
Einfache Häufigkeiten: Festlegen von Ergebnisoptionen **454**
Festlegen von Titeln und Fußnoten **456**
Anzeigen von Eigenschaften **456**

Allgemeines über die Anwendungsroutine 'Einfache Häufigkeiten'

Mit der Anwendungsroutine **Einfache Häufigkeiten** können Sie beispielsweise Häufigkeitstabellen aus Ihren Daten generieren. Sie können diese auch zum Durchführen von binomialen und Chi-Quadrat-Tests verwenden.

Sie können diese Anwendungsroutine z. B. zum Analysieren der Wirksamkeit eines neuen Medikaments verwenden. Angenommen, eine medizinische Forschungsgruppe möchte die Wirksamkeit einer neuen Behandlung für ein Hautproblem beurteilen. Dermatologen der teilnehmenden Kliniken werden für die Durchführung der Studie und zum Bewerten der Bedingung geschult. Nach der Schulung untersuchen zwei Dermatologen die Patienten mit dem Hauptproblem in einer Pilotstudie und bewerten diese Patienten. Die Anwendungsroutine **Einfache Häufigkeiten** kann zum Bewerten der Diagnoseübereinstimmung eingesetzt werden.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	FREQ
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Einfache Häufigkeiten: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Analysevariablen	gibt die zu analysierenden Variablen an. Die Anwendungsroutine erstellt für jede Variable, die Sie dieser Rolle zuweisen, eine einfache Häufigkeitstabelle. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**.

Optionsname	Beschreibung
	Hinweis: Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer anderen Rolle zugewiesen ist.

Einfache Häufigkeiten: Festlegen von Statistikoptionen

Klicken Sie im Auswahlfenster auf **Statistiken**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Häufigkeitstabelle noptionen	Wählen Sie die Statistiken aus, die Sie in der einfachen Häufigkeitstabelle einschließen möchten. Diese Optionen sind verfügbar: ■ **Häufigkeiten und Prozentwerte mit Kumulativen** - erstellt eine Tabelle, die die Häufigkeiten, kumulativen Häufigkeiten, Prozentwerte der Gesamthäufigkeiten und kumulativen Prozentwerte für jeden Wert der Analysevariablen enthält. ■ **Häufigkeiten und kumulative Häufigkeiten** - erstellt eine Tabelle, die die Häufigkeiten und kumulativen Häufigkeiten für jeden Wert der Analysevariablen enthält. ■ **Häufigkeiten und Prozentwerte** - erstellt eine Tabelle, die die Häufigkeiten und Prozentwerte der Gesamthäufigkeiten für jeden Wert der Analysevariablen enthält. ■ **Nur Häufigkeiten** - erstellt eine Tabelle, die nur die Häufigkeiten für jeden Wert der Analysevariablen enthält.
Fehlende Werte	Aktivieren Sie das Kontrollkästchen **Häufigkeiten anzeigen**, um fehlende Werte in den Häufigkeitstabellen einzuschließen. Wählen Sie das Kontrollkästchen **Bei Berechnungen einschließen**, um die Häufigkeiten fehlender Werte in Binomial- oder Chi-Quadrat-Tests einzuschließen.
Binomialverhältnisse	Wählen Sie den asymptotischen Test, exakte *p*-Wertberechnung oder beides. Bei Binomialverhältnissen geben Sie ein Testverhältnis (Nullhypothesenverhältniswert) und Konfidenzniveau an.
Chi-Quadrat Güte der Anpassung	Wählen Sie den asymptotischen Test, exakte *p*-Wertberechnung oder beides.

Optionsname	Beschreibung
Exakte Berechnungen	Die Berechnung exakter *p*-Werte kann viel Arbeitsspeicher und Verarbeitungszeit in Anspruch nehmen. Aktivieren Sie im Bereich **Exakte Berechnungen** das Kontrollkästchen **Max. Berechnungszeit** und legen Sie ein Zeitlimit (in Sekunden) für die Berechnung jedes *p*-Werts für jede Kreuztabelle fest. Der Standardwert beträgt `900` Sekunden. Zur Berechnung der Monte-Carlo-Schätzer für die exakten p-Werte anstelle der direkten Berechnung der exakten *p*-Werte, aktivieren Sie das Kontrollkästchen **Monte-Carlo-Schätzung**. Die Monte-Carlo-Schätzung kann bei großen Problemen hilfreich sein, bei denen die exakte Berechnung zeit- und speicherintensiv ist, asymptotische Annäherungen jedoch eventuell nicht ausreichend sind.

Einfache Häufigkeiten: Erzeugen von Plots

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Markieren Sie die Kontrollkästchen, um Balkendiagramme der Häufigkeiten anzuzeigen. Sie können ein horizontales Balkendiagramm, ein vertikales Balkendiagramm oder beides anzeigen lassen.

Einfache Häufigkeiten: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Speichern von Ausgabedaten	Zum Speichern von Häufigkeitstabellen in einer Ausgabedatei müssen Sie das Kontrollkästchen **Datei mit Häufigkeiten und Prozentwerten erstellen** aktivieren. Um die kumulative Häufigkeit und den kumulativen Prozentwert für einfache Tabellen in die Ausgabedaten einzuschließen, aktivieren Sie das

Optionsname	Beschreibung
	Kontrollkästchen **Kumulative Häufigkeiten und kumulative Prozentwerte zu Datei hinzufügen**. Wählen Sie die in die Datei einzuschließenden Analysevariablen aus, und legen Sie anschließend den Speicherort fest, unter dem Sie die Datei speichern möchten. Der Pfad der Datei muss in der Form \ **\Server\Bibliothek\Member** angegeben werden. Sie können auch auf **Durchsuchen** klicken, um einen Speicherort auszuwählen. Wenn Sie ein Binomialverhältnis oder Chi-Quadrat-Test auswählten, können Sie die Häufigkeitstabellen permanent in eine Ausgabedatei speichern, indem Sie das Kontrollkästchen **Binomialverhältnis und Chi-Quadrat-Statistiken zu Datei hinzufügen** aktivieren. Geben Sie den Speicherort für die Datei an. Der Pfad der Datei muss in der Form **Server\Bibliothek \Member** angegeben werden. Sie können auch auf **Durchsuchen** klicken, um einen Speicherort auszuwählen. **Hinweis:** Sie können das Binomialverhältnis und die Chi-Quadrat-Statistiken nur für die letzte Variable in der Rolle **Analysevariablen** permanent speichern.
Anzeige für gesamte Ergebnisausgabe unterdrücken	Zeigt die Ergebnisse nicht an. **Hinweis:** Um die Anwendungsroutine ausführen zu können, müssen Sie entweder die Ausgabe anzeigen oder die Ergebnisse in einer Datei speichern. Wenn Sie das Kontrollkästchen neben **Datei mit Häufigkeiten und Prozentwerten erstellen** deaktiviert und das Kontrollkästchen neben **Alle angezeigten Ausgaben und Plots unterdrücken** aktiviert haben, können Sie die Anwendungsroutine ausführen oder speichern.
Reihenfolge für die Ausgabedaten auswählen	Um die Reihenfolge der Werte in der Tabelle zu bestimmen, wählen Sie die Reihenfolge für die Ausgabedaten aus der Dropdown-Liste **Ausgabedaten ordnen nach**. Diese Optionen sind verfügbar: ■ **Dateireihenfolge** - ordnet Werte entsprechend der Reihenfolge in der Eingabedatei an. ■ **Formatierte Werte** - ordnet Werte in aufsteigender Reihenfolge nach ihren formatierten Werten an. ■ **Absteigende Häufigkeiten** - ordnet Werte nach der absteigenden Häufigkeiteszählung an. ■ **Unformatierte Werte** - ordnet Werte in aufsteigender Reihenfolge nach ihren unformatierten Werten an.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

59

p-Regelkarte

Allgemeines über die 'p-Regelkarte' 459
p-Regelkarte: Zuweisen von Variablen zu Analyserollen 460
Regelkarten: Festlegen von Regelgrenzen 462
Regelkarte: Auswählen der durchzuführenden Tests 463
Regelkarten: Festlegen von Optionen für die Plotdarstellung 464
Regelkarten: Speichern von Analyseergebnissen 465
Festlegen von Titeln und Fußnoten 466
Anzeigen von Eigenschaften 466

Allgemeines über die 'p-Regelkarte'

Die Shewhart-Regelkarte ist ein Grafik- und Analysetool, mit dem bestimmt werden kann, ob ein Prozess innerhalb der statistischen Kontrolle liegt. Die Anwendungsroutine **p-Regelkarte** erstellt p-Regelkarten für die Proportionen der nichtkonformen (fehlerhaften) Elemente in den Untergruppenstichproben.

Sie können diese Anwendungsroutine zum Überwachen der Anteile der Defekte in einem Fertigungsprozess verwenden. Angenommen, ein Elektronikunternehmen stellt Schaltungen in Chargen zu je 500 Stück her und verwendet eine p-Regelkarte zur Überwachung der Anteile an fehlerhaften Schaltungen. Es werden dreißig Chargen untersucht, wobei die Fehler in den einzelnen Chargen gezählt werden. Die Fehlerzählungen werden in einer SAS-Datei gespeichert, die zum Erstellen der p-Regelkarte verwendet wird. Jeder Punkt auf der p-Regelkarte stellt den Anteil der nichtkonformen Elemente in einer bestimmten Untergruppe dar. Wenn beispielswiese der Wert der fehlerhaften Schaltungen 5 beträgt, wird als Wert für die erste Charge 5/500 = 0,01 dargestellt. Wenn alle Punkte innerhalb der Regelgrenzen liegen (standardmäßig ist dies 3 Sigma), kann daraus geschlossen werden, dass der Prozess innerhalb der statistischen Kontrolle liegt.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SHEWHART
Erforderliche SAS-Produkte	Base SAS, SAS/QC, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

p-Regelkarte: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Verhältnis nichtkonformer Elemente	gibt die Variablen an, welche die Anzahl der nichtkonformen Elemente enthalten. Sie müssen dieser Rolle mindestens eine Variable zuweisen
Untergruppen-Identifier	gibt die Variable an, welche die Untergruppen in den Daten identifiziert. Die Werte der Variablen geben an, wie die Beobachtungen in der Eingabetabelle in rationale Untergruppen gegliedert werden. Dies sind die üblichen Werte: ■ die Indizes, welche die Reihenfolge angeben, in der die Untergruppenstichproben gesammelt wurden ■ die Daten oder Zeiten, zu denen die Untergruppenstichproben gesammelt wurden ■ die Etiketten, welche die Untergruppenstichproben eindeutig identifizieren. Ein Untergruppen-Identifier ist erforderlich. Sie können dieser Rolle nur eine Variable zuweisen.

Optionsname	Beschreibung
	Um die Daten nach den identifizierten Untergruppen zu sortieren, aktivieren Sie das Kontrollkästchen **Nach Untergruppe sortieren**.
Stichprobengröße der Untergruppe	gibt die Stichprobengröße der Untergruppe als Werte einer Variablen an. Die Stichprobengröße der Untergruppe ist erforderlich, und Sie können dieser Rolle nur eine numerische Variable zuweisen. Aktivieren Sie das Kontrollkästchen **Numerischen Wert für Untergruppe verwenden (statt Variable)**, wenn die Stichprobengröße der Untergruppe fest ist und Sie einen konstanten Wert für die Größe angeben müssen. Geben Sie den gewünschten Wert im Feld **Stichprobengröße der Untergruppe** ein. Der Wert für die Stichprobengröße der Untergruppe muss zwischen dem maximalen Wert der Prozessvariablen und 99999 (einschließlich) liegen. **Hinweis:** Diese Rolle ist nicht erforderlich, wenn Sie das Kontrollkästchen **Numerischen Wert für Untergruppe verwenden (statt Variable)** im Fenster **Optionen** aktiviert haben.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.
Blockvariablen	gibt bis zu zwei Variablen an, welche die Daten in Blöcke aufeinanderfolgender Untergruppen gruppieren. Diese Blöcke werden in einer Legende beschriftet. Jede Blockvariable liefert eine Etikettenausprägung in der Legende.

Regelkarten: Festlegen von Regelgrenzen

Klicken Sie im Auswahlfenster auf **Regelgrenzen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Sigma-Grenzen	gibt die Breite der Regelgrenzen als Mehrfaches des Standardfehlers (Sigma) der beschreibenden Untergruppenstatistik an, die im Diagramm dargestellt wird. Die Breite muss positiv sein. Der Standardwert für das Mehrfache ist 3.
Berechnungsmethode auswählen	gibt die Berechnungsmethode für die Regelgrenzen an. Folgende Optionen können ausgewählt werden: ■ **Regelgrenzen aus den aktiven Daten berechnen** ■ **Regelgrenzen aus der ausgewählten Datei berechnen** Klicken Sie auf **Durchsuchen**, um die Datei anzugeben, die die Regelgrenzen bzw. die Parameter zu deren Berechnung enthält. Wählen Sie die gewünschte Datei aus und klicken Sie auf **Öffnen**. **Hinweis:** Bei der ausgewählten Datei muss es sich um eine LIMITS= Datei handeln. Weitere Informationen über die in einer LIMITS= Datei erforderlichen Variablen finden Sie in der Prozedurendokumentation zur erstellten Regelkarte. Wenn Sie die beschreibenden Statistiken für alle Untergruppen darstellen möchten, aktivieren Sie das Kontrollkästchen **Beschreibende Statistiken für alle Untergruppen darstellen, unabhängig davon, ob die Stichprobengröße der Untergruppe der nominalen Stichprobengröße für Regelgrenzen entspricht**. Sie können diese Statistiken beispielsweise darstellen, wenn fast alle Untergruppen eine gemeinsame Stichprobengröße haben und wenn Sie die festen (anstelle der variablen) Regelgrenzen, die der nominalen Stichprobengröße *n* entsprechen, anzeigen möchten. Wenn Sie diese Option auswählen und sich Ihre Stichprobengrößen für die Untergruppen deutlich unterscheiden, ist die Interpretation der Regelgrenzen nur für die Untergruppen sinnvoll, deren Stichprobengröße *n* entspricht.

Optionsname	Beschreibung
	Um spezielle Marker für die Datenpunkte hinzuzufügen, bei denen die Stichprobengröße ungleich `n` ist, aktivieren Sie das Kontrollkästchen **Spezielle Marker für Punkte hinzufügen, die Stichprobengrößen entsprechen, welche ungleich nominalen Stichprobengrößen für feste Regelgrenzen sind**. ■ **Regelgrenzen angeben** Geben Sie in den Feldern **Oben**, **Mitte** und **Unten** die obere Regelgrenze, die Mittellinie bzw. die untere Regelgrenze an. Wenn die Stichprobengrößen Ihrer Untergruppe variieren, müssen Sie eine nominale Stichprobengröße angeben. **Hinweis:** Wenn Sie der Rolle **Analyse gruppieren nach** Variablen zuweisen, können Sie keine Werte für Regelgrenzen angeben.

Regelkarte: Auswählen der durchzuführenden Tests

Klicken Sie im Auswahlfenster auf **Tests**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Tests wenden nicht die Regelgrenzen an, die keine Drei-Sigma-Grenzen sind oder die mit der Untergruppenstichprobengröße variieren.

Optionsname	Beschreibung
Test auswählen	Fordern Sie über die Kontrollkästchen einen oder mehrere Tests auf besondere Ursachen an, die auch als Ausführungstests, Mustertests und Western-Electric-Regeln bekannt sind. Diese Tests entdecken bestimmte nichtzufällige Muster in den Punkten, die in der primären Regelkarte dargestellt werden. Das Vorkommen eines nichtzufälligen Musters wird als Signal bezeichnet. Die Tests können auf das Vorhandensein einer besonderen Variationsursache hinweisen. Wenn Sie die Tests auswählen, erscheint eine Beschreibung des jeweiligen Tests im Feld **Beschreibung**.

Optionsname	Beschreibung
Etikett	Geben Sie in den Textfeldern die Etiketten für Punkte an, bei denen die Tests auf besondere Ursachen positiv sind. Die Etiketten dürfen maximal 16 Zeichen lang sein. Jedes Etikett erscheint bei allen Punkten im Test, an denen ein Signal vorliegt.
Text zum Identifizieren der Testsignale	■ Wählen Sie über die Dropdown-Liste eine Farbe für die angegebenen Etiketten aus. ■ Markieren Sie das Kontrollkästchen **Zonenlinien anzeigen**, um Linien einzuschließen, welche die Zonen A, B und C in einem primären Diagramm kennzeichnen. Diese Zonen werden gebildet, indem die Intervalle zwischen den Regelgrenzen in sechs gleichmäßig verteilte Unterintervalle unterteilt werden. ■ Verwenden Sie das Kontrollkästchen **3 Sigma-Grenzen überschreiben**, um nach speziellen Ursachen zu suchen, wenn Sie eine andere Sigma-Regelgrenze als das Standardvielfache von 3 angeben. Um die Sigma-Grenze zu ändern, klicken Sie im Auswahlfenster auf **Regelgrenzen**. ■ Über das Kontrollkästchen **Test auf überlappende Punktmuster anwenden** wenden Sie Tests auf besondere Ursachen auf überlappende Punktmuster an.

Regelkarten: Festlegen von Optionen für die Plotdarstellung

Einige der Darstellungsoptionen sind für unterschiedliche Regelkarten-Anwendungsroutinen gleich. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Festlegen von Achsenoptionen” auf Seite 759
- “Hinzufügen von Referenzlinien” auf Seite 760
- “Festlegen weiterer Plotoptionen” auf Seite 760
- “Festlegen von Blockvariablenoptionen” auf Seite 762

Regelkarten: Speichern von Analyseergebnissen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

Um die Analyseergebnisse zu speichern, können Sie Ausgabedateien mit diversen Statistiken speichern. Markieren Sie die Kontrollkästchen der Ausgabedaten, die Sie permanent speichern möchten.

Optionsname	Beschreibung
Datei mit Untergruppenstatistiken und Regelgrenzen	erstellt eine Ausgabedatei, welche die im Diagramm dargestellten Informationen enthält. Dies beinhaltet die Werte der Untergruppenvariablen und ihre entsprechenden beschreibenden Statistiken und Regelgrenzen. **Hinweis:** Diese Option erstellt eine Datei OUTTABLE=. Sie können diese Datei mit der Option TABLE= aus der SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Untergruppenstatistiken	erstellt eine Ausgabedatei mit den beschreibenden Statistiken der Untergruppe. **Hinweis:** Diese Option erstellt eine Datei mit OUTHISTORY=. Sie können diese Datei mit der Option HISTORY= aus der SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Regelgrenzen	erstellt eine Ausgabedatei, welche die Regelgrenzen speichert. Sie können diese Datei verwenden, wenn Sie die Berechnungsmethode für die Regelgrenzen später in dieser Anwendungsroutine auswählen.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

60

Pareto-Diagramm

Allgemeines über die Anwendungsroutine 'Pareto-Diagramm' ... **469**

Pareto-Diagramm: Zuweisen von Variablen zu Analyserollen ... **470**

Pareto-Diagramm: Beschränken der Kategorieanzahl ... **471**

Pareto-Diagramm: Festlegen von Optionen für die Plotdarstellung ... **473**
- Festlegen von Balkenoptionen ... 473
- Pareto-Diagramm: Festlegen von Achsenoptionen ... 474
- Pareto-Diagramm: Hinzufügen von Referenzlinien ... 474
- Pareto-Diagramm: Einfügungen hinzufügen ... 474
- Festlegen weiterer Plotoptionen ... 475

Festlegen von Titeln und Fußnoten ... **476**

Anzeigen von Eigenschaften ... **477**

Allgemeines über die Anwendungsroutine 'Pareto-Diagramm'

Pareto-Diagramme sind eines von sieben Basistools zur Qualitätsverbesserung in der amerikanischen und japanischen Industrie. Ein vertikales Pareto-Diagramm zeigt die relative Häufigkeit von Problemen innerhalb eines Prozesses als Balken mit abnehmender Höhe von links nach rechts an. Ein horizontales Pareto-Diagramm zeigt diese Informationen als Balken mit abnehmender Länge von oben nach unten an.

Pareto-Diagramme sind ein Tool zur Visualisierung des Pareto-Prinzips, welches angibt, dass eine kleine Untergruppe von wenigen zentralen Problemen wesentlich häufiger auftritt als die übrigen, zahlreichen, trivialeren Probleme. Ein Pareto-Diagramm unterstützt Sie beim Definieren von Prioritäten für Qualitätsverbesserungsaktivitäten durch Identifizieren der Probleme, welche die größte Aufmerksamkeit erfordern, und deren Lösung die größten Auswirkungen verspricht.

Sie können die Anwendungsroutine **Pareto-Diagramm** beispielsweise verwenden, um ein vergleichendes Pareto-Diagramm zu erstellen, mit dem Sie die Ergebnisse vor und nach einem Ereignis vergleichen können. Angenommen, während der Produktion eines Metaloxid-Halbleiterkondensators werden Fehlerursachen vor und nach der Reinigung einer Röhre im Diffusionsofen festgestellt. In der Ausgabe werden die beiden Pareto-Diagramme gestapelt, sodass Sie bestimmen können, ob durch den Reinigungsprozess mehr oder weniger Fehler auftreten.

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	PARETO
Erforderliche SAS-Produkte	Base SAS, SAS/QC, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

Pareto-Diagramm: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Prozessvariable	gibt die zu analysierende Prozessvariable an. Diese Variable bestimmt die Pareto-Kategorien für das Diagramm. Eine Prozessvariable ist erforderlich. Sie können dieser Rolle nur eine Variable zuweisen.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert.

Optionsname	Beschreibung
	Hinweis: Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt die Variable an, die als relative Gewichtungsvariable verwendet wird. Wenn Sie dieser Rolle eine Variable zuweisen, stellt die Höhe der Balken die gewichteten Häufigkeiten der Kategorien dar. Eine relative Gewichtungsvariable ist bei der Durchführung einer Pareto-Analyse nützlich, die auf Kosten statt auf der Häufigkeit des Vorkommens basiert. Sie können dieser Rolle nur eine Variable zuweisen.
Klassifizierungsvariablen	gibt eine oder zwei Variablen an, die als Klassifizierungsvariablen dienen. Wenn Sie dieser Rolle eine Variable zuweisen, erstellt die Anwendungsroutine ein vergleichendes Pareto-Diagramm, mit dem Sie Häufigkeiten über Ausprägungen der Klassifizierungsvariable hinweg vergleichen können.

Pareto-Diagramm: Beschränken der Kategorieanzahl

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Kategorien anzeigen	Sie können die Anzahl der Pareto-Kategorien beschränken, die im Diagramm angezeigt werden. Standardmäßig werden alle Kategorien für den Prozess von Balken im Pareto-Diagramm dargestellt. Um die angezeigte Anzahl der Kategorien zu ändern, wählen Sie eine der diese Optionen aus: ■ **N höchste Werte** erstellt ein Diagramm, das nur die Pareto-Kategorien mit den *n* höchsten Werten anzeigt. Der Standardwert für *n* ist `5`. ■ **Bis zu N Prozent der Gesamthäufigkeit** erstellt ein Diagramm, das die Pareto-Kategorien anzeigt, die zusammen den kumulativen Prozentwert von weniger als oder gleich *n* Prozent ergeben. Der Standardwert für *n* ist `90`.

Optionsname	Beschreibung
	■ **Mindestens N Prozent der Gesamthäufigkeit** erstellt ein Diagramm, das Pareto-Kategorien anzeigt, deren Prozentsätze größer als oder gleich *n* Prozent sind. Der Standardwert für *n* ist 5. Nachdem Sie die Option ausgewählt haben, geben Sie einen Wert für *n* im Feld **N** ein. Um eine neue Kategorie zu erstellen, die alle auszuschließenden Kategorien zusammenfasst, markieren Sie das Kontrollkästchen **Oben nicht eingeschlossene Kategorien zeigen**. Standardmäßig wird diese Kategorie mit **Andere** bezeichnet. Sie können im Feld **Etikett** ein anderes, bis zu 32 Zeichen langes Kategorieetikett angeben.
Fehlende Werte bei Prozessvariablen einschließen	Markieren Sie dieses Kontrollkästchen, um fehlende Werte als eine Kategorie zu behandeln. Diese Kategorie wird im Pareto-Diagramm von einem Balken dargestellt. Wenn die Prozessvariable eine alphanumerische Variable ist, wird ein fehlender Wert als leerer, interner (unformatierter) Wert definiert. Wenn die Prozessvariable numerisch ist, wird ein fehlender Wert als beliebiger fehlender Wert in SAS definiert. Standardmäßig werden fehlende Werte aus der Analyse ausgeschlossen.
Ausgabe in Datei speichern	Aktivieren Sie dieses Kontrollkästchen, um Ihre Ergebnisse in einer Ausgabedatei zu speichern. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Pareto-Diagramm: Festlegen von Optionen für die Plotdarstellung

Festlegen von Balkenoptionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Balken**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Balkenausrichtung und -format	Wählen Sie **Vertikal** aus, um die Balken vertikal darzustellen. Dies ist die Standardausrichtung. Wählen Sie **Horizontal** aus, um die Balken horizontal darzustellen. Ändern Sie über die Dropdown-Liste **Balken** die Balkenfarbe. Ändern Sie über die Dropdown-Liste **Umriss** die Umrissfarbe.
Balken beschriften	Sie können Etiketten über den Balken positionieren. Wählen Sie **Balkenhöhe**, um die Höhe jedes Balkens zu beschriften. Wählen Sie **Kumulativer Prozentwert**, um den kumulativen Prozentwert jedes Balkens zu beschriften. Standardmäßig wird das Etikett weggelassen.
Kurve anzeigen	Standardmäßig zeigt das Pareto-Diagramm die kumulative Prozentkurve an. Ändern Sie die Farbe der Kurve über die Dropdown-Liste **Kurvenfarbe**. Um die Kurvenpunkte mit den entsprechenden Werten zu beschriften, markieren Sie das Kontrollkästchen **Kurvenpunkte mit Werten beschriften**. Um diese Kurve und die sekundäre Achse zu unterdrücken, deaktivieren Sie das Kontrollkästchen **Kurve anzeigen**.
Balken hervorheben	Markieren Sie das Kontrollkästchen **Höchster**, um die Balken für die *n* höchsten Werte hervorzuheben. Markieren Sie **Niedrigster**, um die Balken für die *n* niedrigsten Werte hervorzuheben. Ändern Sie über die Dropdown-Liste **Farbe** die Hervorhebungsfarbe. Verwenden Sie zum Markieren mehrerer Balken die Pfeiltasten, um einen Wert für *n* im Feld **Anzahl der Balken** anzugeben. Der Wert von *n* muss kleiner oder gleich der Anzahl der Prozesskategorien sein.

Pareto-Diagramm: Festlegen von Achsenoptionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Achsen**, um auf diese Optionen zuzugreifen.

Die Ausrichtung der Balken bestimmt, ob sich die Achsenoptionen auf die primäre vertikale Achse (Y) oder die primäre horizontale Achse (X) beziehen.

Folgende Optionen können angegeben werden:

- die Achsenfarbe und -breite.
- das Etikett für die primäre Achse.
- ob ein Rahmen um das Pareto-Diagramm gezeichnet werden soll und welche Farbe für den Hintergrund verwendet werden soll.
- die Skalierung der primären Achse. Folgende Optionen können ausgewählt werden:
  - **Prozent** – Die Skala wird in Prozent der Gesamthäufigkeit angegeben. Wenn Sie eine relative Gewichtungsvariable angeben, wird die Skala in Prozent der gesamten Gewichtung angegeben. Hierbei handelt es sich um die Standardeinstellung.
  - **Anzahl** – Die Skala wird in Häufigkeitseinheiten angegeben (Anzahl). Diese Option gilt nicht, wenn Sie eine relative Gewichtungsvariable zuweisen.
  - **Gewichtung** – Die Skala wird in den gleichen Einheiten wie die relative Gewichtungsvariable angegeben. Diese Option gilt nur, wenn Sie eine relative Gewichtungsvariable zuweisen.

Pareto-Diagramm: Hinzufügen von Referenzlinien

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Referenzlinien**, um auf diese Optionen zuzugreifen.

Die Ausrichtung der Balken bestimmt, ob die Referenzlinien parallel zur vertikalen Achse oder zur horizontalen Achse erscheinen. Um Referenzlinien zu verwenden, aktivieren Sie das Kontrollkästchen **Referenzlinien verwenden**.

Pareto-Diagramm: Einfügungen hinzufügen

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Einfügung**, um auf diese Optionen zuzugreifen.

Sie können ein Pareto-Diagramm erweitern, indem Sie direkt in der Grafik ein Feld oder eine Tabelle (eine so genannte Einfügung) der beschreibenden Statistiken

hinzufügen. Wenn Sie ein vergleichendes Pareto-Diagramm erstellen, erscheint in jedem Komponentendiagramm eine Einfügung.

Hinweis: Wenn Sie eine Einfügung in das Pareto-Diagramm vornehmen, sollten Sie das Kontrollkästchen **Etikett für die Legende der Stichprobengröße anzeigen** im Fenster **Darstellung > Optionen** deaktivieren. Wenn Sie eine Einfügung in das Pareto-Diagramm vornehmen und der Stichprobengröße eine Legende hinzufügen, kann es passieren, dass die Felder für die Einfügung und die Legende in der Ausgabe überlappen. Es wird deshalb empfohlen, entweder eine Einfügung vorzunehmen oder die Legende der Stichprobengröße anzuzeigen, jedoch nicht beides gleichzeitig.

Markieren Sie das Kontrollkästchen **Einfügung einschließen**, um dem Pareto-Diagramm beschreibende Statistiken hinzuzufügen. Wählen Sie dann die Statistiken, die Sie in die Einfügung einschließen möchten. Welche Statistiken verfügbar sind, hängt von den Optionen ab, den Sie zuvor ausgewählt haben.

- Standardmäßig ist **Stichprobengröße** die einzige Statistik, die in der Einfügung erscheint.
- Wenn Sie der Rolle **Relative Gewichtung** eine Variable zugewiesen haben, können Sie **Summe der Gewichtungen** auswählen.
- Wenn Sie die Anzahl der Kategorien eingeschränkt haben, sind auch folgende Statistiken verfügbar:
  - **In beschränktem Diagramm ausgeschlossene Beobachtungen** ist verfügbar, wenn es keine Kategorie **Other** gibt, welche die ausgeschlossenen Kategorien darstellt.
  - **Anzahl der Beobachtungen in OTHER=** ist verfügbar, wenn es eine Kategorie **Other** gibt, welche die ausgeschlossenen Kategorien darstellt.
  - **Anzahl der "OTHER="-Kategorien** ist verfügbar, wenn es eine Kategorie **Other** gibt, welche die ausgeschlossenen Kategorien darstellt.

Festlegen weiterer Plotoptionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Stichprobenlegende	Aktivieren Sie das Kontrollkästchen **Etikett für die Legende der Stichprobengröße anzeigen**, um eine Stichprobenlegende anzuzeigen. Standardmäßig zeigt die Legende N=n an, wobei *n* der Gesamtzahl der Pareto-Kategorien entspricht. Wenn Sie ein vergleichendes Pareto-Diagramm erstellen, erscheint in jedem Komponentendiagramm eine Legende. Die Ausrichtung der Balken bestimmt die Position der Legende im Diagramm.

Optionsname	Beschreibung
	Hinweis: Wenn Sie eine Einfügung im Pareto-Diagramm hinzufügen, sollten Sie das Kontrollkästchen **Etikett für die Legende der Stichprobengröße anzeigen** deaktivieren. Wenn Sie eine Einfügung im Pareto-Diagramm und der Stichprobengröße eine Legende hinzufügen, kann es passieren, dass die Felder für die Einfügung und die Legende in der Ausgabe überlappen. Es wird deshalb empfohlen, entweder eine Einfügung vorzunehmen oder die Legende der Stichprobengröße anzuzeigen, jedoch nicht beides gleichzeitig. Um die Legende mit einem Rahmen zu umgeben, wählen Sie aus der Dropdown-Liste **Hintergrund** eine Farbe aus. Diese Farbe füllt den Rahmen aus. Um die Stichprobenlegende zu beschriften, geben Sie eine bis zu 32 Zeichen lange Zeichenkette in das Feld **Etikett** ein.
Vergleichsoptionen	Wenn Sie einer Analyserolle mindestens eine Klassifizierungsvariable zuweisen, können Sie über die Felder **Pro Seite** die Anordnung der Spalten und Zeilen im vergleichenden Pareto-Diagramm ändern. Wenn Sie eine Klassifizierungsvariable zuweisen, ist die Standardanordnung eine Spalte und zwei Zeilen pro Seite. Wenn Sie zwei Klassifizierungsvariablen zuweisen, ist die Standardanordnung zwei Spalten und zwei Zeilen pro Seite. Sie können auch über die Dropdown-Listen **Etikettenhintergrund** Farben zum Füllen des Rahmenbereichs der Zeilenetiketten und Spaltenetiketten angeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

61

Kreisdiagramm

Allgemeines über die Anwendungsroutine 'Kreisdiagramm' ... **479**

Kreisdiagramm: Auswählen des Diagrammtyps ... **480**

Kreisdiagramm: Zuweisen von Variablen zu Analyserollen ... **481**

Kreisdiagramm: Festlegen von Darstellungsoptionen ... **482**
- Festlegen der Optionen für Kreisdiagramme ... 482
- Festlegen des Layouts ... 483
- Festlegen erweiterter Optionen ... 484
- Weitere Darstellungsoptionen ... 484

Festlegen von Titeln und Fußnoten ... **484**

Anzeigen von Eigenschaften ... **485**

Allgemeines über die Anwendungsroutine 'Kreisdiagramm'

Die Anwendungsroutine **Kreisdiagramm** erstellt einfache, gruppierte oder gestapelte Diagramme, mit denen der relative Anteil bestimmter Daten auf das Ganze dargestellt wird. Hierzu werden die Daten als keilförmige Kreissegmente angezeigt. Jedes Segment stellt eine Datenkategorie dar. Die Segmentgröße stellt den Beitrag der Daten zur Gesamtdiagrammstatistik dar. Beispielsweise kann ein Kreisdiagramm die Verkaufszahlen der einzelnen Läden als Bruchteil der Gesamtverkaufszahlen der Ladenkette darstellen.

Weitere Informationen, siehe “Auswählen des Grafikausgabeformats” auf Seite 6.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GCHART

Name der Anforderung	Prozedur- und Produktnamen
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Kreisdiagramm: Auswählen des Diagrammtyps

Klicken Sie im Auswahlfenster auf **Kreisdiagramm**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Einfacher Kreis	erstellt ein Diagramm, das den relativen Beitrag der Teile zum Ganzen aufzeigt. Die Daten erscheinen als keilförmige "Segmente" eines Kreises. Jedes Segment stellt eine Datenkategorie dar. Die Segmentgröße stellt den Beitrag der Daten zur Gesamtdiagrammstatistik dar.
Gruppierter Kreis	erstellt ein separates Kreisdiagramm für jeden eindeutigen Wert der Variablen, die der Rolle **Kreise gruppieren nach** zugewiesen wurde.
Gestapelter Kreis	teilt das Diagramm in konzentrische Kreise gemäß den Werten der Variablen, die der Rolle **Kreise stapeln nach** zugewiesen wurden. Die Breite der Ringe, die für alle Untergruppen identisch ist, wird vom Radius des Kreises bestimmt.
Gruppierter/ Gestapelter Kreis	erstellt ein separates Kreisdiagramm für jeden eindeutigen Wert der Variablen, die der Rolle **Kreise gruppieren nach** zugewiesen wurde. Jedes Diagramm wird gemäß den Werten der Variablen, die der Rolle **Kreise stapeln nach** zugewiesen wurden, in konzentrische Kreise aufgeteilt.

Kreisdiagramm: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Hinweis: Die jeweils verfügbaren Rollen hängen vom ausgewählten Diagrammtyp ab.

Optionsname	Beschreibung
Spalte für Diagramm	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die verschiedenen Segmente. Die Spalte kann alphanumerisch oder numerisch sein.
Kreise gruppieren nach	Für jeden eindeutigen Wert der Spalte, die Sie dieser Rolle zuweisen, wird ein Kreisdiagramm erstellt. Alle Diagramme werden auf der gleichen Seite angezeigt.
Kreise stapeln nach	Für jeden eindeutigen Wert der Spalte, die Sie dieser Rolle zuweisen, wird ein konzentrischer Kreis (oder Stapel) im Kreisdiagramm erstellt. Diese Spalte wird als Untergruppenvariable bezeichnet. **Hinweis:** Wenn Sie als Grafikausgabeformat Java gewählt haben, werden die Werte der Untergruppenvariable nicht als konzentrische Kreise angezeigt. stattdessen führen Sie im entstandenen Diagramm einen Drilldown auf ein Kreissegment durch, um die Werte für die Untergruppe anzuzeigen. Wenn Sie jedoch Java-Bild als Grafikausgabeformat gewählt haben, ist das Diagramm ein statisches Bild und die beim Grafikausgabeformat Java verfügbare Drill-down-Funktionalität geht verloren.

Optionsname	Beschreibung
Summe von	Die Spalte, die Sie dieser Rolle zuweisen, bestimmt die Größe der Segmente.
Diagramme gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Diagramme, die erstellt werden. Für jeden eindeutigen Wert wird eine separate Grafik erzeugt. Jedes Kreisdiagramm wird auf einer neuen Seite angezeigt.

Kreisdiagramm: Festlegen von Darstellungsoptionen

Festlegen der Optionen für Kreisdiagramme

Klicken Sie im Auswahlfenster unter der Überschrift "Darstellung" auf **Kreise**, um auf die zugehörigen Optionen zuzugreifen.

Optionsname	Beschreibung
Farben angeben	gibt die Farbe der Segmente an. Folgende Optionen können ausgewählt werden: ■ **Standardfarbschema** - Verwendet für die Segmente das Standardfarbschema. Das Standardfarbschema wird durch den aktuellen Stil bestimmt. Sie können den Stil mit Hilfe des Stil-Managers oder des Dialogfelds **Eigenschaften** für das Ergebnis festlegen. In SAS Enterprise Guide können Sie im Dialogfeld **Optionen** auch den Stil für jeden Ergebnistyp festlegen. Im SAS Add-In for Microsoft Office können Sie im Dialogfeld **SAS Add-In - Optionen** Stile auf die SAS Report- und HTML-Formate anwenden. ■ **Existierende Farben (in einer früheren Version angegeben)** ermöglicht das Verwenden von Farben, die in einer früheren Version von SAS Enterprise Guide oder SAS Add-In for Microsoft Office festgelegt wurden. Diese Option ist nur für SAS-Inhalte verfügbar, die Sie von einer früheren Version migriert haben. ■ **Benutzerdefinierte Farben** - Ermöglicht die Angabe benutzerdefinierter Farben für die Segmente. Mit den

Optionsname	Beschreibung
	Dropdown-Listen können Sie bis zu 12 Farben für das Balkendiagramm bestimmen. wenn das Diagramm mehr als 12 Farben erfordert, werden die zusätzlichen Farben aus dem aktuellen Standardfarbschema ausgewählt. Klicken Sie auf **Zurücksetzen**, um die benutzerdefinierten Farben wieder auf den Standard zurückzusetzen. **Hinweis:** Je nach Ringdiagrammtyp werden eventuell nicht alle benutzerdefinierten Farben verwendet.
Anzahl der Segmente angeben	gibt die Anzahl der Segmente im Diagramm an. Standardmäßig wird die Anzahl der Segmente automatisch bestimmt. Bei stetigen Spalten werden die Werte in Bereiche geteilt, wobei für jeden Bereich ein Segment zur Verfügung steht. Segmente, die 4 Prozent oder weniger des Gesamtbereichs darstellen würden, werden in einem Segment mit der Bezeichnung "Sonstige" zusammengefasst. Um die Anzahl der Segmente anzugeben, die im Diagramm angezeigt werden sollen, aktivieren Sie das Kontrollkästchen **Anzahl der Segmente angeben**.

Festlegen des Layouts

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Layout**, um auf diese Optionen zuzugreifen.

Wählen Sie aus, ob im Etikett jedes Segments der Namen, der Prozentwert, der Statistikwert oder eine Kombination aus allem angezeigt werden soll. Wählen Sie für jede Komponente des Etiketts eine der folgenden Positionen:

- im Segment
- außerhalb des Kreises oder Rings
- außerhalb des Segments mit einem darauf zeigenden Pfeil

Sie können für die Schrift der Segmentetiketten und für die Segmente die gleiche Farbe wählen. Diese Option sollten Sie nicht verwenden, wenn Sie die Etiketten innerhalb der Segmente selbst anzeigen möchten.

Sie können den Prozentwert und das Etikett für das Segment "Sonstige" angeben. Dieses Segment ist eine Sammlung aller Midpoints, deren statistische Werte des Diagramms kleiner oder gleich dem Gesamtprozentwert sind. Der Standardwert für den prozentualen Anteil an der Summe beträgt 4. Das bedeutet, dass jedes Segment, das 4 Prozent oder weniger von der Gesamtprozentzahl enstpricht, im Segment "Sonstige" dargestellt wird.

Sie können für das Etikett des Segments "Sonstige" einen Text mit max. 16 Zeichen eingeben. Das Standardetikett lautet "Sonstige". Das Segment "Sonstige" ist das letzte Kreissegment, unabhängig von der Reihenfolge der Segmente.

Wenn nur ein Midpoint in die Kategorie "Sonstige" fällt, bleibt das Segment für diesen Midpoint an seiner ursprünglichen Position im Kreis und behält sein ursprüngliches Etikett bei. Angenommen ein Kreis zeigt die Getränkeumsätze in Prozentwerten und enthält die folgenden Segmente und Prozentwerte: Kaffee 35%, Tee 15%, Mineralwasser 5% und Milch 45%. Wenn der Gesamtprozentwert 5 beträgt, bleibt das Segment für Wasser das dritte Segment und wird nicht zum letzten Segment, und dieses Segment bekommt das Etikett "Wasser" und nicht "Sonstige".

Verwenden Sie bei gruppierten Diagrammen den Bereich **Gruppiertes Layout**, um die Diagramme anzuordnen. Geben Sie die Anzahl der Spalten oder Zeilen im Ausgabebereich an.

Festlegen erweiterter Optionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Erweitert**, um auf diese Optionen zuzugreifen.

Wählen Sie aus der Dropdown-Liste die Statistik, mit der die Diagrammsegmente berechnet werden sollen. Wenn Sie der Rolle **Summe von** keine Variable zuweisen, können Sie die Häufigkeit oder den Prozentwert für jedes Segment berechnen. Standardmäßig wird die Häufigkeit berechnet. Wenn Sie der Rolle **Summe von** eine Variable zuweisen, können Sie die Summe oder den Mittelwert für jedes Segment berechnen. Standardmäßig wird die Summe berechnet.

Weitere Darstellungsoptionen

Einige Darstellungsoptionen gelten für alle Grafik-Anwendungsroutinen. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Festlegen von Legendenoptionen” auf Seite 749
- “Anpassen des Diagrammbereichs” auf Seite 750

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

62

Assistent für Kreisdiagramme

Allgemeines über die Anwendungsroutine 'Assistent für Kreisdiagramme' **487**
Assistent für Kreisdiagramme: Auswählen der Daten **487**
Assistent für Kreisdiagramme: Zuweisen von Variablen zu Rollen **488**
Assistent für Kreisdiagramme: Festlegen der Darstellungsoptionen **489**
Assistent für Kreisdiagramme: Festlegen von Titeln und Fußnoten **489**

Allgemeines über die Anwendungsroutine 'Assistent für Kreisdiagramme'

Der **Assistent für Kreisdiagramme** unterstützt Sie bei der Erstellung zwei- oder dreidimensionaler Kreisdiagramme mit gruppierten oder gestapelten Kreisen. Jedes Kreissegment stellt eine Datenkategorie dar. Die Segmentgröße stellt den Beitrag der Daten zur Gesamtdiagrammstatistik dar.

Assistent für Kreisdiagramme: Auswählen der Daten

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im

Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Klicken Sie auf **Weiter**, um Variablen Rollen zuzuweisen.

Assistent für Kreisdiagramme: Zuweisen von Variablen zu Rollen

1. Bevor Sie Variablen zu Rollen zuweisen können, müssen Sie die Daten auswählen.
2. So weisen Sie den folgenden Rollen Variablen zu:
   - **Segment** - Die Werte der Variablen, die Sie dieser Rolle zuweisen, legen die verschiedenen Segmente fest. Die Variable kann entweder alphanumerisch oder numerisch sein.

     Sie können auch diese Optionen angeben:

     - **Ein Segment pro eindeut. Datenwert** - Diese Auswahl hat keine Auswirkung auf alphanumerische Variablen. Bei numerischen Variablen wird für jeden eindeutigen Wert ein Segment gezeichnet.
     - **Segmente automatisch erstellen** - Numerische Variable werden wie kontinuierliche Variable behandelt, und die Werte werden in Bereiche unterteilt, wobei für jeden Bereich ein Segment verwendet wird. Die Anzahl der Bereiche wird automatisch festgelegt. Segmente, die 4 Prozent oder weniger des Gesamtbereichs darstellen würden, werden in einem Segment mit der Bezeichnung "Sonstige" zusammengefasst.
     - **Segmentgröße** - Die Variable oder die Statistik, die Sie dieser Rolle zuweisen, bestimmt die Größe der Segmente.
       - Wenn Sie dieser Rolle eine Variable zuweisen, können Sie auswählen, ob die Summe oder der Mittelwert jedes Segments berechnet wird. Standardmäßig wird die Summe berechnet.
       - Wenn Sie dieser Rolle keine Variable zuweisen, können Sie die Häufigkeit oder den Prozentwert für jedes Segment berechnen. Standardmäßig wird die Häufigkeit berechnet.
3. (Optional) Weisen Sie folgenden Rollen Variablen zu:
   - **Gruppieren nach** - Die Werte der Variablen, die Sie dieser Rolle zuweisen, sind in Gruppen angeordnet. Für jeden eindeutigen Wert wird ein separater Kreis erzeugt.
   - **Stapeln nach** - Die Werte der Variablen, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der gestapelten, konzentrischen Kreise für jedes Kreisdiagramm. Wenn es sich beim Grafikausgabetyp um Java handelt, besitzt das Diagramm eher Drill-Down-Funktionalität und nicht konzentrische Kreise.

- **Gruppiert nach** - Für jede BY-Gruppe werden separate Diagramme generiert. Die Gruppen werden von den Werten der Variablen bestimmt, die Sie dieser Rolle zuweisen. Es wird für jede BY-Gruppe eine Ausgabe generiert.

4 Klicken Sie auf **Weiter**, um die Darstellungsoptionen festzulegen.

Assistent für Kreisdiagramme: Festlegen der Darstellungsoptionen

Der **Assistent für Kreisdiagramme** bestimmt automatisch die im Diagramm verwendeten Farben. Sie können dieses Farbschema in der Dropdown-Liste **Schema** anzeigen.

Geben Sie diese Darstellungsoptionen an:

- zwei- oder dreidimensionales Diagramm
- Legende
- Variablennamen in den Segmenten des Diagramms
- Prozentwert der einzelnen Segmente im Diagramm
- Datenwerte für die einzelnen Segmente im Diagramm

Klicken Sie auf **Weiter**, um Titel und Fußnoten anzugeben.

Assistent für Kreisdiagramme: Festlegen von Titeln und Fußnoten

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

1 Vor dem Angeben der Titel und Fußnoten müssen Sie die Darstellungsoptionen festlegen.

2 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld. Sie können in Titeln und Fußnoten Makrovariablen verwenden.

3 Klicken Sie auf **Fertig stellen**, um die Anwendungsroutine auszuführen.

63

P-P-Plots

Allgemeines über die Anwendungsroutine 'P-P-Plots' ... **491**

Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen ... **492**

P-P-Plots: Auswählen der Verteilung ... **494**

P-P-Plots: Anpassen der Verteilung ... **495**

Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung ... **495**

Festlegen von Achsenoptionen ... 495
Hinzufügen von Referenzlinien ... 496
Einfügungen hinzufügen ... 496
Festlegen weiterer Plotoptionen ... 496
Festlegen von Optionen für Spezifikationsgrenzen ... 499

Prozessfähigkeitsanalyse: Festlegen der Analysetypen ... **499**

Festlegen von Titeln und Fußnoten ... **500**

Anzeigen von Eigenschaften ... **500**

Allgemeines über die Anwendungsroutine 'P-P-Plots'

Eine Prozessfähigkeitsanalyse vergleicht die Verteilung der Ausgabe eines Prozesses in statistischer Kontrolle mit den Spezifikationsgrenzen, um die Konsistenz zu bestimmen, mit der die Spezifikationen erfüllt werden. Ein P-P-Plot (Probability-Probability-Plot, auch als Prozent-Plot bezeichnet) wird bei der Prozessfähigkeitsanalyse eingesetzt, um die empirische kumulative Verteilungsfunktion (ECDF) einer Variablen mit einer angegebenen theoretischen kumulativen Verteilungsfunktion, wie z.B. der Normalverteilung, zu vergleichen. Wenn die beiden Verteilungen übereinstimmen, bilden die Punkte im Plot ein lineares Muster, das durch den Ursprung geht und eine Einheitssteigung aufweist. Daher können Sie mit einem P-P-Plot ermitteln, wie gut eine theoretische Verteilung einen Satz an Messwerten modelliert.

Sie können einen P-P-Plot auch zum Analysieren eines Fertigungsprozesses verwenden. Angenommen, es werden die Abstände zwischen zwei Bohrungen, die in 50 Stahlbleche geschnitten wurden, gemessen und in einer Datei gespeichert. Der Schneidprozess befindet sich innerhalb der statistischen Kontrolle. Sie möchten überprüfen, ob die Abstände normal verteilt sind. Hierzu können Sie mithilfe der Anwendungsroutine **P-P-Plot** einen Plot erstellen, der auf der normalen kumulativen Verteilungsfunktion basiert. Anhand dieses Plots können Sie bestimmen, ob die Abstände normal verteilt sind.

Hinweis: P-P-Plots sind nicht mit Wahrscheinlichkeitsplots zu verwechseln, die einen Satz geordneter Messwerte mit Perzentilen aus einer angegebenen Verteilung vergleichen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	CAPABILITY
Erforderliche SAS-Produkte	Base SAS, SAS/QC, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Analysevariablen	gibt die zu analysierenden numerischen Variablen an. Sie müssen dieser Rolle mindestens eine Variable zuweisen.

Optionsname	Beschreibung
	Bei allen Plots, außer dem P-P-Plot, können Sie eine Obergrenze, einen Zielwert und eine Untergrenze im Bereich **Spezifikationsgrenzen für *Variablenname*** für jede Analysevariable angeben. Wenn Sie diese Grenzen angeben, enthält die Ausgabe automatisch eine Tabelle mit den Tests auf Normalverteilung, eine Tabelle mit den Spezifikationsgrenzen und eine Tabelle mit den Prozessfähigkeitsindizes.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt die Variable an, die als relative Gewichtungsvariable verwendet werden soll. Wenn Sie dieser Rolle eine Variable zuweisen, werden mit dem Wert der Variablen in jeder Beobachtung gewichtete Statistiken berechnet. Sie können dieser Rolle nur eine Variable zuweisen. **Hinweis:** Sie sollten der Rolle **Relative Gewichtung** keine Variable zuweisen, wenn Sie ein Histogramm, Wahrscheinlichkeitsplot, P-P-Plot, Q-Q-Plot oder CDF-Plot erstellen. Falls Sie dieser Rolle ein Variable zugewiesen haben, müssen Sie eine eventuelle Ausgabe mit Vorsicht interpretieren.
Klassifizierungsvariablen (nur verfügbar für die Anwendungsroutine Histogramm)	gibt eine oder zwei Variablen an, die als Klassifizierungsvariablen dienen. Wenn Sie dieser Rolle eine Variable zuweisen, erstellt die Anwendungsroutine ein Vergleichshistogramm, mit dem Sie die Verteilung einer Analysevariablen über Ausprägungen der Klassifizierungsvariablen hinweg vergleichen können. **Hinweis:** Um vergleichende Wahrscheinlichkeitsplots oder vergleichende Q-Q-Plots zu erstellen, verwenden Sie die Anwendungsroutine Verteilungsanalyse.

P-P-Plots: Auswählen der Verteilung

Klicken Sie im Auswahlfenster unter der Überschrift 'Verteilungen' auf **Zusammenfassung**, um auf diese Optionen zuzugreifen.

Sie können eine theoretische Verteilung auf einem P-P-Plot anpassen. Wenn Sie die Verteilung ausgewählt haben, wählen Sie den Namen der Verteilung im Auswahlfenster aus, um die Parameterwerte und Linieneigenschaften für diese Verteilung anzugeben.

Optionsname	Beschreibung
Verteilungen	Verwenden Sie die Kontrollkästchen, um eine oder mehrere der folgenden geschätzten Dichtekurven auszuwählen: ■ **Normal** passt eine normale Dichte mithilfe eines Mittelwerts (My) und einer Standardabweichung (Sigma) an. ■ **Lognormal** passt eine Lognormal-Dichte mithilfe eines Skalierungsparameters (Zeta), eines Schwellenwertparameters (Theta) und eines Formparameters (Sigma) an. ■ **Exponentiell** passt eine exponentielle Dichte mithilfe eines Skalierungsparameters (Sigma) und eines Schwellenwertparameters (Theta) an. ■ **Weibull** passt eine Weibull-Dichte mit drei Parametern mit einem Skalierungsparameter (sigma), einem Schwellenwertparameter (theta) und einem Formparameter (c) an. ■ **Beta** passt eine Beta-Dichte mit einem Skalierungsparameter (Sigma), einem Schwellenwertparameter (Theta) und Formparametern (Alpha und Beta) an. ■ **Gamma** passt eine Gamma-Dichte mit einem Skalierungsparameter (Sigma), einen Schwellenwertparameter (Theta) und einen Formparameter (Alpha) an.
Grafikstil	Wählen Sie den Grafikstil, der in der Ausgabe verwendet werden soll. Sie haben die Auswahl zwischen üblichen SAS-Grafiken und ODS Statistical Graphs (statistischen ODS-Grafiken). Weitere Informationen zu ODS-Grafiken finden Sie unter *SAS Output Delivery System: User's Guide*.

P-P-Plots: Anpassen der Verteilung

Klicken Sie im Auswahlfenster unter der Überschrift **Verteilungen** auf den Namen der Verteilung, die Sie anpassen möchten, um auf diese Optionen zuzugreifen.

Hinweis: Diese Verteilungsoptionen sind nur durch vorherige Auswahl der Verteilung auf der Seite **Zusammenfassung** verfügbar, oder durch Aktivieren des Kontrollkästchens ***Name der Verteilung*** auf der Seite dieser Verteilung.

Für die Normal-, Lognormal-, Exponential-, Weibull-, Beta- und Gamma-Verteilungen können Sie die Parameter für jede Analysevariable oder für alle Analysevariablen angeben. Standardmäßig schätzt die Anwendungsroutine die Parameterwerte für alle Analysevariablen.

So legen Sie für jede Analysevariable Parameter fest:

1 Deaktivieren Sie das Kontrollkästchen **Verteilung auf alle Variablen anwenden**.

2 Aktivieren Sie das Kontrollkästchen der Variablen im Feld **Analysevariablen**.

3 Geben Sie die Parameterwerte an.

Hinweis: Um zu den geschätzten Parameterwerten zurückzukehren, klicken Sie auf **Schätzer verwenden**.

Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung

Festlegen von Achsenoptionen

Klicken Sie im Auswahlfenster unter **Darstellung** ⇨ **Achsen** auf **Achsen**, um auf die Optionen zuzugreifen.

Folgende Optionen können angegeben werden:

- die Farbe der Achsen
- die Farbe der Achsenmarkierungen

- ein Etikett für die horizontalen und vertikalen Achsen

Hinzufügen von Referenzlinien

Klicken Sie im Auswahlfenster unter **Darstellung** ⇨ **Achsen** auf **Horizontal** bzw. **Vertikal**, um auf die Optionen zuzugreifen.

Um Referenzlinien zu verwenden, wählen Sie **Referenzlinien verwenden**. Sie können die Referenzlinie formatieren und die Position jeder Referenzlinie festlegen.

Einfügungen hinzufügen

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Einfügung**, um auf diese Optionen zuzugreifen.

Sie können einen Plot erweitern, indem Sie direkt in der Grafik ein Feld oder eine Tabelle (eine so genannte Einfügung) der beschreibenden Statistiken hinzufügen. Wenn Sie ein Vergleichshistogramm erstellen, erscheint in jedem Komponentenhistogramm eine Einfügung.

Wählen Sie **Einfügung einschließen**, um eine beschreibende Statistik zum Plot hinzuzufügen. Wählen Sie dann die Statistiken, die Sie in die Einfügung einschließen möchten.

Hinweis: Die Statistiken für die Berechnung der Prozessfähigkeitsindizes sind nur verfügbar, wenn Sie Spezifikationsgrenzen festgelegt und dann der Rolle **Analysevariablen** zugewiesen haben. Sie können Spezifikationsgrenzen für Histogramm, CDF-Plots, Wahrscheinlichkeitsplots und Q-Q-Plots erstellen. Weitere Informationen zur Berechnung der Prozessfähigkeitsindizes finden Sie in der Hilfe zur Prozedur CAPABILITY.

Festlegen weiterer Plotoptionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Optionen**, um auf diese Optionen zuzugreifen.

Die verfügbaren Plotdarstellungsoptionen hängen vom Typ des Plots ab, den die Anwendungsroutine erzeugt.

Optionsname	Beschreibung
Hintergrund	bestimmt die Hintergrundfarbe des Plots.

Optionsname	Beschreibung
Balkenumriss und **Balkenfüllung**	legt die Farben der Balken in einem Histogramm fest. **Hinweis:** Diese Optionen sind nur für Vergleichshistogramme verfügbar.
Beide Achsen umkehren (Perzentile auf entgegengesetzte Achsen)	kehrt die Achsen auf dem Plot um. Wenn Sie diese Option auswählen, funktionieren die von Ihnen ausgewählten Achsenoptionen und Referenzlinien eventuell nicht mehr, da sich die Ausrichtung des Plots ändert. **Hinweis:** Diese Option ist nur für Wahrscheinlichkeitsplots und Q-Q-Plots verfügbar.
Keine Standardlegenden	verhindert die Anzeige der Standardlegende mit dem Plot.
Midpoints verwenden	geben Sie Midpoints für das Histogramm an. Die in der Dropdown-Liste verfügbaren Optionen hängen davon ab, ob Sie ein Histogramm oder ein Vergleichshistogramm erstellen. Wenn Sie der Rolle **Klassifizierungsvariablen** eine Variable zuweisen, erstellen Sie ein Vergleichshistogramm. Für Vergleichshistogramme stehen folgende Optionen zur Verfügung: ■ **Midpoints angeben** - Ermöglicht das Angeben der Start- und End-Midpoints. Sie können auch das Schrittintervall festlegen, das für die Berechnung der anderen Midpoints verwendet werden soll. Der Midpoint ist der Mittelwert des Wertebereichs, der von den einzelnen Balken dargestellt wird. ■ **Schlüssel** - Bestimmt die Midpoints für die Daten in der Schlüsselzelle. Standardmäßig werden die Ausprägungen der Klassifizierungsvariablen von oben nach unten (links nach rechts) in aufsteigender Reihenfolge der internen (nicht formatierten) Werte der ersten Klassifizierungsvariablen angezeigt. Wenn Sie nur eine Klassifizierungsvariable angegeben haben, ist die Schlüsselzelle die Ausprägung, die als erstes in dieser Reihenfolge erscheint. Wenn Sie zwei Klassifizierungsvariablen angegeben haben, ist die Schlüsselzelle die Kombination der Ausprägungen von Variable 1 und Variable 2, die als erstes in dieser Reihenfolge erscheint. Die Auswahl der Schlüsselzelle bestimmt die einheitliche Horizontalachse, die für alle Zellen verwendet wird. Die Midpoint-Liste für die Schlüsselzelle wird dann wie erforderlich in beide Richtungen erweitert, bis sie die Daten in den übrigen Zellen abdeckt. ■ **Einheitlich** - Bestimmt die Anzahl der Midpoints abhängig von der Gesamt-Stichprobengröße.

Optionsname	Beschreibung
	Wenn Sie der Rolle **Klassifizierungsvariablen** keine Variable zuweisen, erstellen Sie ein Histogramm. Für Histogramme sind die folgenden Optionen verfügbar: ■ **Midpoints angeben** - Ermöglicht das Angeben der Start- und End-Midpoints. Sie können auch das Schrittintervall festlegen, das für die Berechnung der anderen Midpoints verwendet werden soll. Der Midpoint ist der Mittelwert des Wertebereichs, der von den einzelnen Balken dargestellt wird. ■ **Anzahl der Intervalle angeben** - Ermöglicht, die Anzahl der Intervalle (auch Histogrammintervalle genannt) für die Daten anzugeben. Standardmäßig basiert die Anzahl der angezeigten Intervalle auf dem Datenbereich. Sie können jedoch eine Maximalzahl anzuzeigender Intervalle angeben. **Hinweis:** Sie können nur dann Midpoints angeben, wenn Sie die Anwendungsroutine **Histogramme** verwenden.
Größe ben.def. Diagramm angeben (in Pixel)	legt die Bildgröße fest. Zum Vergleich wird die Gesamtbildschirmgröße angezeigt in den Feldern **Breite** und **Höhe** angezeigt.

Hinweis: Vergleichsoptionen sind für Vergleichshistogramme verfügbar. Sie können ein Vergleichshistogramm erstellen, indem Sie in der Anwendungsroutine **Histogramme** der Rolle **Klassifizierungsvariablen** eine Variable zuweisen.

Optionsname	Beschreibung
Zeilen und **Spalten**	Wenn Sie einer Analyserolle mindestens eine Klassifizierungsvariable zuweisen, können Sie über die Felder **Zeilen pro Seite** und **Spalten pro Seite** die Anordnung der Spalten und Zeilen im Vergleichshistogramm ändern. Wenn Sie eine Klassifizierungsvariable zuweisen, ist die Standardanordnung zwei Zeilen und eine Spalte pro Seite. Wenn Sie zwei Klassifizierungsvariablen zuweisen, ist die Standardanordnung zwei Spalten und zwei Zeilen pro Seite. Sie können auch über die Dropdown-Listen **Rahmenseite** und **Rahmenoberseite** Farben zum Füllen des Rahmenbereichs der Zeilenetiketten und Spaltenetiketten angeben.
Anzahl der Intervalle	Standardmäßig basiert die Anzahl der angezeigten Intervalle auf dem Datenbereich. Sie können jedoch eine Maximalzahl anzuzeigender Intervalle angeben. Sie können auch eine Standardabweichung angeben, so dass die Anzahl der angezeigten Intervalle von einem Bereich des

Optionsname	Beschreibung
	Standardabweichungswerts über und unter dem Mittelwert der Daten abhängt.

Festlegen von Optionen für Spezifikationsgrenzen

Hinweis: Diese Optionen sind für Histogramme, Wahrscheinlichkeitsplot, Q-Q-Plots und CDF-Plots verfügbar.

Klicken Sie im Auswahlfenster unter **Darstellung** auf **Spezifikationsgrenzen**, um die untere Spezifikationsgrenze, das Ziel und die obere Spezifikationsgrenze anzugeben.

Hinweis: Die Optionen für **Spezifikationsgrenzen** sind nur verfügbar, wenn Sie eine oder mehrere Grenzen für mindestens eine Analysevariable angeben. Um eine Spezifikationsgrenze für eine Analysevariable anzugeben, klicken Sie im Auswahlfenster auf **Daten** und wählen dann die Analysevariable aus.

Prozessfähigkeitsanalyse: Festlegen der Analysetypen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

Standardmäßig enthält der Prozessfähigkeitsanalysebericht die folgenden Tabellen: eine mit Grundkonfidenzintervallen, eine mit Grundmaßen, eine mit Tests auf Lageparametern und eine mit Momenten.

Um eine Tabelle zum Bericht hinzuzufügen, wählen Sie die Tabelle im Feld **Tabellen** aus.

Um die Analyseoptionen festzulegen, wählen Sie die Tabelle im Feld **Tabellen** aus. Sind Optionen verfügbar, erscheint rechts ein Feld, in dem Sie die Optionen angeben können.

Um alle Tabellen mit beschreibenden Statistiken im Bericht zu unterdrücken, aktivieren Sie die Option **Beschreibende Statistiken und Tabellen für Prozessfähigkeitsindizes unterdrücken**. Folgende Tabellen werden durch Auswahl dieser Option nicht unterdrückt:

- Tabellen, die von der ausgewählten Verteilung erstellt wurden. Sie können diese Tabellen unterdrücken, indem Sie im Fenster für diese Verteilung das Kontrollkästchen **Verteilungstabellen unterdrücken** aktivieren.
- Tabellen, die mit der INTERVALS-Anweisung erstellt wurden. Um diese Tabellen zu unterdrücken, können Sie die Option NOPRINT angeben. Weitere Informationen zur Anweisung INTERVALS finden Sie in der Hilfe zur Prozedur CAPABILITY.

Um Ihre Ergebnisse in eine Ausgabedatei zu speichern, aktivieren Sie das Kontrollkästchen **Ausgabestatistiken in einer Datei speichern**. Die Datei enthält univariate Statistiken und Prozessfähigkeitsindizes.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

   Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

64

Zeitreihendaten vorbereiten

Allgemeines über die Anwendungsroutine 'Zeitreihendaten vorbereiten' ... 503
Zeitreihendaten vorbereiten: Zuweisen von Variablen zu Analyserollen ... 504
Zeitreihendaten vorbereiten: Festlegen der Ausgabehäufigkeit ... 505
Zeitreihendaten vorbereiten: Festlegen von Zeitreihenmerkmalen ... 506
Zeitreihendaten vorbereiten: Angeben der Transformationen für jede Zeitreihenvariable ... 508
Zeitreihendaten vorbereiten: Festlegen von Ergebnisoptionen ... 509
Anzeigen von Eigenschaften ... 509
Erstellen der Beobachtungszeilen für fehlende Zeilen in der Datei ... 510
Konvertieren von Zeitreihendaten von einer Stichprobenhäufigkeit in eine andere ... 511
Ändern der Beobachtungsmerkmale der Eingabedaten ... 512
Konvertieren aperiodischer Reihen in periodische Schätzer ... 513
Interpolieren fehlender Werte ... 513
Anwenden generischer Transformationen auf die Daten ... 515

Allgemeines über die Anwendungsroutine 'Zeitreihendaten vorbereiten'

Mit der Anwendungsroutine **Zeitreihendaten vorbereiten** können Sie Daten für die Analyse durch andere Zeitreihen-Anwendungsroutinen aufbereiten. Sie kann auch zum Durchführen generischer Transformationen an Daten verwendet werden, die in anderen Anwendungsroutinen verwendet werden sollen. Beispielsweise können Sie mathematische Operationen, Funktionen und andere Transformationen auf Variablen in einer Datei anwenden.

Vorsicht beim Verwenden dieser Anwendungsroutine zum Bearbeiten von Zeitreihendaten, da sich die ausgewählten Optionen auf die Güte der Analysen auswirken, die später für die Daten ausgeführt werden.

Mit dieser Anwendungsroutine können folgende Aufgaben erledigt werden:

- Interpolation vierteljährlicher Schätzer aus einer jährlichen Reihe oder Aggregation vierteljährlicher Werte zum Erstellen einer jährlichen Reihe
- Kombination von Zeitreihen, die mit unterschiedlichen Stichprobenhäufigkeiten gemessen werden. Angenommen, Ihnen liegen Daten zu monatlichen Geldmengen, zum vierteljährlichen Bruttoinlandsprodukt und zu wöchentlichen Zinssätzen vor, anhand derer Sie die Analyse eines Modells vornehmen möchten, bei der alle Variablen verwendet werden. Zum Ausführen der Analyse müssen Sie zunächst die Reihen in eine allgemeine Häufigkeit konvertieren und die Variablen in einer Datei zusammenfassen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	EXPAND
Erforderliche SAS-Produkte	Base SAS, SAS/ETS
Empfohlene zusätzliche SAS-Produkte	Ohne

Zeitreihendaten vorbereiten: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Zeit-ID-Variable	gibt die Variable an, mit der den Zeilen Daten zugeordnet werden. Sie müssen dieser Rolle genau eine Variable zuweisen. Die erste Variable, die ein SAS-Datums-, Uhrzeit- oder Datetime-Format enthält, wird dieser Rolle automatisch zugewiesen. Sie können auch eine Zeit-ID-Variable erstellen.
Zeitreihenvariablen	gibt die Variablen an, die für die Analyse vorbereitet werden, indem mathematische Operationen, die in dieser Anwendungsroutine verfügbar sind, angewendet werden. Wenn keine Variablen angegeben werden, konvertiert die Anwendungsroutine alle numerischen Variablen in den Eingabedaten entsprechend der von Ihnen gesetzten Merkmalsoptionen.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits für eine andere Rolle ausgewählt ist.

Zeitreihendaten vorbereiten: Festlegen der Ausgabehäufigkeit

Klicken Sie im Auswahlfenster unter der Überschrift 'Merkmale' auf **Häufigkeiten**, um auf diese Optionen zuzugreifen.

- Um die Ausgabehäufigkeit unverändert zu lassen, wählen Sie **Gleich wie Eingabehäufigkeit**.
- Um das Häufigkeitsintervall zu ändern, wählen Sie **Intervall** und dann das Ausgabeintervall aus der Dropdown-Liste.
- Um das Verhältnis der Eingabehäufigkeit zur Ausgabehäufigkeit zu ändern, wählen Sie **Verhältnis**. Verwenden Sie die Felder **Eingabebeobachtungen** und **Ausgabebeobachtungen**, um das Verhältnis von Eingabebeobachtungen zu Ausgabebeobachtungen anzugeben.

Zeitreihendaten vorbereiten: Festlegen von Zeitreihenmerkmalen

Klicken Sie im Auswahlfenster unter der Überschrift 'Merkmale' auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Interpolationsmethode	gibt die Interpolationsmethode an. Diese Methoden sind verfügbar: ■ **Kubischer Spline** - passt eine kubische Splinekurve an die Eingabewerte an. Ein kubischer Spline ist eine segmentierte Funktion, die aus polynomialen Funktionen dritten Grades (kubisch) besteht, die so gejoint wurden, dass die gesamte Kurve und ihre ersten und zweiten Ableitungen kontinuierlich sind. ■ **Linearer Spline** - passt eine kontinuierliche Kurve an die Daten an, indem aufeinander folgende Liniensegmente miteinander verbunden werden. ■ **Stufenfunktion** - passt eine nichtkontinuierliche, stückweise konstante Kurve an. ■ **Einfache Aggregation** - führt eine einfache Aggregation der Zeitreihen ohne Interpolation der fehlenden Werte durch. ■ **Keine Interpolation** - gibt an, dass keine Interpolation durchgeführt wird. **Hinweis:** Die Methode **Keine Interpolation** ist nicht verfügbar, wenn eine Häufigkeitskonvertierung ausgewählt wurde. Weitere Informationen, siehe "Zeitreihendaten vorbereiten: Festlegen der Ausgabehäufigkeit" auf Seite 505. Für die Methode "Kubischer Spline" können Sie die unterste und die oberste Endpunktbedingung angeben. Hier sind die Optionen für Endpunktbedingungen: ■ **Kein Knoten** - bedeutet, dass die ersten beiden und die letzten beiden Splinestücke als Teil derselben kubischen Kurve beschränkt sind. Hierbei handelt es sich um die Standardeinstellung. ■ **Natürlicher Spline** - gibt die Bedingung für natürliche Splines an. Die zweite Ableitung der Splinekurve ist am Endpunkt auf Null beschränkt. ■ **Steigung** - gibt die erste Ableitung der Splinekurve am Endpunkt an.

Optionsname	Beschreibung
	■ **Krümmung** - gibt die zweite Ableitung der Splinekurve am Endpunkt an. Bei den Optionen **Steigung** und **Krümmung** müssen Sie den ersten Ableitungswert bzw. den zweiten Ableitungswert angeben. Der Standard für beide Ableitungswerte ist 1.
Beobachtungsmerkmale	gibt das Beobachtungsmerkmal für die Eingabe- und Ausgabezeitreihe an. Das Beobachtungsmerkmal definiert, welchen Beobachtungstyp jeder Wert in der Zeitreihe hat. Dies stellt sicher, dass die Interpolationen richtig durchgeführt werden. Folgende Optionen stehen zur Verfügung: ■ **Summe** - gibt an, dass die Datenwerte Periodensummen für das Zeitintervall darstellen, das der Beobachtung entspricht. ■ **Mittelwert** - gibt an, dass die Datenwerte Periodenmittelwerte darstellen. ■ **Anfang** - gibt an, dass die Datenwerte Anfangswerte der Periode sind. Hierbei handelt es sich um die Standardeinstellung. ■ **Mitte** - gibt an, dass die Datenwerte Periodenmittelpunkte sind. ■ **Ende** - gibt an, dass die Datenwerte Endwerte der Periode sind. ■ **Ableitung** - gibt an, dass die Ausgabedaten die Ableitungen der kubischen Splinekurve sind, die an die Eingabedaten angepasst wurden. **Hinweis:** Die Option **Ableitung** ist nur in der Dropdown-Liste **Ausgabe** verfügbar. Der Unterschied zwischen der Datumsausrichtung und den Beobachtungsmerkmalen kann verwirrend sein. Die Option zur Datumsausrichtung steuert die Datumswerte, die der Zeit-ID-Variable für jede Beobachtung zugewiesen sind. Die Beobachtungsmerkmale steuern, wie die Zeitreihenwerte mathematisch interpretiert werden, wenn Interpolationen erfolgen.
Datumsausrichtung im Intervall	legt fest, wie Datumsangaben innerhalb der einzelnen Intervalle ausgerichtet werden sollen. Wenn Sie beispielweise ein monatliches Intervall angegeben haben, beginnt das Zeitintervall bei der Option **Anfang** am ersten Tag des Monats, bei der Option **Mitte** zur Monatsmitte und bei der Option **Ende** am letzten Tag des Monats.

Zeitreihendaten vorbereiten: Angeben der Transformationen für jede Zeitreihenvariable

Klicken Sie im Auswahlfenster unter der Überschrift "Transformationen" auf **Vor Interpolation** bzw. **Nach Interpolation**, um auf diese Optionen zuzugreifen.

Einzelne Transformationen können für jede Variable angewendet werden, die der Rolle **Zeitreihenvariablen** zugewiesen ist.

Optionsname	Beschreibung
Vor Interpolation-/ Nach Interpolation-Aktionen	Klicken Sie auf **Hinzufügen**, um eine Liste verfügbarer Transformationen anzuzeigen. Markieren Sie die Kontrollkästchen aller Transformationen, die Sie hinzufügen möchten und klicken Sie auf **OK**. Um eine Transformation zu entfernen, wählen Sie sie aus und klicken anschließend auf **Entfernen**. Die Transformationen erfolgen in der Reihenfolge, in der sie in der Liste erscheinen. Um die Reihenfolge der Transformationen zu ändern, wählen Sie eine Transformation aus und verschieben Sie diese in der Liste mithilfe der Schaltflächen **Nach oben** und **Nach unten**. Um die Transformation y = 5 + log(x + 10) auf die Eingabezeitreihe anzuwenden, würden Sie also folgende Transformationen in die Liste in dieser Reihenfolge eingeben: 1 Konstante hinzufügen (n = 10) 2 log 3 3 Konstante hinzufügen (n = 5)
Optionen für gewählten Vorgang	Bei einigen Transformationen müssen Sie zusätzliche Optionen angeben, die im Bereich **Optionen für ausgewählte Aktion** angezeigt werden. Die Schaltflächen **Ausführen** und **Speichern** sind erst verfügbar, wenn alle Transformationsparameter angegeben sind.

Zeitreihendaten vorbereiten: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Speichern der Ausgabedatei	Die Datei mit den konvertierten oder transformierten Zeitreihenvariablen wird immer gespeichert. Der Name dieser Ausgabedatei wird im Feld **Modifizierte Daten** angezeigt. Um die Koeffizienten der Splinekurven zu speichern, die an die Eingabezeitreihe angepasst wurden, markieren Sie das Kontrollkästchen **Parameterschätzer**. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Berücksichtigen von Zeitreihenplots in der Ausgabe	Aktivieren Sie das Kontrollkästchen **Plots für Eingabe- und Ausgabezeitreihen erstellen**, um Plots der transformierten Eingabe- und Ausgabezeitreihen sowie der konvertierten Zeitreihen in der Ausgabe zu berücksichtigen. Plots der Zeitreihen und der Zwischenreihen werden für jede Variable erstellt, die Sie der Rolle **Zeitreihenvariable** zugewiesen haben.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

Erstellen der Beobachtungszeilen für fehlende Zeilen in der Datei

Die meisten Anwendungsroutinen für Zeitreihenanalyse gehen davon aus, dass Zeitreihenbeobachtungen aus gleichmäßigen Zeitintervallen oder Zeiträumen stammen. Wenn Zeitintervalle fehlen, sind die Daten nicht gleichmäßig verteilt und die Analyse schlägt eventuell fehl oder erzeugt ungültige Ergebnisse. Folgende Daten sind nicht gleichmäßig verteilt:

Zeile	Datum	Value
1	JAN10	100
2	FEB10	110
3	MAR10	120
4	MAY10	140
5	JUN10	150

Diese Daten sind nicht gleichmäßig verteilt, da das Intervall zwischen Beobachtung 3 und 4 nicht mit dem Intervall zwischen Beobachtung 2 und 3 übereinstimmt.

So erstellen Sie Beobachtungen für die fehlenden Intervalle:

1. Weisen Sie die Datumsvariable der Rolle **Zeit-ID-Variable** zu.
2. Setzen Sie das Intervall auf die Häufigkeit der Daten. Im Beispiel sind die Daten **monatlich**.
3. Klicken Sie auf **Ausführen**.

Eine neue Datei wird erstellt, die alle Spalten in der Originaldatei enthält. Neue Beobachtungen für die ausgelassenen Intervalle werden mit fehlenden Werten hinzugefügt. Diese Daten können nun in Anwendungsroutinen verwendet werden, die fehlende Werte verarbeiten. Sie können jedoch auch die Anwendungsroutine Zeitreihendaten vorbereiten für die neue Datei verwenden, um die fehlenden Werte zu interpolieren. Weitere Informationen, siehe “Interpolieren fehlender Werte” auf Seite 513.

Sie können die fehlenden Werte auch interpolieren, um Schätzer der fehlenden Beobachtungen zu liefern. Weitere Informationen, siehe “Konvertieren aperiodischer Reihen in periodische Schätzer” auf Seite 513.

Konvertieren von Zeitreihendaten von einer Stichprobenhäufigkeit in eine andere

Mit der Anwendungsroutine Zeitreihendaten vorbereiten können Zeitreihendaten von einer Stichprobenhäufigkeit in eine andere konvertiert werden. Beispielsweise können wöchentliche Daten in monatliche Daten aggregiert werden, oder wöchentliche Daten können aus monatlichen Daten interpoliert werden.

So konvertieren Sie die Häufigkeit der Zeitreihendaten:

1. Weisen Sie die Datumsvariable der Rolle **Zeit-ID-Variable** zu. Im Bereich **Best. Zeit-ID-Variable verwenden** setzen Sie das Intervall auf die Häufigkeit der Eingabedaten.
2. Weisen Sie die zu konvertierenden Zeitreihenvariablen der Rolle **Zeitreihenvariablen** zu.
3. Klicken Sie im Auswahlfenster unter der Überschrift ‘Merkmale’ auf **Häufigkeiten** und wählen Sie entweder **Verhältnis** oder **Intervall** im Bereich **Ausgabehäufigkeit**. Wählen Sie z.B. **Intervall** aus und setzen Sie den Intervallwert auf **Wöchentlich**, um die Eingabedaten in eine wöchentliche Häufigkeit zu konvertieren.

4 Klicken Sie im Auswahlfenster unter der Überschrift 'Merkmale' auf **Optionen**, und ändern Sie die Interpolationsmethode in eine Methode, die für die Daten geeignet ist. Wählen Sie z.B. **Kubischer Spline** aus, um eine Splinekurve dritter Ordnung an die Eingabedaten anzupassen.

5 Wählen Sie die Merkmale der Eingabedatenbeobachtung und die gewünschten Merkmale der Ausgabedatenbeobachtung. Beispielsweise können Sie Daten für Monatsanfang in Quartalsmittelwerte konvertieren.

6 Wählen Sie die gewünschte Ausgabedatenausrichtung aus.

7 Klicken Sie auf **Ausführen**.

Ändern der Beobachtungsmerkmale der Eingabedaten

Mit der Anwendungsroutine Zeitreihendaten vorbereiten können die Beobachtungsmerkmale der Eingabedaten geändert werden. Beispielsweise können Sie Daten für monatliche Summen in Monatsmittelwerte konvertieren.

So ändern Sie die Beobachtungsmerkmale der Zeitreihendaten:

1 Weisen Sie die Datumsvariable der Rolle **Zeit-ID-Variable** zu. Im Bereich **Best. Zeit-ID-Variable verwenden** setzen Sie das Intervall auf die Häufigkeit der Eingabedaten.

2 Weisen Sie die zu konvertierenden Zeitreihenvariablen der Rolle **Zeitreihenvariablen** zu.

3 Klicken Sie im Auswahlfenster unter der Überschrift 'Merkmale' auf **Häufigkeiten**. Geben Sie im Bereich **Ausgabehäufigkeit** die gleichen Werte für Ausgabe- und Eingabehäufigkeit ein, es sei denn, Sie ändern auch die Häufigkeit der Daten.

4 Klicken Sie im Auswahlfenster unter der Überschrift 'Merkmale' auf **Optionen**, und ändern Sie die Interpolationsmethode in eine Methode, die für die Daten geeignet ist. Wählen Sie z.B. **Kubischer Spline** aus, um eine Splinekurve dritter Ordnung an die Eingabedaten anzupassen.

5 Wählen Sie die Merkmale der Eingabedatenbeobachtung und die gewünschten Merkmale der Ausgabedatenbeobachtung. Beispielsweise können Sie Daten für Monatsanfang in Quartalsmittelwerte konvertieren.

6 Wählen Sie die gewünschte Ausgabedatenausrichtung aus.

7 Klicken Sie auf **Ausführen**.

Konvertieren aperiodischer Reihen in periodische Schätzer

Die meisten Anwendungsroutinen für Zeitreihenanalyse gehen davon aus, dass Zeitreihenbeobachtungen aus gleichmäßigen Zeitintervallen oder Zeiträumen stammen und dass keine Zeitintervalle fehlen. Daher müssen aperiodische Reihen, wie Transaktionsdaten, vor dem Analysieren mit anderen Zeitreihenanwendungsroutinen in periodische Daten konvertiert werden.

So konvertieren Sie aperiodische Daten in periodische Schätzer:

1. Weisen Sie die Datumsvariable der Rolle **Zeit-ID-Variable** zu. Im Bereich **Best. Zeit-ID-Variable verwenden** wählen Sie die Option **Aperiodische Daten** aus.
2. Weisen Sie die zu konvertierenden Zeitreihenvariablen der Rolle **Zeitreihenvariablen** zu.
3. Klicken Sie im Auswahlfenster unter der Überschrift **Merkmale** auf **Häufigkeiten** und wählen Sie das gewünschten Ausgabeintervall aus.
4. Klicken Sie im Auswahlfenster unter der Überschrift **Merkmale** auf **Optionen**, und ändern Sie die Interpolationsmethode in eine Methode, die für die Daten geeignet ist. Wählen Sie z.B. **Kubischer Spline** aus, um eine Splinekurve dritter Ordnung an die Eingabedaten anzupassen.
5. Wählen Sie die gewünschte Ausgabedatenausrichtung aus.
6. Wählen Sie die Merkmale der Eingabedatenbeobachtung und die gewünschten Merkmale der Ausgabedatenbeobachtung. Beispielsweise können Sie aperiodische Werte eines bestimmten Zeitpunkts in monatliche Mittelwerte konvertieren.
7. Klicken Sie auf **Ausführen**.

Interpolieren fehlender Werte

Mit der Anwendungsroutine Zeitreihendaten vorbereiten können fehlende Werte in den Eingabezeitreihen interpoliert werden. Diese Daten enthalten fehlende Werte:

Zeile	Datum	Value

1	JAN10	100
2	FEB10	110
3	MAR10	120
4	APR10	.
5	MAY10	140
6	JUN10	150
7	JUL10	.
8	AUG10	150

So interpolieren Sie die fehlenden Werte in den Eingabedaten:

1. Weisen Sie die Datumsvariable der Rolle **Zeit-ID-Variable** zu. Im Bereich **Best. Zeit-ID-Variable verwenden** setzen Sie das Intervall auf die Häufigkeit der Eingabedaten.
2. Weisen Sie die zu konvertierenden Zeitreihenvariablen der Rolle **Zeitreihenvariablen** zu.
3. Klicken Sie im Auswahlfenster unter der Überschrift 'Merkmale' auf **Häufigkeiten** und wählen Sie **Gleich wie Eingabehäufigkeit** als Ausgabereihenhäufigkeit.
4. Klicken Sie im Auswahlfenster unter der Überschrift **Merkmale** auf **Optionen**, und ändern Sie die Interpolationsmethode in eine Methode, die für die Daten geeignet ist. Wählen Sie z.B. **Kubischer Spline** aus, um eine Splinekurve dritter Ordnung an die Eingabedaten anzupassen.
5. Wählen Sie die Merkmale der Eingabedatenbeobachtung und die gewünschten Merkmale der Ausgabedatenbeobachtung. Beispielsweise können Sie Daten für Tagesende in Monatsmittelwerte konvertieren.
6. Wählen Sie die gewünschte Ausgabedatenausrichtung aus.
7. Klicken Sie auf **Ausführen**.

Anwenden generischer Transformationen auf die Daten

Mit der Anwendungsroutine Zeitreihendaten vorbereiten können Sie eine Vielzahl von mathematischen Transformationen für die Eingabezeitreihen durchführen. Diese Transformationen können in Verbindung mit den verfügbaren Interpolationsoperationen durchgeführt werden, oder ohne Werte zu interpolieren.

Um Daten in Verbindung mit einer Interpolationsoperation zu transformieren, führen Sie alle Schritte für die Interpolationsoption außer dem Abschicken der Anwendungsroutine aus. (Weitere Informationen, siehe “Interpolieren fehlender Werte” auf Seite 513.) Definieren Sie anschließend auf den Seiten **Transformationen** die Transformationen, die Sie für die Daten vor und nach der Interpolation durchführen möchten. Beide Listen können leer sein.

So transformieren Sie Daten, ohne eine Interpolation durchzuführen:

1. Weisen Sie die Datumsvariable der Rolle **Zeit-ID-Variable** zu. Im Bereich **Best. Zeit-ID-Variable verwenden** setzen Sie das Intervall auf die Häufigkeit der Eingabedaten.
2. Weisen Sie die zu konvertierenden Zeitreihenvariablen der Rolle **Zeitreihenvariablen** zu.
3. Klicken Sie im Auswahlfenster unter der Überschrift **Merkmale** auf **Optionen** und wählen Sie **Keine Interpolation** als Interpolationsmethode.
4. Wählen Sie die gewünschte Ausgabedatenausrichtung aus.
5. Klicken Sie im Auswahlfenster unter der Überschrift **Transformationen** auf **Vor Interpolation**, um die Transformationen einzufügen, die Sie ausführen möchten. Obwohl die Liste **Nach Interpolation** im Allgemeinen leer gelassen wird, können Sie auch auf der Seite **Nach Interpolation** Transformationen definieren. Diese Transformationen erfolgen nach Durchführung der Transformationen in der Liste **Vor Interpolation**.
6. Klicken Sie auf **Ausführen**.

65

Hauptkomponenten

Allgemeines über die Anwendungsroutine 'Hauptkomponenten' ... 517
Hauptkomponenten: Zuweisen von Variablen zu Analyserollen ... 518
Hauptkomponenten: Festlegen der Analyseoptionen ... 519
Hauptkomponenten: Erzeugen von Plots ... 520
Hauptkomponenten: Festlegen von Ergebnisoptionen ... 521
Festlegen von Titeln und Fußnoten ... 522
Anzeigen von Eigenschaften ... 523

Allgemeines über die Anwendungsroutine 'Hauptkomponenten'

Die Hauptkomponentenanalyse ist eine multivariate Technik zur Untersuchung von Beziehungen zwischen verschiedenen quantitativen Variablen. Sie sollten die Hauptkomponentenanalyse einsetzen, wenn Sie Daten zusammenfassen und lineare Beziehungen erkennen möchten.

Sie können die Anwendungsroutine **Hauptkomponenten** verwenden, wenn mehr Variablen vorliegen, als gleichzeitig gedruckt werden können. Angenommen, Ihnen liegen die Verbrechensraten pro 100.000 Menschen in 7 Kategorien für alle 50 amerikanischen Bundesstaaten im Jahr 1977 vor. Da es sieben numerische Variablen gibt, können nicht alle Variablen gleichzeitig dargestellt werden. In diesem Fall können Sie mithilfe der Hauptkomponentenanalyse die Daten in zwei oder drei Dimensionen zusammenfassen und so leichter darstellen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	PRINCOMP
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Hauptkomponenten: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Analysevariablen	führt eine Analyse für jede Variable in dieser Liste durch. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Partielle Variablen	gibt die Variablen an, die beim Analysieren einer partiellen Korrelation oder Kovarianzmatrix herausgefiltert werden sollen. Die Anwendungsroutine **Hauptkomponenten** berechnet die Hauptkomponenten der Residuen aus der Vorhersage der Analysevariablen nach den Variablen in dieser Rolle.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Sie können Analysen nicht nach einer Variable gruppieren, die Sie als Analysevariable ausgewählt haben.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie

Optionsname	Beschreibung
	n Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt die Werte an, die zum Gewichten der entsprechenden Werte in der Analysevariable verwendet werden sollen. Sie können dieser Rolle maximal eine Variable zuweisen.

Hauptkomponenten: Festlegen der Analyseoptionen

Klicken Sie im Auswahlfenster auf **Analyse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Analysieren	Wählen Sie die auszuführende Analyse aus: ■ **Korrelationen** berechnet die Hauptkomponenten aus der Korrelationsmatrix. ■ **Kovarianzen** berechnet die Hauptkomponenten aus der Kovarianzmatrix. Wenn Sie diese Option angeben, werden die Variablen mit großen Varianzen eher Komponenten mit großen Eigenwerten zugewiesen, während Variablen mit kleinen Varianzen eher Komponenten mit kleinen Eigenwerten zugewiesen werden. Sie sollten diese Option nicht angeben, wenn die Einheiten, in denen die Variablen gemessen werden, vergleichbar sind oder wenn die Variablen in irgendeiner Weise standardisiert wurden. ■ **Unkorrigierte Korrelationen** berechnet die Hauptkomponenten aus der Korrelationsmatrix. Allerdings wird die Konstante im Modell weggelassen, sodass die Korrelationsmatrix und die Standardabweichungen für den Mittelwert eventuell nicht korrigiert werden können. ■ **Unkorrigierte Kovarianzen** berechnet die Hauptkomponenten aus der Kovarianzmatrix. Allerdings wird die Konstante im Modell weggelassen, sodass die Kovarianzmatrix und die Standardabweichungen für den Mittelwert eventuell nicht korrigiert werden können.

Optionsname	Beschreibung
Zu berechnende Hauptkomponenten	Standardeinstellung ist die Anzahl der Variablen. Klicken Sie auf die nach oben bzw. unten weisenden Pfeile, um diese Zahl zu ändern.
Präfix für die Benennung von Hauptkomponenten	Geben Sie das Präfix für die Hauptkomponenten an. Der Name, den Sie angeben, darf maximal 26 Zeichen lang sein. Wenn Sie kein Präfix angeben, verwendet die Anwendungsroutine den Standardwert. Standardmäßig ist das Präfix für die Benennung der Hauptkomponenten PRIN1, PRIN2, ..., PRIN`n`.
Singularitätskriterium	Geben Sie das Singularitätskriterium an, wobei 0 < p < 1. Wenn eine Variable, die Sie der Rolle **Partielle Variablen** zugewiesen haben, für R^2 einen Wert größer 1-p aufweist (bei der Prognose mit den vor ihr aufgelisteten Variablen in der PARTIAL-Anweisung), wird der Variablen ein standardisierter Koeffizient von 0 zugewiesen. Der Standardwert ist `1E-08`.
Teiler für Varianz	Standardmäßig ist der Teiler für die Varianzberechnung der Freiheitsgrad. Sie können einen anderen Teiler verwenden, indem Sie die gewünschte Methode aus der Dropdown-Liste **Teiler für Varianz** auswählen. **Hinweis:** Für die partielle Extraktion müssen Sie der Rolle **Partielle Variablen** eine Variable zuweisen.

Hauptkomponenten: Erzeugen von Plots

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Scree- und Varianz-Plots erstellen	Erstellt einen Scree-Plot der Eigenwerte und einen Plot der erklärten Verhältnisvarianz. Sie können mit diesem Scree-Plot festlegen, wie viele Komponenten in der Analyse verwendet werden sollen. Die Eigenwerte entsprechen den einzelnen Hauptkomponenten und stellen eine Partitionierung der gesamten Variation in der Stichprobe dar. Die Eigenwerte werden vom größten zum kleinsten Wert sortiert, wobei die ersten Werte den größten Teil der Variation ausmachen. Der Scree-Plot zeigt im Allgemeinen bei den ersten Eigenwerten einen starken Abfall, der sich bei den übrigen Eigenwerten abschwächt. Sie können

Optionsname	Beschreibung
	mit der Anzahl der Eigenwerte in diesem Übergangsbereich die passende Anzahl der einzuschließenden Komponenten festlegen.
Streuungsplotmatrix der Hauptkomponenten-Scorewerte erstellen	erstellt eine Streuungsdiagramm-Matrix der Hauptkomponenten-Scorewerte. Das Histogramm jeder Komponente wird im diagonalen Element der Matrix angezeigt.
Musterprofil-Plot erstellen	erstellt einen Muster-Profilplot. Es gibt ein Profil für jede Komponente. Der Wert auf der Y-Achse ist die Korrelation zwischen der Variable und der Hauptkomponente.
Streuungsplot der Hauptkomponenten-Scorewerte erstellen	erstellt ein Streuungsdiagramm für jeden Hauptkomponenten-Scorewert. Sie können diese Plots anpassen, indem Sie entscheiden, die Vorhersageellipsen für die Hauptkomponenten-Scorewerte einer neuen Beobachtung anzuzeigen. Standardmäßig werden keine Ellipsen angezeigt.
Musterkomponentenplots erstellen	erstellt paarweise Komponentenmuster-Plots. Jede Beobachtung im Plot stellt die Korrelationen zwischen der Variablen und den beiden entsprechenden Komponenten dar. Sie können die Muster auch als Vektoren darstellen. Standardmäßig wird dann ein Kreis mit einem 100%-Varianzkreis für den Vektormusterplot gezeichnet.

Hauptkomponenten: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Tabelle für Originaldaten und Komponenten-Scorewerte erstellen	Mit der Option **Tabelle für Originaldaten und Hauptkomponenten-Scorewerte erstellen** erstellen Sie eine Tabelle, in der Ihre Einlesedaten enthalten sind sowie die Scorewerte der Hauptkomponenten. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und

Optionsname	Beschreibung
	speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Tabelle für Statistiken erstellen	Mit der Option **Tabelle für Statistiken erstellen** erstellen Sie eine Tabelle, in der Mittelwerte, Standardabweichungen, Anzahl der Beobachtungen, Korrelationen oder Kovarianzen, Eigenwerte und Eigenvektoren enthalten sind.
Hauptkomponenten-Scorewerte standardisieren	Aktivieren Sie das Kontrollkästchen **Hauptkomponenten-Scorewerte standardisieren**, wenn Sie die Scorewerte der Hauptkomponenten auf gleiche Varianz standardisieren möchten. Diese Option hat Einfluss auf die Tabelle, die Sie im Feld **Tabelle für Originaldaten und Hauptkomponenten-Scorewerte erstellen** angegeben haben.
Bericht und Plots unterdrücken	Bewirkt, dass keine Ergebnisse angezeigt werden.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

66

Wahrscheinlichkeitsplots

Allgemeines über die Anwendungsroutine 'Wahrscheinlichkeitsplots' ... **525**

Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen ... **526**

Wahrscheinlichkeitsplots: Auswählen der Verteilung ... **528**

Wahrscheinlichkeitsplots: Anpassen der Verteilung ... **529**

Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung ... **530**

Festlegen von Achsenoptionen ... 530
Hinzufügen von Referenzlinien ... 530
Einfügungen hinzufügen ... 530
Festlegen weiterer Plotoptionen ... 531
Festlegen von Optionen für Spezifikationsgrenzen ... 533

Prozessfähigkeitsanalyse: Festlegen der Analysetypen ... **534**

Festlegen von Titeln und Fußnoten ... **535**

Anzeigen von Eigenschaften ... **535**

Allgemeines über die Anwendungsroutine 'Wahrscheinlichkeitsplots'

Eine Prozessfähigkeitsanalyse vergleicht die Verteilung der Ausgabe eines Prozesses in statistischer Kontrolle mit den Spezifikationsgrenzen, um die Konsistenz zu bestimmen, mit der die Spezifikationen erfüllt werden. Ein Wahrscheinlichkeitsplot wird bei der Prozessfähigkeitsanalyse verwendet, um geordnete Werte einer Variablen mit den Perzentilen einer angegebenen theoretischen Verteilung, wie der Normalverteilung, zu vergleichen. Wenn die Datenverteilung der theoretischen Verteilung entspricht, bilden die Punkte im Plot ein lineares Muster. Daher können Sie mit einem Wahrscheinlichkeitsplot ermitteln, wie gut eine theoretische Verteilung einen Satz an Messwerten modelliert.

Hinweis: Wahrscheinlichkeitsplots ähneln Q-Q-Plots. Wahrscheinlichkeitsplots sind bei der grafischen Schätzung von Perzentilen vorzuziehen, wohingegen Q-Q-Plots bei der grafischen Schätzung von Verteilungsparametern und Prozessfähigkeitsindizes geeigneter sind.

Sie können ein Wahrscheinlichkeitsplot auch zum Analysieren eines Fertigungsprozesses verwenden. Angenommen, die Durchmesser von 50 Stahlstangen werden gemessen und in einer Datei gespeichert. Der Prozess zur Fertigung der Stangen befindet sich innerhalb der statistischen Kontrolle, und Sie möchten überprüfen, ob die Durchmesser normal verteilt sind. Die Linearität des Punktmusters weist darauf hin, ob die Messungen eine Normalverteilung aufweisen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	CAPABILITY
Erforderliche SAS-Produkte	Base SAS, SAS/QC, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Analysevariablen	gibt die zu analysierenden numerischen Variablen an. Sie müssen dieser Rolle mindestens eine Variable zuweisen.

Optionsname	Beschreibung
	Bei allen Plots, außer dem P-P-Plot, können Sie eine Obergrenze, einen Zielwert und eine Untergrenze im Bereich **Spezifikationsgrenzen für *Variablenname*** für jede Analysevariable angeben. Wenn Sie diese Grenzen angeben, enthält die Ausgabe automatisch eine Tabelle mit den Tests auf Normalverteilung, eine Tabelle mit den Spezifikationsgrenzen und eine Tabelle mit den Prozessfähigkeitsindizes.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt die Variable an, die als relative Gewichtungsvariable verwendet werden soll. Wenn Sie dieser Rolle eine Variable zuweisen, werden mit dem Wert der Variablen in jeder Beobachtung gewichtete Statistiken berechnet. Sie können dieser Rolle nur eine Variable zuweisen. **Hinweis:** Sie sollten der Rolle **Relative Gewichtung** keine Variable zuweisen, wenn Sie ein Histogramm, Wahrscheinlichkeitsplot, P-P-Plot, Q-Q-Plot oder CDF-Plot erstellen. Falls Sie dieser Rolle ein Variable zugewiesen haben, müssen Sie eine eventuelle Ausgabe mit Vorsicht interpretieren.
Klassifizierungsvariablen (nur verfügbar für die Anwendungsroutine Histogramm)	gibt eine oder zwei Variablen an, die als Klassifizierungsvariablen dienen. Wenn Sie dieser Rolle eine Variable zuweisen, erstellt die Anwendungsroutine ein Vergleichshistogramm, mit dem Sie die Verteilung einer Analysevariablen über Ausprägungen der Klassifizierungsvariablen hinweg vergleichen können. **Hinweis:** Um vergleichende Wahrscheinlichkeitsplots oder vergleichende Q-Q-Plots zu erstellen, verwenden Sie die Anwendungsroutine Verteilungsanalyse.

Wahrscheinlichkeitsplots: Auswählen der Verteilung

Klicken Sie im Auswahlfenster unter der Überschrift 'Verteilungen' auf **Zusammenfassung**, um auf diese Optionen zuzugreifen.

Sie können eine theoretische Verteilung auf einem Wahrscheinlichkeitsplot anpassen. Wenn Sie die Verteilung ausgewählt haben, wählen Sie den Namen der Verteilung im Auswahlfenster aus, um die Parameterwerte und Linieneigenschaften für diese Verteilung anzugeben.

Optionsname	Beschreibung
Verteilungen	Verwenden Sie die Kontrollkästchen, um eine oder mehrere der folgenden geschätzten Dichtekurven auszuwählen: ■ **Normal** passt eine normale Dichte mithilfe eines Mittelwerts (My) und einer Standardabweichung (Sigma) an. ■ **Lognormal** passt eine Lognormal-Dichte mithilfe eines Skalierungsparameters (Zeta), eines Schwellenwertparameters (Theta) und eines Formparameters (Sigma) an. ■ **Exponentiell** passt eine exponentielle Dichte mithilfe eines Skalierungsparameters (Sigma) und eines Schwellenwertparameters (Theta) an. ■ **Beta** passt eine Beta-Dichte mit einem Skalierungsparameter (Sigma), einem Schwellenwertparameter (Theta) und Formparametern (Alpha und Beta) an. ■ **Gamma** passt eine Gamma-Dichte mit einem Skalierungsparameter (Sigma), einen Schwellenwertparameter (Theta) und einen Formparameter (Alpha) an. ■ **Weibull (3-Parameter)** passt eine 3-Parameter-Weibull-Dichte mit einem Skalierungsparameter (sigma), einem Schwellenwertparameter (theta) und einem Formparameter (c) an. ■ **Weibull (2-Parameter)** passt eine 2-Parameter-Weibull-Dichte mit einem Skalierungsparameter (sigma) und einem Formparameter (c) an. Bei Weibull (2-Parameter) wird angenommen, dass die Daten einen bekannten unteren Schwellenwertparameter, theta=0, haben.

Optionsname	Beschreibung
Grafikstil	Wählen Sie den Grafikstil, der in der Ausgabe verwendet werden soll. Sie haben die Auswahl zwischen üblichen SAS-Grafiken und ODS Statistical Graphs (statistischen ODS-Grafiken). Weitere Informationen zu ODS-Grafiken finden Sie unter *SAS Output Delivery System: User's Guide*.

Wahrscheinlichkeitsplots: Anpassen der Verteilung

Klicken Sie im Auswahlfenster unter der Überschrift **Verteilungen** auf den Namen der Verteilung, die Sie anpassen möchten, um auf diese Optionen zuzugreifen.

Hinweis: Diese Verteilungsoptionen sind nur durch vorherige Auswahl der Verteilung auf der Seite **Zusammenfassung** verfügbar, oder durch Aktivieren des Kontrollkästchens ***Name der Verteilung*** auf der Seite dieser Verteilung.

Für die Normal-, Lognormal-, Exponential-, Weibull-, Beta- und Gamma-Verteilungen können Sie die Parameter für jede Analysevariable oder für alle Analysevariablen angeben. Standardmäßig schätzt die Anwendungsroutine die Parameterwerte für alle Analysevariablen.

So legen Sie für jede Analysevariable Parameter fest:

1. Deaktivieren Sie das Kontrollkästchen **Verteilung auf alle Variablen anwenden**.
2. Aktivieren Sie das Kontrollkästchen der Variablen im Feld **Analysevariablen**.
3. Geben Sie die Parameterwerte an.

Hinweis: Um zu den geschätzten Parameterwerten zurückzukehren, klicken Sie auf **Schätzer verwenden**.

Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung

Festlegen von Achsenoptionen

Klicken Sie im Auswahlfenster unter **Darstellung** ⇨ **Achsen** auf **Achsen**, um auf die Optionen zuzugreifen.

Folgende Optionen können angegeben werden:

- die Farbe der Achsen
- die Farbe der Achsenmarkierungen
- ein Etikett für die horizontalen und vertikalen Achsen

Hinzufügen von Referenzlinien

Klicken Sie im Auswahlfenster unter **Darstellung** ⇨ **Achsen** auf **Horizontal** bzw. **Vertikal**, um auf die Optionen zuzugreifen.

Um Referenzlinien zu verwenden, wählen Sie **Referenzlinien verwenden**. Sie können die Referenzlinie formatieren und die Position jeder Referenzlinie festlegen.

Einfügungen hinzufügen

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Einfügung**, um auf diese Optionen zuzugreifen.

Sie können einen Plot erweitern, indem Sie direkt in der Grafik ein Feld oder eine Tabelle (eine so genannte Einfügung) der beschreibenden Statistiken hinzufügen. Wenn Sie ein Vergleichshistogramm erstellen, erscheint in jedem Komponentenhistogramm eine Einfügung.

Wählen Sie **Einfügung einschließen**, um eine beschreibende Statistik zum Plot hinzuzufügen. Wählen Sie dann die Statistiken, die Sie in die Einfügung einschließen möchten.

Hinweis: Die Statistiken für die Berechnung der Prozessfähigkeitsindizes sind nur verfügbar, wenn Sie Spezifikationsgrenzen festgelegt und dann der Rolle

Analysevariablen zugewiesen haben. Sie können Spezifikationsgrenzen für Histogramm, CDF-Plots, Wahrscheinlichkeitsplots und Q-Q-Plots erstellen. Weitere Informationen zur Berechnung der Prozessfähigkeitsindizes finden Sie in der Hilfe zur Prozedur CAPABILITY.

Festlegen weiterer Plotoptionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Optionen**, um auf diese Optionen zuzugreifen.

Die verfügbaren Plotdarstellungsoptionen hängen vom Typ des Plots ab, den die Anwendungsroutine erzeugt.

Optionsname	Beschreibung
Hintergrund	bestimmt die Hintergrundfarbe des Plots.
Balkenumriss und **Balkenfüllung**	legt die Farben der Balken in einem Histogramm fest. **Hinweis:** Diese Optionen sind nur für Vergleichshistogramme verfügbar.
Beide Achsen umkehren (Perzentile auf entgegengesetzte Achsen)	kehrt die Achsen auf dem Plot um. Wenn Sie diese Option auswählen, funktionieren die von Ihnen ausgewählten Achsenoptionen und Referenzlinien eventuell nicht mehr, da sich die Ausrichtung des Plots ändert. **Hinweis:** Diese Option ist nur für Wahrscheinlichkeitsplots und Q-Q-Plots verfügbar.
Keine Standardlegenden	verhindert die Anzeige der Standardlegende mit dem Plot.
Midpoints verwenden	geben Sie Midpoints für das Histogramm an. Die in der Dropdown-Liste verfügbaren Optionen hängen davon ab, ob Sie ein Histogramm oder ein Vergleichshistogramm erstellen. Wenn Sie der Rolle **Klassifizierungsvariablen** eine Variable zuweisen, erstellen Sie ein Vergleichshistogramm. Für Vergleichshistogramme stehen folgende Optionen zur Verfügung: ■ **Midpoints angeben** - Ermöglicht das Angeben der Start- und End-Midpoints. Sie können auch das Schrittintervall festlegen, das für die Berechnung der anderen Midpoints verwendet werden soll. Der Midpoint ist der Mittelwert des Wertebereichs, der von den einzelnen Balken dargestellt wird. ■ **Schlüssel** - Bestimmt die Midpoints für die Daten in der Schlüsselzelle. Standardmäßig werden die Ausprägungen

Optionsname	Beschreibung
	der Klassifizierungsvariablen von oben nach unten (links nach rechts) in aufsteigender Reihenfolge der internen (nicht formatierten) Werte der ersten Klassifizierungsvariablen angezeigt. Wenn Sie nur eine Klassifizierungsvariable angegeben haben, ist die Schlüsselzelle die Ausprägung, die als erstes in dieser Reihenfolge erscheint. Wenn Sie zwei Klassifizierungsvariablen angegeben haben, ist die Schlüsselzelle die Kombination der Ausprägungen von Variable 1 und Variable 2, die als erstes in dieser Reihenfolge erscheint. Die Auswahl der Schlüsselzelle bestimmt die einheitliche Horizontalachse, die für alle Zellen verwendet wird. Die Midpoint-Liste für die Schlüsselzelle wird dann wie erforderlich in beide Richtungen erweitert, bis sie die Daten in den übrigen Zellen abdeckt. ■ **Einheitlich** - Bestimmt die Anzahl der Midpoints abhängig von der Gesamt-Stichprobengröße. Wenn Sie der Rolle **Klassifizierungsvariablen** keine Variable zuweisen, erstellen Sie ein Histogramm. Für Histogramme sind die folgenden Optionen verfügbar: ■ **Midpoints angeben** - Ermöglicht das Angeben der Start- und End-Midpoints. Sie können auch das Schrittintervall festlegen, das für die Berechnung der anderen Midpoints verwendet werden soll. Der Midpoint ist der Mittelwert des Wertebereichs, der von den einzelnen Balken dargestellt wird. ■ **Anzahl der Intervalle angeben** - Ermöglicht, die Anzahl der Intervalle (auch Histogrammintervalle genannt) für die Daten anzugeben. Standardmäßig basiert die Anzahl der angezeigten Intervalle auf dem Datenbereich. Sie können jedoch eine Maximalzahl anzuzeigender Intervalle angeben. **Hinweis:** Sie können nur dann Midpoints angeben, wenn Sie die Anwendungsroutine **Histogramme** verwenden.
Größe ben.def. Diagramm angeben (in Pixel)	legt die Bildgröße fest. Zum Vergleich wird die Gesamtbildschirmgröße angezeigt in den Feldern **Breite** und **Höhe** angezeigt.

Hinweis: Vergleichsoptionen sind für Vergleichshistogramme verfügbar. Sie können ein Vergleichshistogramm erstellen, indem Sie in der Anwendungsroutine **Histogramme** der Rolle **Klassifizierungsvariablen** eine Variable zuweisen.

Optionsname	Beschreibung
Zeilen und **Spalten**	Wenn Sie einer Analyserolle mindestens eine Klassifizierungsvariable zuweisen, können Sie über die Felder **Zeilen pro Seite** und **Spalten pro Seite** die Anordnung der Spalten und Zeilen im Vergleichshistogramm ändern. Wenn Sie eine Klassifizierungsvariable zuweisen, ist die Standardanordnung zwei Zeilen und eine Spalte pro Seite. Wenn Sie zwei Klassifizierungsvariablen zuweisen, ist die Standardanordnung zwei Spalten und zwei Zeilen pro Seite. Sie können auch über die Dropdown-Listen **Rahmenseite** und **Rahmenoberseite** Farben zum Füllen des Rahmenbereichs der Zeilenetiketten und Spaltenetiketten angeben.
Anzahl der Intervalle	Standardmäßig basiert die Anzahl der angezeigten Intervalle auf dem Datenbereich. Sie können jedoch eine Maximalzahl anzuzeigender Intervalle angeben. Sie können auch eine Standardabweichung angeben, so dass die Anzahl der angezeigten Intervalle von einem Bereich des Standardabweichungswerts über und unter dem Mittelwert der Daten abhängt.

Festlegen von Optionen für Spezifikationsgrenzen

Hinweis: Diese Optionen sind für Histogramme, Wahrscheinlichkeitsplot, Q-Q-Plots und CDF-Plots verfügbar.

Klicken Sie im Auswahlfenster unter **Darstellung** auf **Spezifikationsgrenzen**, um die untere Spezifikationsgrenze, das Ziel und die obere Spezifikationsgrenze anzugeben.

Hinweis: Die Optionen für **Spezifikationsgrenzen** sind nur verfügbar, wenn Sie eine oder mehrere Grenzen für mindestens eine Analysevariable angeben. Um eine Spezifikationsgrenze für eine Analysevariable anzugeben, klicken Sie im Auswahlfenster auf **Daten** und wählen dann die Analysevariable aus.

Prozessfähigkeitsanalyse: Festlegen der Analysetypen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

Standardmäßig enthält der Prozessfähigkeitsanalysebericht die folgenden Tabellen: eine mit Grundkonfidenzintervallen, eine mit Grundmaßen, eine mit Tests auf Lageparametern und eine mit Momenten.

Um eine Tabelle zum Bericht hinzuzufügen, wählen Sie die Tabelle im Feld **Tabellen** aus.

Um die Analyseoptionen festzulegen, wählen Sie die Tabelle im Feld **Tabellen** aus. Sind Optionen verfügbar, erscheint rechts ein Feld, in dem Sie die Optionen angeben können.

Um alle Tabellen mit beschreibenden Statistiken im Bericht zu unterdrücken, aktivieren Sie die Option **Beschreibende Statistiken und Tabellen für Prozessfähigkeitsindizes unterdrücken**. Folgende Tabellen werden durch Auswahl dieser Option nicht unterdrückt:

- Tabellen, die von der ausgewählten Verteilung erstellt wurden. Sie können diese Tabellen unterdrücken, indem Sie im Fenster für diese Verteilung das Kontrollkästchen **Verteilungstabellen unterdrücken** aktivieren.
- Tabellen, die mit der INTERVALS-Anweisung erstellt wurden. Um diese Tabellen zu unterdrücken, können Sie die Option NOPRINT angeben. Weitere Informationen zur Anweisung INTERVALS finden Sie in der Hilfe zur Prozedur CAPABILITY.

Um Ihre Ergebnisse in eine Ausgabedatei zu speichern, aktivieren Sie das Kontrollkästchen **Ausgabestatistiken in einer Datei speichern**. Die Datei enthält univariate Statistiken und Prozessfähigkeitsindizes.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

67

Proportionale Hazards

Allgemeines über die Anwendungsroutine 'Proportionale Hazards' **537**
Proportionale Hazards: Zuweisen von Variablen zu Analyserollen **538**
Proportionale Hazards: Festlegen von Modelloptionen **540**
Proportionale Hazards: Optionen für das Festlegen der Methoden **542**
Proportionale Hazards: Erzeugen von Plots **543**
Proportionale Hazards: Festlegen von Ergebnisoptionen **544**
Festlegen von Titeln und Fußnoten **545**
Anzeigen von Eigenschaften **546**

Allgemeines über die Anwendungsroutine 'Proportionale Hazards'

Die Anwendungsroutine **Proportionale Hazards** führt basierend auf Cox' Modell für proportionale Hazards Regressionsanalysen von Überlebensdaten durch. Dieses semiparametrische Modell findet breite Anwendung bei der Analyse von Überlebensdaten, um die Effekte erklärender Variablen in Überlebenszeiten zu erklären.

Die Überlebenszeitanalyse wird vor allem in klinischen und epidemiologischen Nachuntersuchungen verwendet. Doch auch auf anderen Gebieten, wie Soziologie, Maschinenbau und Ökonomie, werden die Methoden zur Überlebenszeitanalyse eingesetzt. Unabhängig vom Einsatzgebiet wird bei allen Überlebenszeitanalysestudien bestimmt, ob und wann ein Ereignis aufgetreten ist. Beispielsweise unterscheiden sich Personen, die fünf Jahre nach einer Operation sterben, von Personen, die einen Monat nach einer Operation sterben. Eine Analyse, die einfach die Todesfälle nach einer Operation zählt, würde wertvolle Informationen über die Überlebenszeit ignorieren. Die Überlebenszeitanalyse kann

auch zum Analysieren anderer, zeitunabhängiger Ergebnisse verwendet werden. So kann z. B. ein Ingenieur die Kilometerzahl bis zum Ausrangieren eines Reifens oder die Anzahl der Zyklen bis zur Reparatur einer Maschine analysieren. Bei allen diesen Studien analysieren Sie ein Ergebnis bis zum Auftreten eines Ereignisses.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	PHREG
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Proportionale Hazards: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Überlebenszeit	gibt die als Überlebenszeitvariable verwendete Variable an. Überlebenszeit ist erforderlich. Sie können dieser Rolle nur eine Variable zuweisen.
Zensurvariable	gibt die Variable an, die als Zensurvariable verwendet wird. Wählen Sie einen Wert im Bereich **Rechts zensierende Werte**, um eine Liste mit mindestens einem Wert zu erstellen, mit dem/denen die Überlebenszeit zensiert wird. Geben Sie in das Feld **Angepasste Werte eingeben** einen bestimmten Wert ein, und klicken Sie auf **Hinzufügen**, um ihn in die Liste der Zensurwerte einzuschließen. Um einen Wert aus der Liste der Zensurwerte zu

Optionsname	Beschreibung
	entfernen, deaktivieren Sie das Kontrollkästchen neben diesem Wert. Die Zensurwerte sollten nicht fehlende, numerische Werte sein. Die Zensurvariable ist nicht erforderlich. Sie können dieser Rolle jedoch nur eine Variable zuweisen.
Erklärende Variablen	gibt die erklärenden Variablen für das Modell an.
Schichtvariablen	gibt die Variablen an, deren Werte die Schichtausprägungen bestimmen. Wenn Sie keine Endpunkte für die numerischen Variablen festlegen, werden die Schichten durch die Kombination der Ausprägungen der eindeutigen Werte der Schichtvariablen ermittelt. Um fehlende Werte als gültige Schichtausprägungen zuzulassen, aktivieren Sie das Kontrollkästchen. Wenn Sie das Kontrollkästchen nicht aktivieren, werden Beobachtungen mit fehlenden Werten aus der Analyse entfernt. Bei numerischen Schichtvariablen können Sie die Schichtintervalle angeben, indem Sie im Bereich **Schichtausprägungen auf Intervalle angeben** klicken. Anschließend definieren Sie die Intervalle, indem Sie Endpunkte hinzufügen. Standardmäßig bilden die eindeutigen Werte Schichtausprägungen. Wenn Sie dieser Rolle mehrere Variablen zuweisen, definieren Sie für jede Variable die Intervalle separat.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
ID-Variablen	gibt mindestens eine Variable an, die als ID-Variable in der Ausgabedatei verwendet werden soll, die die Überlebensschätzer enthält. Um die Ausgabedatei zu erstellen, geben Sie auf der Seite **Ergebnisse** an, dass Sie die Überlebensschätzer speichern möchten.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.

Proportionale Hazards: Festlegen von Modelloptionen

Klicken Sie im Auswahlfenster auf **Modell**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Optionen sind nur verfügbar, wenn Sie der Rolle **Erklärende Variablen** eine Variable zugewiesen haben.

Optionsname	Beschreibung
Festlegen der Modellauswahlmethode	Wählen Sie aus der Dropdown-Liste **Auswahlmethode** eine Methode aus, um festzulegen, welche erklärenden Variablen im Modell erscheinen. Wenn Sie keine Methode auswählen, wird das volle Modell verwendet. **Volles Modell (keine Auswahl)** Diese Methode ist der Standardwert und bietet keine Modellauswahlfunktion. Alle erklärenden Variablen werden in das Modell eingeschlossen. **Aufwärtsverfahren** Das Vorwärtsverfahren schätzt zuerst Parameter für die Variablen, die in das Modell aufgenommen werden. Bei diesen Variablen handelt es sich um die Konstanten und um die ersten *n* erklärenden Variablen, die in der Liste **Einzuschließende Variablen** aufgeführt werden. Der Standard für *n* ist null, sodass keine Variablen im Modell enthalten sind. Danach berechnet dieses Verfahren die angepassten Chi-Quadrat-Statistiken für jede nicht im Modell enthaltene Variable und untersucht die größte dieser Statistiken. Die *p*-Werte für diese Statistiken werden mit dem Signifikanzniveau verglichen, das bei Einschließen einer Variablen in das Modell angegeben wird. Standardmäßig beträgt dieser Wert `0,05`. Um das Signifikanzniveau zu ändern, geben Sie den Wert in das Textfeld **Für Aufnahme ins Modell** ein. Wenn keine angepasste Chi-Quadrat-Statistik einen größeren Wert als diesen aufweist, hält das Vorwärtsverfahren an. Andernfalls fügt das Vorwärtsverfahren die Variable mit der größten Chi-Quadrat-Statistik dem Modell hinzu. Das Vorwärtsverfahren berechnet dann erneut die angepassten Chi-Quadrat-Statistiken für die Variablen, die außerhalb des Modells bleiben, und das Bewertungsverfahren wird fortgesetzt. Die Variablen werden also solange sequentiell dem Modell hinzugefügt, bis keine verbleibende Variable mehr eine signifikante angepasste Chi-Quadrat-Statistik erzeugt. Eine Variable, die zum Modell hinzugefügt wurde, verbleibt im Modell.

Optionsname	Beschreibung
	Rückwärtsverfahren Das Rückwärtsverfahren schätzt zuerst Parameter für das ganze Modell und eliminiert nicht signifikante Variablen, wenn keine erklärenden Variablen in der Liste **Einzuschließende Variablen** enthalten sind. Ergebnisse des Wald-Tests werden für einzelne Parameter untersucht. Die am wenigsten signifikante Variable, die nicht das im Feld **Für Verbleiben im Modell** angegebene Signifikanzniveau erreicht, wird entfernt. (Standardmäßig beträgt dieser Wert `0,05`.) Wenn eine Variable aus dem Modell entfernt wurde, bleibt sie ausgeschlossen. Der Vorgang wird solange wiederholt, bis keine andere Variable im Modell das angegebene Niveau erfüllt, um entfernt zu werden. **Schrittweise Auswahl** Die schrittweise Methode ist eine Modifikation des Vorwärtsverfahrens mit der Ausnahme, dass die Variablen, die bereits im Modell enthalten sind, nicht unbedingt beibehalten werden müssen. Variablen werden dem Modell so hinzugefügt und daraus entfernt, dass auf jeden Schritt im Vorwärtsverfahren mindestens ein Schritt im Rückwärtsverfahren folgen kann. Der schrittweise Auswahlprozess endet, wenn keine weitere Variable dem Modell hinzugefügt werden kann, oder wenn die Variable, die gerade dem Modell hinzugefügt wurde, die einzige Variable ist, die im folgenden Schritt des Rückwärtsverfahrens entfernt wird. **Auswahl der besten Teilmenge** Diese Auswahlmethode sortiert die Modelle nach der Likelihood-Scorewert-Statistik (Chi-Quadrat). Sie identifiziert eine bestimmte Anzahl an Modellen, welche die für alle möglichen Modellgrößen die Chi-Quadrat-Statistik mit den höchsten Scorewerten aufweisen, die von einer erklärenden Variablen bis zu einem einzelnen Modell mit allen erklärenden Variablen reichen. Wenn es nicht mehr als 10 erklärende Variablen gibt, werden alle möglichen Modelle für jede Modellgröße aufgeführt. Gibt es mehr als 10 erklärende Variablen, so entspricht die Anzahl der Modelle, die für jede Modellgröße ausgewählt werden, der Gesamtzahl der erklärenden Variablen. Die kleinsten Modelle enthalten *n* erklärende Variablen, wobei *n* von der Anzahl der erklärenden Variablen in der Liste **Einzuschließende Variablen** bis zur Gesamtzahl der erklärenden Variablen im Modell reicht. Wenn die Liste **Einzuschließende Variablen** leer ist, so ist *n* `1`. **Hinweis:** Konstantenparameter bleiben immer im Modell.
Angeben des Signifikanzniveaus	Bei dem Vorwärtsverfahren oder der schrittweisen Auswahlmethode bestimmt das Signifikanzniveau, das Sie im Feld **Für Aufnahme ins Modell** angeben, ob eine erklärende Variable in das Modell eingeschlossen wird. Variablen, die nicht im Modell vorkommen, werden geprüft, und die Variable mit dem kleinsten *p*-Wert wird aufgenommen, wenn der *p*-Wert kleiner oder gleich dem angegebenen Signifikanzniveau ist. Das Standard-Signifikanzniveau beträgt `0,05`. Bei dem Rückwärtsverfahren oder der schrittweisen Auswahlmethode bestimmt das Signifikanzniveau, das Sie im Feld **Für Verbleiben im Modell** angeben, ob eine erklärende

Optionsname	Beschreibung
	Variable im Modell beibehalten wird. Variablen im Modell werden geprüft, und die Variable mit dem größten *p*-Wert wird entfernt, wenn der *p*-Wert über dem angegebenen Signifikanzniveau liegt. Das Standard-Signifikanzniveau beträgt `0,05`.
Auswählen der im Modell verbleibenden Variablen	Das Feld **Einzuschließende Variablen** führt die erklärenden Variablen auf, die während des Auswahlprozesses der Variablen in das Modell eingeschlossen oder daraus entfernt werden sollen. Um eine Variable immer in das Modell einzuschließen, aktivieren Sie das Kontrollkästchen neben dem Variablennamen. Wenn Sie eine Variable nicht mehr in das Modell einschließen möchten, deaktivieren Sie das Kontrollkästchen neben dem Variablennamen.

Proportionale Hazards: Optionen für das Festlegen der Methoden

Klicken Sie im Auswahlfenster auf **Methoden**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Konfidenzgrenzen für Hazard-Quotienten berechnen	zeigt die Konfidenzgrenzen für jede erklärende Variable an. Sie können entweder einen Wert als Konfidenzniveau eingeben oder einen Wert aus der Liste **Konfidenzniveau** auswählen. Der Standardwert ist ein Konfidenzniveau von 95%.
Korrelationen der Parameterschätzer	zeigt die geschätzte Korrelationsmatrix der Parameterschätzer an.
Kovarianzen der Parameterschätzer	zeigt die geschätzte Kovarianzmatrix der Parameterschätzer an.
Methode für gleiche Ausfallzeiten	gibt an, wie gleiche Ausfallzeiten behandelt werden. Folgende Optionen können ausgewählt werden: ■ **Breslows approximierte Likelihood** verwendet Breslows approximierte Likelihood-Methode. Wenn nicht zu viele Verbindungen vorliegen, bietet diese Methode zufrieden stellende Approximationen für die Berechnung exakter

Optionsname	Beschreibung
	bedingter Wahrscheinlichkeiten für das kontinuierliche Zeit-Skala-Modell. Dies ist die Standardmethode. ■ **Diskretes logistisches Modell** ersetzt das proportionale Hazard-Modell durch das diskrete logistische Modell. Wenn die Zeitskala tatsächlich diskret ist, müssen Sie diese Methode verwenden. Wenn die Analyse der Fallkontrollstudien mehr als einen Fall in einem entsprechenden Satz aufweist, ist diese Methode erforderlich. ■ **Efrons approximierte Likelihood** verwendet Efrons approximierte Likelihood-Methode. Wenn nicht zu viele Verbindungen vorliegen, bietet diese Methode zufrieden stellende Approximationen für die Berechnung exakter bedingter Wahrscheinlichkeiten für das kontinuierliche Zeit-Skala-Modell. ■ **Exakte bedingte Wahrscheinlichkeiten berechnen** berechnet die exakte bedingte Wahrscheinlichkeit unter der proportionalen Hazard-Annahme, dass alle überlappenden Ereigniszeiten entweder vor zensierten Zeiten des gleichen Werts oder vor größeren Werten auftreten. **Hinweis:** Diese Methode kann eine beträchtliche Menge der Computerressourcen in Anspruch nehmen. Wenn es keine Verbindungen gibt, erzeugen alle vier Methoden die gleiche Likelihood und identische Schätzer.

Proportionale Hazards: Erzeugen von Plots

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Hinweis: Bevor Sie einen Plot zur Modellbeurteilung erstellen können, müssen Sie der Rolle **Erklärende Variablen** eine Variable zuweisen.

Optionsname	Beschreibung
Kumulatives Martingale-Residuen-Plot	Für jede erklärende Variable zeichnet diese Option die beobachteten kumulativen Martingale-Residuen gegen die Werte der erklärenden Variablen gemeinsam mit 20 simulierten Residuenmustern. Wenn Sie die Option **Panel-Plot der kumulativen Martingale-Residuen** auswählen, können Sie einen

Optionsname	Beschreibung
	Plot mit vier Panels erstellen. Jeder Panel enthält die beobachteten kumulativen Martingale-Residuen und zwei simulierte Residuenmuster.
Standardisiertes Score-Process-Plot	Für jede erklärende Variable zeichnet diese Option die beobachtete Score-Process-Komponente gegen die Nachfolgezeit, gemeinsam mit 20 simulierten Mustern.
Überlebensfunktionsplot	Diese Option stellt die geschätzten Überlebensfunktion für eine Verweismenge der erklärenden Variablen dar. Diese Menge besteht aus den entsprechenden Mittelwerten der Stichproben.
Kumulatives Hazard-Funktionsplot	Diese Option stellt die geschätzten Hazard-Funktion für eine Verweismenge der erklärenden Variablen dar. Diese Menge besteht aus den entsprechenden Mittelwerten der Stichproben. Wenn es keine Verbindungen gibt, erzeugen alle vier Methoden die gleiche Likelihood und identische Schätzer.

Proportionale Hazards: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Speichern von Ausgabedaten	speichert Ihre Ergebnisse teilweise oder vollständig in einer Ausgabetabelle. Aktivieren Sie die Kontrollkästchen der Statistiken, die in Ihrer Ausgabe enthalten sein sollen. ■ Die Option **Überlebensschätzer** erstellt eine Ausgabedatei, in der Überlebensschätzer, Diagnosen und geschätzte lineare Prädiktoren enthalten sind. ■ Die Option **Regressionskoeffizientenschätzer** erstellt eine Ausgabedatei, in der die Regressionskoeffizientenschätzer enthalten sind. Wenn Sie die Kovarianzmatrix in der Datei einschließen möchten, markieren Sie das Kontrollkästchen **Geschätzte Kovarianzmatrix einschließen**. ■ Die Option **Schätzer für Baseline-Überlebensfunktion** erstellt eine Ausgabedatei, in der die Überlebensfunktionsschätzer enthalten sind, die den Mittelwerten der Erklärungsvariablen für jede Schicht

Optionsname	Beschreibung
	entsprechen. Darüber hinaus sind in der Datei die Schätzer der Überlebensfunktion, linearer Prädiktor, Standardfehler, Konfidenzintervalle, der Log des Überlebens sowie der Log des negativen Log des Überlebens enthalten. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Alle angezeigten Ausgaben und Plots unterdrücken	Bewirkt, dass keine Ergebnisse angezeigt werden. Wenn Sie dieses Kontrollkästchen aktivieren, sind die Optionen im Fenster **Plots** nicht verfügbar. Die Ergebnisse werden standardmäßig angezeigt.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

68

Q-Q-Plots

Allgemeines über die Anwendungsroutine 'Q-Q-Plots' ... **547**

Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen ... **548**

Q-Q-Plots: Auswählen der Verteilung ... **550**

Q-Q-Plots: Anpassen der Verteilung ... **551**

Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung ... **551**

Festlegen von Achsenoptionen ... 551
Hinzufügen von Referenzlinien ... 552
Einfügungen hinzufügen ... 552
Festlegen weiterer Plotoptionen ... 552
Festlegen von Optionen für Spezifikationsgrenzen ... 555

Prozessfähigkeitsanalyse: Festlegen der Analysetypen ... **555**

Festlegen von Titeln und Fußnoten ... **556**

Anzeigen von Eigenschaften ... **556**

Allgemeines über die Anwendungsroutine 'Q-Q-Plots'

Eine Prozessfähigkeitsanalyse vergleicht die Verteilung der Ausgabe eines Prozesses in statistischer Kontrolle mit den Spezifikationsgrenzen, um die Konsistenz zu bestimmen, mit der die Spezifikationen erfüllt werden. Ein Quantil-Quantil-Plot (Q-Q-Plot) wird bei der Prozessfähigkeitsanalyse verwendet, um geordnete Werte einer Variablen mit den Quantilen einer angegebenen theoretischen Verteilung, wie der Normalverteilung, zu vergleichen. Wenn die Datenverteilung der theoretischen Verteilung entspricht, bilden die Punkte im Plot ein lineares Muster. Daher können Sie mit einem Q-Q-Plot ermitteln, wie gut eine theoretische Verteilung einen Satz an Messwerten modelliert.

Hinweis: Q-Q-Plots ähneln Wahrscheinlichkeitsplots. Q-Q-Plots sind bei der grafischen Schätzung von Verteilungsparametern und Prozessfähigkeitsindizes

vorzuziehen, wohingegen Wahrscheinlichkeitsplots bei der grafischen Schätzung von Perzentilen geeigneter sind.

Sie können einen Q-Q-Plot auch zum Analysieren eines Fertigungsprozesses verwenden. Angenommen, es werden die Abstände zwischen zwei Bohrungen, die in 50 Stahlbleche geschnitten wurden, gemessen und in einer Datei gespeichert. Der Schneidprozess befindet sich innerhalb der statistischen Kontrolle. Sie möchten überprüfen, ob die Prozessverteilung normal ist. Der Q-Q-Plot vergleicht die Beobachtungen in der Datei mit den Quantilen der Normalverteilung. Die Linearität des Punktmusters weist darauf hin, ob die Messungen eine Normalverteilung aufweisen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	CAPABILITY
Erforderliche SAS-Produkte	Base SAS, SAS/QC, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

Prozessfähigkeitsanalyse: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Analysevariablen	gibt die zu analysierenden numerischen Variablen an. Sie müssen dieser Rolle mindestens eine Variable zuweisen. Bei allen Plots, außer dem P-P-Plot, können Sie eine Obergrenze, einen Zielwert und eine Untergrenze im Bereich

Optionsname	Beschreibung
	Spezifikationsgrenzen für *Variablenname* für jede Analysevariable angeben. Wenn Sie diese Grenzen angeben, enthält die Ausgabe automatisch eine Tabelle mit den Tests auf Normalverteilung, eine Tabelle mit den Spezifikationsgrenzen und eine Tabelle mit den Prozessfähigkeitsindizes.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt die Variable an, die als relative Gewichtungsvariable verwendet werden soll. Wenn Sie dieser Rolle eine Variable zuweisen, werden mit dem Wert der Variablen in jeder Beobachtung gewichtete Statistiken berechnet. Sie können dieser Rolle nur eine Variable zuweisen. **Hinweis:** Sie sollten der Rolle **Relative Gewichtung** keine Variable zuweisen, wenn Sie ein Histogramm, Wahrscheinlichkeitsplot, P-P-Plot, Q-Q-Plot oder CDF-Plot erstellen. Falls Sie dieser Rolle ein Variable zugewiesen haben, müssen Sie eine eventuelle Ausgabe mit Vorsicht interpretieren.
Klassifizierungsvariablen (nur verfügbar für die Anwendungsroutine Histogramm)	gibt eine oder zwei Variablen an, die als Klassifizierungsvariablen dienen. Wenn Sie dieser Rolle eine Variable zuweisen, erstellt die Anwendungsroutine ein Vergleichshistogramm, mit dem Sie die Verteilung einer Analysevariablen über Ausprägungen der Klassifizierungsvariablen hinweg vergleichen können. **Hinweis:** Um vergleichende Wahrscheinlichkeitsplots oder vergleichende Q-Q-Plots zu erstellen, verwenden Sie die Anwendungsroutine Verteilungsanalyse.

Q-Q-Plots: Auswählen der Verteilung

Klicken Sie im Auswahlfenster unter der Überschrift 'Verteilungen' auf **Zusammenfassung**, um auf diese Optionen zuzugreifen.

Sie können eine theoretische Verteilung auf einem Q-Q-Plot anpassen. Wenn Sie die Verteilung ausgewählt haben, wählen Sie den Namen der Verteilung im Auswahlfenster aus, um die Parameterwerte und Linieneigenschaften für diese Verteilung anzugeben.

Optionsname	Beschreibung
Verteilungen	Verwenden Sie die Kontrollkästchen, um eine oder mehrere der folgenden geschätzten Dichtekurven auszuwählen: ■ **Normal** passt eine normale Dichte mithilfe eines Mittelwerts (My) und einer Standardabweichung (Sigma) an. ■ **Lognormal** passt eine Lognormal-Dichte mithilfe eines Skalierungsparameters (Zeta), eines Schwellenwertparameters (Theta) und eines Formparameters (Sigma) an. ■ **Exponentiell** passt eine exponentielle Dichte mithilfe eines Skalierungsparameters (Sigma) und eines Schwellenwertparameters (Theta) an. ■ **Beta** passt eine Beta-Dichte mit einem Skalierungsparameter (Sigma), einem Schwellenwertparameter (Theta) und Formparametern (Alpha und Beta) an. ■ **Gamma** passt eine Gamma-Dichte mit einem Skalierungsparameter (Sigma), einen Schwellenwertparameter (Theta) und einen Formparameter (Alpha) an. ■ **Weibull (3-Parameter)** passt eine 3-Parameter-Weibull-Dichte mit einem Skalierungsparameter (sigma), einem Schwellenwertparameter (theta) und einem Formparameter (c) an. ■ **Weibull (2-Parameter)** passt eine 2-Parameter-Weibull-Dichte mit einem Skalierungsparameter (sigma) und einem Formparameter (c) an. Bei Weibull (2-Parameter) wird angenommen, dass die Daten einen bekannten unteren Schwellenwertparameter, theta=0, haben.
Grafikstil	Wählen Sie den Grafikstil, der in der Ausgabe verwendet werden soll. Sie haben die Auswahl zwischen üblichen SAS-Grafiken und ODS Statistical Graphs (statistischen ODS-Grafiken). Weitere Informationen zu ODS-Grafiken finden Sie unter *SAS Output Delivery System: User's Guide*.

Q-Q-Plots: Anpassen der Verteilung

Klicken Sie im Auswahlfenster unter der Überschrift **Verteilungen** auf den Namen der Verteilung, die Sie anpassen möchten, um auf diese Optionen zuzugreifen.

Hinweis: Diese Verteilungsoptionen sind nur durch vorherige Auswahl der Verteilung auf der Seite **Zusammenfassung** verfügbar, oder durch Aktivieren des Kontrollkästchens ***Name der Verteilung*** auf der Seite dieser Verteilung.

Für die Normal-, Lognormal-, Exponential-, Weibull-, Beta- und Gamma-Verteilungen können Sie die Parameter für jede Analysevariable oder für alle Analysevariablen angeben. Standardmäßig schätzt die Anwendungsroutine die Parameterwerte für alle Analysevariablen.

So legen Sie für jede Analysevariable Parameter fest:

1 Deaktivieren Sie das Kontrollkästchen **Verteilung auf alle Variablen anwenden**.

2 Aktivieren Sie das Kontrollkästchen der Variablen im Feld **Analysevariablen**.

3 Geben Sie die Parameterwerte an.

Hinweis: Um zu den geschätzten Parameterwerten zurückzukehren, klicken Sie auf **Schätzer verwenden**.

Prozessfähigkeitsanalyse: Festlegen von Optionen für die Plotdarstellung

Festlegen von Achsenoptionen

Klicken Sie im Auswahlfenster unter **Darstellung** ⇨ **Achsen** auf **Achsen**, um auf die Optionen zuzugreifen.

Folgende Optionen können angegeben werden:

- die Farbe der Achsen
- die Farbe der Achsenmarkierungen

- ein Etikett für die horizontalen und vertikalen Achsen

Hinzufügen von Referenzlinien

Klicken Sie im Auswahlfenster unter **Darstellung** ⇨ **Achsen** auf **Horizontal** bzw. **Vertikal**, um auf die Optionen zuzugreifen.

Um Referenzlinien zu verwenden, wählen Sie **Referenzlinien verwenden**. Sie können die Referenzlinie formatieren und die Position jeder Referenzlinie festlegen.

Einfügungen hinzufügen

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Einfügung**, um auf diese Optionen zuzugreifen.

Sie können einen Plot erweitern, indem Sie direkt in der Grafik ein Feld oder eine Tabelle (eine so genannte Einfügung) der beschreibenden Statistiken hinzufügen. Wenn Sie ein Vergleichshistogramm erstellen, erscheint in jedem Komponentenhistogramm eine Einfügung.

Wählen Sie **Einfügung einschließen**, um eine beschreibende Statistik zum Plot hinzuzufügen. Wählen Sie dann die Statistiken, die Sie in die Einfügung einschließen möchten.

Hinweis: Die Statistiken für die Berechnung der Prozessfähigkeitsindizes sind nur verfügbar, wenn Sie Spezifikationsgrenzen festgelegt und dann der Rolle **Analysevariablen** zugewiesen haben. Sie können Spezifikationsgrenzen für Histogramm, CDF-Plots, Wahrscheinlichkeitsplots und Q-Q-Plots erstellen. Weitere Informationen zur Berechnung der Prozessfähigkeitsindizes finden Sie in der Hilfe zur Prozedur CAPABILITY.

Festlegen weiterer Plotoptionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Optionen**, um auf diese Optionen zuzugreifen.

Die verfügbaren Plotdarstellungsoptionen hängen vom Typ des Plots ab, den die Anwendungsroutine erzeugt.

Optionsname	Beschreibung
Hintergrund	bestimmt die Hintergrundfarbe des Plots.

Optionsname	Beschreibung
Balkenumriss und **Balkenfüllung**	legt die Farben der Balken in einem Histogramm fest. **Hinweis:** Diese Optionen sind nur für Vergleichshistogramme verfügbar.
Beide Achsen umkehren (Perzentile auf entgegengesetzte Achsen)	kehrt die Achsen auf dem Plot um. Wenn Sie diese Option auswählen, funktionieren die von Ihnen ausgewählten Achsenoptionen und Referenzlinien eventuell nicht mehr, da sich die Ausrichtung des Plots ändert. **Hinweis:** Diese Option ist nur für Wahrscheinlichkeitsplots und Q-Q-Plots verfügbar.
Keine Standardlegenden	verhindert die Anzeige der Standardlegende mit dem Plot.
Midpoints verwenden	geben Sie Midpoints für das Histogramm an. Die in der Dropdown-Liste verfügbaren Optionen hängen davon ab, ob Sie ein Histogramm oder ein Vergleichshistogramm erstellen. Wenn Sie der Rolle **Klassifizierungsvariablen** eine Variable zuweisen, erstellen Sie ein Vergleichshistogramm. Für Vergleichshistogramme stehen folgende Optionen zur Verfügung: ■ **Midpoints angeben** - Ermöglicht das Angeben der Start- und End-Midpoints. Sie können auch das Schrittintervall festlegen, das für die Berechnung der anderen Midpoints verwendet werden soll. Der Midpoint ist der Mittelwert des Wertebereichs, der von den einzelnen Balken dargestellt wird. ■ **Schlüssel** - Bestimmt die Midpoints für die Daten in der Schlüsselzelle. Standardmäßig werden die Ausprägungen der Klassifizierungsvariablen von oben nach unten (links nach rechts) in aufsteigender Reihenfolge der internen (nicht formatierten) Werte der ersten Klassifizierungsvariablen angezeigt. Wenn Sie nur eine Klassifizierungsvariable angegeben haben, ist die Schlüsselzelle die Ausprägung, die als erstes in dieser Reihenfolge erscheint. Wenn Sie zwei Klassifizierungsvariablen angegeben haben, ist die Schlüsselzelle die Kombination der Ausprägungen von Variable 1 und Variable 2, die als erstes in dieser Reihenfolge erscheint. Die Auswahl der Schlüsselzelle bestimmt die einheitliche Horizontalachse, die für alle Zellen verwendet wird. Die Midpoint-Liste für die Schlüsselzelle wird dann wie erforderlich in beide Richtungen erweitert, bis sie die Daten in den übrigen Zellen abdeckt. ■ **Einheitlich** - Bestimmt die Anzahl der Midpoints abhängig von der Gesamt-Stichprobengröße.

Optionsname	Beschreibung
	Wenn Sie der Rolle **Klassifizierungsvariablen** keine Variable zuweisen, erstellen Sie ein Histogramm. Für Histogramme sind die folgenden Optionen verfügbar: ■ **Midpoints angeben** - Ermöglicht das Angeben der Start- und End-Midpoints. Sie können auch das Schrittintervall festlegen, das für die Berechnung der anderen Midpoints verwendet werden soll. Der Midpoint ist der Mittelwert des Wertebereichs, der von den einzelnen Balken dargestellt wird. ■ **Anzahl der Intervalle angeben** - Ermöglicht, die Anzahl der Intervalle (auch Histogrammintervalle genannt) für die Daten anzugeben. Standardmäßig basiert die Anzahl der angezeigten Intervalle auf dem Datenbereich. Sie können jedoch eine Maximalzahl anzuzeigender Intervalle angeben. **Hinweis:** Sie können nur dann Midpoints angeben, wenn Sie die Anwendungsroutine **Histogramme** verwenden.
Größe ben.def. Diagramm angeben (in Pixel)	legt die Bildgröße fest. Zum Vergleich wird die Gesamtbildschirmgröße angezeigt in den Feldern **Breite** und **Höhe** angezeigt.

Hinweis: Vergleichsoptionen sind für Vergleichshistogramme verfügbar. Sie können ein Vergleichshistogramm erstellen, indem Sie in der Anwendungsroutine **Histogramme** der Rolle **Klassifizierungsvariablen** eine Variable zuweisen.

Optionsname	Beschreibung
Zeilen und **Spalten**	Wenn Sie einer Analyserolle mindestens eine Klassifizierungsvariable zuweisen, können Sie über die Felder **Zeilen pro Seite** und **Spalten pro Seite** die Anordnung der Spalten und Zeilen im Vergleichshistogramm ändern. Wenn Sie eine Klassifizierungsvariable zuweisen, ist die Standardanordnung zwei Zeilen und eine Spalte pro Seite. Wenn Sie zwei Klassifizierungsvariablen zuweisen, ist die Standardanordnung zwei Spalten und zwei Zeilen pro Seite. Sie können auch über die Dropdown-Listen **Rahmenseite** und **Rahmenoberseite** Farben zum Füllen des Rahmenbereichs der Zeilenetiketten und Spaltenetiketten angeben.
Anzahl der Intervalle	Standardmäßig basiert die Anzahl der angezeigten Intervalle auf dem Datenbereich. Sie können jedoch eine Maximalzahl anzuzeigender Intervalle angeben. Sie können auch eine Standardabweichung angeben, so dass die Anzahl der angezeigten Intervalle von einem Bereich des

Optionsname	Beschreibung
	Standardabweichungswerts über und unter dem Mittelwert der Daten abhängt.

Festlegen von Optionen für Spezifikationsgrenzen

Hinweis: Diese Optionen sind für Histogramme, Wahrscheinlichkeitsplot, Q-Q-Plots und CDF-Plots verfügbar.

Klicken Sie im Auswahlfenster unter **Darstellung** auf **Spezifikationsgrenzen**, um die untere Spezifikationsgrenze, das Ziel und die obere Spezifikationsgrenze anzugeben.

Hinweis: Die Optionen für **Spezifikationsgrenzen** sind nur verfügbar, wenn Sie eine oder mehrere Grenzen für mindestens eine Analysevariable angeben. Um eine Spezifikationsgrenze für eine Analysevariable anzugeben, klicken Sie im Auswahlfenster auf **Daten** und wählen dann die Analysevariable aus.

Prozessfähigkeitsanalyse: Festlegen der Analysetypen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

Standardmäßig enthält der Prozessfähigkeitsanalysebericht die folgenden Tabellen: eine mit Grundkonfidenzintervallen, eine mit Grundmaßen, eine mit Tests auf Lageparametern und eine mit Momenten.

Um eine Tabelle zum Bericht hinzuzufügen, wählen Sie die Tabelle im Feld **Tabellen** aus.

Um die Analyseoptionen festzulegen, wählen Sie die Tabelle im Feld **Tabellen** aus. Sind Optionen verfügbar, erscheint rechts ein Feld, in dem Sie die Optionen angeben können.

Um alle Tabellen mit beschreibenden Statistiken im Bericht zu unterdrücken, aktivieren Sie die Option **Beschreibende Statistiken und Tabellen für Prozessfähigkeitsindizes unterdrücken**. Folgende Tabellen werden durch Auswahl dieser Option nicht unterdrückt:

- Tabellen, die von der ausgewählten Verteilung erstellt wurden. Sie können diese Tabellen unterdrücken, indem Sie im Fenster für diese Verteilung das Kontrollkästchen **Verteilungstabellen unterdrücken** aktivieren.
- Tabellen, die mit der INTERVALS-Anweisung erstellt wurden. Um diese Tabellen zu unterdrücken, können Sie die Option NOPRINT angeben. Weitere Informationen zur Anweisung INTERVALS finden Sie in der Hilfe zur Prozedur CAPABILITY.

Um Ihre Ergebnisse in eine Ausgabedatei zu speichern, aktivieren Sie das Kontrollkästchen **Ausgabestatistiken in einer Datei speichern**. Die Datei enthält univariate Statistiken und Prozessfähigkeitsindizes.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

   Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

69

Radardiagramm

Allgemeines über die Anwendungsroutine 'Radardiagramm' ... **559**

Radardiagramm: Zuweisen von Variablen zu Analyserollen ... **560**

Radardiagramm: Festlegen von Darstellungsoptionen ... **561**

Festlegen von Optionen für Radardiagramme ... 561

Weitere Darstellungsoptionen ... 562

Festlegen von Titeln und Fußnoten ... **562**

Anzeigen von Eigenschaften ... **562**

Allgemeines über die Anwendungsroutine 'Radardiagramm'

Die Anwendungsroutine **Radardiagramm** erstellt Radardiagramme (oder Sterndiagramme), welche die relative Häufigkeit von Datenmessungen bei Problemen in der Qualitätskontrolle oder Marktforschung aufzeigen. Die Diagrammstatistiken werden auf Speichen angezeigt, die von der Diagrammmitte ausgehen.

Beispielsweise kann ein Radardiagramm die Gesamtumsatzzahlen für drei Produktionsstätten anzeigen. Die Umsatzzahlen für einen Standort werden durch die Länge der Speiche dargestellt.

Sie können kein Radardiagramm erstellen, wenn Sie die Ausgabeformate Java-Applet oder Java-Bild ausgewählt haben.

Weitere Informationen, siehe “Auswählen des Grafikausgabeformats” auf Seite 6.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GRADAR
Erforderliche SAS-Produkte	Base SAS, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	Ohne

Radardiagramm: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Spalte für Diagramm	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Speichen des Radardiagramms. Die Spalte muss mindestens drei Beobachtungen aufweisen, da für die Definition einer Ebene drei Punkte nötig sind.
Diagramme gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Diagramme, die erstellt werden. Für jeden eindeutigen Wert wird eine separate Grafik erzeugt.
Häufigkeitszählung	Die Werte der Spalte, die Sie dieser Rolle zuweisen, stellen die Häufigkeit (Anzahl der Vorkommen) der Werte in der Rolle **Darzustellende Spalten** zur Verfügung. Die Spalte muss numerisch sein. Wenn Sie diese Rolle weglassen, wird jeder Wert der Spalte **Darzustellende Spalten** genau einmal gezählt.
Überlagerung	Die Werte der Spalte, die Sie dieser Rolle zuweisen, erzeugen ein vergleichendes Radardiagramm untr Verwendung der Ausprägungen der angegebenen Spalte. Alle Diagramme

Optionsname	Beschreibung
	werden im gleichen Speichensatz dargestellt. Sie können diese Rolle nicht verwenden, wenn Sie der Rolle **Quer** oder **Nach unten** Spalten zuweisen.
Quer	Die Spalte, die Sie dieser Rolle zuweisen, erzeugt ein Radardiagramm für jeden Wert der angegebenen Spalte. Die Diagramme werden von links nach rechts quer über den Grafikbereich gezeichnet. Wenn Sie der Rolle **Nach unten** auch eine Spalte zuweisen, werden die Diagramme von links nach rechts und von oben nach unten gezeichnet.
Nach unten	Die Spalte, die Sie dieser Rolle zuweisen, erzeugt ein Radardiagramm für jeden Wert der angegebenen Spalte. Die Diagramme werden von oben nach unten abwärts über den Grafikbereich gezeichnet. Wenn Sie der Rolle **Quer** auch eine Spalte zuweisen, werden die Diagramme von links nach rechts und von oben nach unten gezeichnet.

Radardiagramm: Festlegen von Darstellungsoptionen

Festlegen von Optionen für Radardiagramme

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Radar**, um auf diese Optionen zuzugreifen.

Wählen Sie den Stil für das Diagramm aus.

- Mit der Option **Korona** wird ein Polygon mit Sternscheitelpunkten erstellt, die vom Kreismittelpunkt her ausstrahlen.
- Mit der Option **Polygon** wird ein geschlossenes Polygon erstellt.
- Mit der Option **Strahlen** werden vom Mittelpunkt ausgehende Strahlen erstellt.
- Mit der Option **Speiche** werden vom Kreismittelpunkt ausgehende Strahlen erstellt.
- Mit der Option **Keil** wird ein geschlossenes Polygon mit Strahlen erstellt, die vom Mittelpunkt zur vollen Speichenlänge verlaufen (dies ist der Standardwert).

Weitere Darstellungsoptionen

Einige Darstellungsoptionen gelten für alle Grafik-Anwendungsroutinen. Weitere Informationen zu diesen Optionen, siehe “Anpassen des Diagrammbereichs” auf Seite 750.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine

- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

70

Zufällige Stichprobe

Allgemeines über die Anwendungsroutine 'Zufällige Stichprobe' **565**
Zufällige Stichprobe: Erstellen einer zufälligen Stichprobe **566**

Allgemeines über die Anwendungsroutine 'Zufällige Stichprobe'

Die Anwendungsroutine **Zufällige Stichprobe** erstellt eine Ausgabedatei, die eine zufällige Stichprobe der Zeilen in der Eingabedatei enthält.

Sie können diese Anwendungsroutine verwenden, wenn Sie eine Teilmenge von Daten benötigen. Angenommen, Sie möchten die Reisekosten von Angestellten untersuchen, um das Verfahren zum Anfertigen der Reisekostenabrechnungen zu verbessern und so möglicherweise die Reisekosten senken zu können. Sie verfügen nicht über ausreichend Ressourcen, um alle Reisekostenabrechnungen zu untersuchen, sodass Sie auf statistische Stichproben zurückgreifen, um die Reisekostenabrechnungen für die Prüfung objektiv auszuwählen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SURVEYSELECT
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	Ohne

Zufällige Stichprobe: Erstellen einer zufälligen Stichprobe

1 Geben Sie einen Namen für die Anwendungsroutine ein.

Hinweis: Diese Option ist nur im SAS Enterprise Guide verfügbar.

2 Wählen Sie die Eingabedatenquelle für die Anwendungsroutine aus.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

3 Wählen Sie die Variablen aus, die in der Ausgabedatei enthalten sein sollen. Standardmäßig werden alle Variablen in die Ausgabedatei einbezogen.

Hinweis: Wenn Sie der Rolle **Schichtvariablen** eine Variable zuweisen, wird diese automatisch auch der Rolle **Ausgabevariablen** zugewiesen. Wenn Sie eine Variable aus der Ausgabe entfernen, wird diese auch aus der Schichtvariable entfernt.

4 Geben Sie die Stichprobengröße in den gewünschten Zeilen bzw. im gewünschten Prozentsatz der Eingabezeilen an. Wenn Sie z.B. 3 Prozent angeben und es 400 Eingabezeilen gibt, besteht die resultierende Stichprobe aus 12 Zeilen.

Hinweis: Wenn Sie Variablen zur Rolle **Schichtvariablen** zuweisen, gilt die hier vorgenommene Angabe zur Stichprobengröße für alle Schichten statt für die gesamte Eingabedatei.

5 Wählen Sie eine Stichprobenmethode aus. Folgende Optionen können ausgewählt werden:

Einfach (Keine Dubletten)
Wenn eine Zeile ausgewählt wird, steht sie für eine weitere Auswahl nicht mehr zur Verfügung. Dadurch kann dieselbe Zeile nicht mehrmals ausgewählt werden.

Unbeschränkt (Dubletten zulässig)
Wenn eine Zeile ausgewählt wird, bleibt sie weiterhin auswählbar. Dadurch kann dieselbe Zeile mehrmals ausgewählt werden. Sie können angeben, wie

mehrfaches Auswählen der gleichen Zeile in der Ausgabedatei protokolliert werden soll. Sie können folgende Optionen wählen:

- **Jede Beobachtung einmal in Ausgabe einbeziehen (ohne Dubletten)** - Eine Zeile, die *n*-Mal ausgewählt wurde, erscheint nur einmal in der Stichprobe. In der Ausgabe wird für die Variable **NumberHits** (die automatisch von der Anwendungsroutine **Zufällige Stichprobe** berechnet wird) angegeben, wie häufig die Beobachtung in der Eingabedatei vorkommt.
- **Alle Beobachtungen in Ausgabe einbeziehen (inklusive Dubletten)** - Eine Zeile, die *n*-Mal ausgewählt wurde, erscheint in der Stichprobe *n*-Mal.

6 Legen Sie die Variablen fest, anhand derer die Eingabedatei in sich gegenseitig ausschließende, nicht überlappende Teilmengen (= Schichten) unterteilt wird. Jede Schicht wird anhand einer Menge von Werten der Schichtvariablen definiert, und von jeder Schicht wird separat eine Stichprobe genommen. Die vollständige Stichprobe ist die Summe aller Stichproben, die aus allen Schichten genommen werden.

Hinweis: Wenn Sie dieser Rolle keine Variablen zuweisen, wird die gesamte Eingabedatei als einzelne Schicht behandelt.

Sie können die gesamte Stichprobengröße proportional zur Größe einer einzelnen Schicht auf die Schichten aufteilen. Beispiel: Die Variable GESCHLECHT kann die beiden Werte M und F aufweisen und die Variable GEWÄHLT die Werte Y und N. Wenn Sie sowohl GESCHLECHT als auch GEWÄHLT als **Schichtvariablen** definieren, wird die Eingabedatei in vier Schichten untergliedert – Männer, die wählten, Männer, die nicht wählten, Frauen, die wählten, und Frauen, die nicht wählten.

Die Eingabedatei enthält 20.000 Zeilen. Hier sehen Sie die Verteilung der Daten:

- 7.000 Männer, die gewählt haben
- 4.000 Männer, die nicht gewählt haben
- 5.000 Frauen, die gewählt haben
- 4.000 Frauen, die nicht gewählt haben

Das Verhältnis der Männer, die gewählt haben, beträgt 7.000/20.000=0,35 oder 35%. Die Verhältnisse der Stichprobe sollten den Verhältnissen der Schicht in der Eingabedatei entsprechen. Beispiel: Wenn Ihre Stichprobendatei 100 Beobachtungen enthält, müssen 35% der Werte in der Stichprobe aus der Schicht mit den Männern die gewählt haben stammen, damit die Verhältnisse der Eingabedatei eingehalten werden.

Wenn Sie die Option **Wenn nicht genügend Beobachtungen zum Auffüllen der Schicht vorhanden sind, alle auswählen** gewählt haben, werden alle Beobachtungen in einer Schicht ausgewählt, sofern die Größe der Schicht kleiner als die angegebene Stichprobengröße ist.

7 Geben Sie den Namen und den Speicherort für die Ausgabedatei an. Standardmäßig wird die Datei in der Bibliothek WORK gespeichert.

8 Geben Sie den Ausgangswert für die Erstellung der Zufallszahlen fest. Wenn Sie keinen zufälligen Ausgangswert angeben, wird zur Erstellung der Stichprobe ein Ausgangswert verwendet, der auf der Systemuhr beruht.

9 Sie können eine Zusammenfassungstabelle erstellen, die den Ausgangwert enthält, der zur Generierung der Stichprobe verwendet wurde. Durch späteres Angeben desselben Ausgangswerts bei derselben Eingabedatei können Sie dieselbe Stichprobe erzeugen.

71

Rangzuweisung für Daten

Allgemeines zur Anwendungsroutine 'Rangzuweisung für Daten' **569**
Rangzuweisung für Daten: Zuweisen von Variablen zu Analyserollen **570**
Rangzuweisung für Daten: Festlegen der Analyseoptionen **570**
Rangwert: Festlegen von Ergebnisoptionen **572**
Anzeigen von Eigenschaften **573**

Allgemeines zur Anwendungsroutine 'Rangzuweisung für Daten'

Die Anwendungsroutine **Rangzuweisung für Daten** berechnet Rangwerte für eine oder mehr numerische Variablen in den Beobachtungen einer SAS-Datei und sendet die Rangwerte als Ausgabe an eine neue SAS-Datei.

Beispielsweise können Sie die Rangwerte des Umsatzes mit den verschiedenen Produkten Ihres Unternehmens bestimmen. In diesem Fall würde die Rangwertvariable die Reihenfolge der Produktverkäufe anzeigen. Das Produkt mit den höchsten Umsatzzahlen würde den höchsten Rangwert erhalten.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	RANK
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

Rangzuweisung für Daten: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Rangspalten	Jede Spalte, die dieser Rolle zugewiesen ist, wird nach Rang angeordnet. Sie müssen dieser Rolle mindestens eine Variable zuweisen. Standardmäßig erhält die Rangwertspalte den Namen rank_*Spaltenname*, wobei *Spaltenname* der Name der Originalspalte ist. Wenn Sie einen neuen Namen angeben möchten, geben Sie diesen in das Textfeld **Rangspaltenname** ein.
Rang nach	Wenn Sie dieser Rolle mindestens eine Spalte zuweisen, wird die Eingabetabelle nach der ausgewählten Spalte oder den ausgewählten Spalten sortiert, während die Rangwerte innerhalb der einzelnen Gruppen berechnet werden.

Rangzuweisung für Daten: Festlegen der Analyseoptionen

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Rangzuweisungsmethode	gibt die Rangzuweisungsmethode an. Folgende Optionen können ausgewählt werden: ■ Mit der Option **Kleinster zum größten** werden die Originalwerte vom kleinsten zum größten Rangwert sortiert. ■ Mit der Option **Perzentilrangwerte** werden die ursprünglichen Werte in 100 Gruppen eingeteilt, in denen den kleinsten Werten ein Perzentilwert von 0 und den größten Werten ein Perzentilwert von 99 zugeordnet wird. ■ Mit der Option **Dezile** werden die ursprünglichen Werte in 10 Gruppen eingeteilt, in denen den kleinsten Werten der Dezil-Wert 0 und den größten Werten der Dezil-Wert 9 zugeordnet wird. ■ Mit der Option **Quartile** werden die ursprünglichen Werte in 4 Gruppen unterteilt, in denen den kleinsten Werten der Quartilwert 0 und den größten Werten der Quartilwert 3 zugewiesen wird. ■ **Gruppe = n (NTILES)** teilt die ursprünglichen Werte in *n* Gruppen ein, in denen den kleinsten Werten der Wert 0 und den größten Werten der Wert *n*-1 zugeordnet wird. Geben Sie den Wert von *n* im Textfeld **Anzahl der Gruppen** an. ■ Mit der Option **Bruchteil-Rangwerte mit Nenner = n** werden die Bruchteil-Rangwerte durch Division jedes Rangwerts durch die Anzahl der Beobachtungen dividiert, die über nichtfehlende Werte der Rangwertvariablen verfügen. ■ Mit der Option **Bruchteil-Rangwerte mit Nenner = n+1** werden die Bruchteil-Rangwerte durch Division jedes Rangwerts durch den Nenner n+1 berechnet. Dabei ist n die Anzahl der Beobachtungen, die über nichtfehlende Werte der Rangwert-Variablen verfügen. ■ Mit der Option **Prozente** wird jeder Rang durch die Anzahl der Beobachtungen dividiert, die über nichtfehlende Variablenwerte verfügen. Anschließend wird das Ergebnis durch 100 dividiert, um einen Prozentwert zu erhalten. ■ Mit der Option **Normalscorewerte (Blom-Formel)** werden aus den Rangwerten mithilfe der Blom-Formel Normalscorewerte berechnet. ■ Mit der Option **Normalscorewerte (Tukey-Formel)** werden aus den Rangwerten mithilfe der Tukey-Formel Normalscorewerte berechnet. ■ Mit der Option **Normalscorewerte (van-der-Waerden-Formel)** werden aus den Rangwerten mithilfe der van-der-Waerden-Formel Normalscorewerte berechnet.

Optionsname	Beschreibung
	■ Mit der Option **Savage-Scorewerte (exponentiell)** werden die Savage-Scorewerte (oder exponentiellen Scorewerte) aus den Rangwerten berechnet.
Wenn Werte überlappen	gibt die Methode zum Auflösen sich überlappender Werte an. Wenn Sie keine Methode auswählen, verwendet die Anwendungsroutine die Methode Mittelwert (Midrank). ■ **Mittelwert (Midrank)** bewirkt, dass überlappende Werte den Mittelwert der entsprechenden Rangwerte erhalten. ■ **Hoher Rangwert** bewirkt, dass der größte der entsprechenden Rangwerte verwendet wird. ■ **Niedriger Rangwert** bewirkt, dass der kleinste der entsprechenden Rangwerte verwendet wird.
Rangwerte vom größten zum kleinsten umkehren	legt fest, dass die Rangfolge umgekehrt wird.

Rangwert: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Standardmäßig enthält die Ausgabetabelle die Originalspalten und die Rangspalten. Wenn Sie die Originalspalte durch die Rangspalten ersetzen möchten, deaktivieren Sie das Kontrollkästchen **Rangwerte einschließen**.

Der Name der Rangspalte wird beim Zuweisen der Originalspalte zur Rolle **Rangspalten** angegeben. Standardmäßig erhält die Rangwertspalte den Namen rank_*Spaltenname*, wobei *Spaltenname* der Name der Originalspalte ist. Um den Spaltennamen zu ändern, wählen Sie im Auswahlfenster **Daten**.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

72

Recency-, Frequency- und Monetary-Analyse

Allgemeines zur Anwendungsroutine 'RFM-Analyse' **575**

Recency-, Frequency- und Monetary-Analyse: Auswählen des zu analysierenden Datentyps **576**

Recency-, Frequency- und Monetary-Analyse: Zuweisen von Variablen zu Analyserollen **577**

Recency-, Frequency- und Monetary-Analyse: Kategorisieren von Daten nach Intervallen **579**

Recency-, Frequency- und Monetary-Analyse: Speichern von Ergebnissen **580**

RFM-Analyse: Auswählen von Plots **580**

Allgemeines zur Anwendungsroutine 'RFM-Analyse'

Die Anwendungsroutine **RFM-Analyse (RFM, Recency, Frequency, and Monetary)** ist eine Technik, mit der Sie unter Ihren bestehenden Kunden die Kunden herausfiltern können, die am wahrscheinlichsten Interesse an einer neuen Kampagne oder einem Produktangebot haben werden. Dieses Verfahren ist eine der am häufigsten verwendeten Techniken im Direktmarketing.

Die Eingabedaten für diese Anwendungsroutine können Transaktionsdaten oder Kundendaten sein. Mit der Anwendungsroutine RFM-Analyse werden die Transaktionsdaten in Kundendaten aggregiert. Bei Transaktionsdaten stellt jeder Datensatz eine Transaktion dar.

Die Scorewerte für Recency, Frequency und Monetary werden auf die folgende Art und Weise ermittelt:

- Der Recency-Scorewert wird bestimmt, indem die Werte zum Datum der letzten Transaktion in aufsteigender Reihenfolge sortiert und anschließend in Intervalle gruppiert werden. Dem Intervall mit den ältesten Daten wird ein Recency-Score von 1 zugewiesen usw..
- Der Frequency-Scorewert wird bestimmt, indem die Werte der Anzahl der Transaktionen in aufsteigender Reihenfolge sortiert und anschließend in Intervalle gruppiert werden. Dem Intervall mit den wenigsten Transaktionen wird ein Frequency-Score von 1 zugewiesen usw..
- Der Monetary-Scorewert wird bestimmt, indem die Werte des Gesamtbetrags der Transaktion in aufsteigender Reihenfolge sortiert und anschließend in Intervalle gruppiert werden. Dem Intervall mit dem geringsten Betrag wird ein Monetary-Score von 1 zugewiesen usw..

Der RFM-Score wird nach folgender Formel berechnet: RFM-Score = 100 * Recency-Score + 10 * Frequency-Score + Monetary-Score. Das am wenigsten vielversprechende Kundensegment erhält den niedrigsten RFM-Score. Das am meisten versprechende Kundensegment erhält den höchsten RFM-Score.

Hinweis: Für die Anwendungsroutine **RFM-Analyse** ist SAS 9.4 oder höher erforderlich.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	

Recency-, Frequency- und Monetary-Analyse: Auswählen des zu analysierenden Datentyps

Klicken Sie im Auswahlfenster auf **RFM-Typ**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Daten auf Transaktionsebene	Jeder Datensatz in den Daten stellt eine Transaktion dar. Mit der Anwendungsroutine **RFM-Analyse** werden die einzelnen Transaktionen nach Kunden-ID aggregiert.
Kundendaten	Jeder Datensatz stellt einen Kunden dar. Die einzelnen Transaktionen werden vor dem Öffnen der Anwendungsroutine **RFM-Analyse** miteinander kombiniert.

Recency-, Frequency- und Monetary-Analyse: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Welche Analyserollen zur Verfügung stehen, ist vom zu analysierenden Datentyp abhängig.

Tabelle 72.1 *Optionen für Daten auf Transaktionsebene*

Optionsname	Beschreibung
Transaktionsdatum	die Variable, mit der das Datum der Transaktion festgelegt wird. Mit dieser Angabe wird das aktuellste (letzte) Transaktionsdatum bestimmt.
Transaktionsbetrag	die Variable, mit der der Betrag der Transaktion festgelegt wird. Diese Daten werden verwendet, um den Gesamtbetrag der Transaktionen zu berechnen, etwa die Summe, den Mittelwert, den Median oder das Maximum der Transaktionen.

Optionsname	Beschreibung
Kunden-ID	die Variable, die eine Zahl oder eine Zeichenkette enthält, durch die ein Kunde eindeutig identifiziert werden kann.

Da die Anwendungsroutine RFM-Analyse transaktionale Daten in Kundendaten aggregiert, müssen Sie eine Aggregationsmethode für die Transaktionen angeben. Standardmäßig lautet die Aggregationsmethode Summe.

Sie können angeben, ob Kunden mit unvollständigen Transaktionsdatensätzen (oder Datensätze mit fehlenden Werten) in die Aggregation einbezogen werden sollen.

- Wenn Sie diese Option auswählen, enthält die Ausgabedatei die Scorewerte für Kunden mit unvollständigen Transaktionsdatensätzen. Bei jedem Kunden werden nur die Transaktionsdatensätze aggregiert, die nichtfehlende Werte für das Transaktionsdatum und den Betrag aufweisen. Alle Datensätze mit fehlenden Werten werden ignoriert.
- Wenn Sie diese Option nicht auswählen, enthält die Ausgabedatei keine Scorewerte für Kunden mit unvollständigen Transaktionsdatensätzen. Es werden nur Transaktionsdatensätze für Kunden mit vollständigen Datensätzen aggregiert.

Tabelle 72.2 *Optionen für Kundendaten*

Optionsname	Beschreibung
Letzte Transaktion	die Variable, mir der das Datum der letzten Transaktion des Kunden beim Unternehmen festgelegt wird. Diese Angabe wird zur Berechnung des Recency-Scorewerts verwendet.
Anzahl der Transaktionen	die Variable, mit der die Anzahl der Transaktionen für die einzelnen Kunden festgelegt wird. Diese Angabe wird zur Berechnung des Frequency-Scorewerts verwendet.
Transaktionsbetrag	die Variable, mit der der gesamte Geldbetrag festgelegt wird, den ein Kunde ausgegeben hat. Diese Angabe wird zur Berechnung des Monetary-Scorewerts verwendet.
Kunden-ID	die Variable, die eine Zahl oder eine Zeichenkette enthält, durch die ein Kunde eindeutig identifiziert werden kann.

Recency-, Frequency- und Monetary-Analyse: Kategorisieren von Daten nach Intervallen

Klicken Sie im Auswahlbereich auf **Intervalle**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Intervallbildungsverfahren	gibt die Methode zu Kategorisierung der Daten an. Folgende Optionen können ausgewählt werden: ■ Bei der Option **Verschachtelt** wird jedem Recency-Wert ein einfacher Rangwert zugewiesen. Innerhalb jedes Recency-Rangwerts wird den Kunden ein Frequency-Rangwert und innerhalb jedes Frequency-Rangwerts ein Monetary-Rangwert zugewiesen. Dadurch ergibt sich eine gleichmäßige Verteilung der kombinierten RFM-Scorewerte und es wird zudem berücksichtigt, dass sich das Kaufverhalten der Kunden (z.B. Frequency und Monetary) mit der Zeit ändern kann. Der Nachteil dieses Verfahrens liegt hingegen darin, dass die Interpretation der Frequency- und Monetary-Scorewerte länger dauert, da alle Scores voneinander abhängig. ■ Bei der Option **Unabhängig** werden die Rangwerte für Recency, Frequency und Monetary unabhängig voneinander vergeben. Der Vorteil dieses Verfahrens ist, dass die Interpretation der Scorewerte eindeutig ist. Ein Frequency-Score von 5 bei dem einen Kunden bedeutet dasselbe wie der Frequency-Score von 5 bei einem anderen Kunden. Bei kleineren Stichproben treten bei diesem Verfahren jedoch weniger gleichmäßige Verteilungen der kombinierten RFM-Scorewerte auf. Ebenso besteht die Möglichkeit, dass manche Segmente keine Kunden aufweisen.
Anzahl der Intervalle	Sie können die Anzahl der Intervalle für Recency-, Frequency- und Monetary-Werte selbst festlegen. Die angegebenen Werte werden anschließend miteinander multipliziert, um die Gesamtanzahl der Intervalle zu berechnen. Beispiel: Wenn Sie 5 Intervalle für Recency, 5 Intervalle für Frequency und 5 für Monetary angeben, ergibt sich eine Gesamtanzahl von 125 Intervallen.

Recency-, Frequency- und Monetary-Analyse: Speichern von Ergebnissen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Speicherort für RFM-Scores	Standardmäßig werden die Scorewerte für Recency, Frequency und Monetary (RFM) in einer Datei in der WORK-Bibliothek gespeichert. Wenn Sie diesen Speicherort ändern möchten, klicken Sie auf **Bearbeiten**.
Variablen der Ausgabedatei	Sie können die Namen der Variablen in der Ausgabedatei anpassen.

RFM-Analyse: Auswählen von Plots

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Monetarisierungs-Heatmap	erstellt eine Heatmap, in der die monetären Werte für jede Kombination aus Recency- und Frequency-Intervallwerten zusammengefasst werden. Dieser Plot wird standardmäßig angezeigt.
RFM-Verteilungsdiagramme	erstellt ein Balkendiagramm, mit dem Sie die Gleichmäßigkeit der Verteilung der Intervallwerte überprüfen können. Dieser Plot wird standardmäßig angezeigt.
Zusammenfassung der RFM-Segmente	erstellt eine Tabelle, mit der Sie die Gleichmäßigkeit der Verteilung der Intervallwerte überprüfen können.

73

Regressionsanalyse mit autoregressiven Fehlern

Allgemeines über die Anwendungsroutine 'Regressionsanalyse mit autoregressiven Fehlern' ... 581
Regressionsanalyse mit autoregressiven Fehlern: Zuweisen von Variablen zu Analyserollen ... 582
Regressionsanalyse mit autoregressiven Fehlern: Festlegen von Modelloptionen ... 583
Regressionsanalyse mit autoregressiven Fehlern: Festlegen von Statistikoptionen ... 584
Regressionsanalyse mit autoregressiven Fehlern: Erzeugen von Plots ... 587
Regressionsanalyse mit autoregressiven Fehlern: Festlegen von Ergebnisoptionen ... 588
Festlegen von Titeln und Fußnoten ... 590
Anzeigen von Eigenschaften ... 590

Allgemeines über die Anwendungsroutine 'Regressionsanalyse mit autoregressiven Fehlern'

Die Anwendungsroutine "Regressionsanalyse mit autoregressiven Fehlern" schätzt und prognostiziert lineare Regressionsmodelle für Zeitreihendaten, wenn die Fehler über die Zeit nicht unabhängig sind (Autokorrelation) oder die Fehlervarianz nicht konstant ist (Ungleichheit der Varianzen). Diese Anwendungsroutine erstellt Vorhersagen und Vorhersagekonfidenzgrenzen unabhängiger Variablen, die in der Eingabedatei enthalten sind. Es handelt sich dabei um ein nützliches Vorhersagetool, da es den Zeitreihenteil des Modells und den systematischen Teil beim Generieren von Prognosewerten verwendet. Dieses Modell berücksichtigt beim Erstellen der Vorhersagen die neueste Abweichung vom Trend.

Mit dieser Anwendungsroutine können folgende Aufgaben erledigt werden:

- Vergleichen von Schätzern aus verschiedenen Schätzmethoden und -modellen in einer Zeitreihe. Beispielsweise können Sie mithilfe dieser Anwendungsroutine Investitionsmodelle über einen Zeitraum hinweg vergleichen.
- Erstellen einer kurzfristigen Vorhersage. Zahlreiche Zeitreihen weisen eine hohe positive Autokorrelation auf, da sie ebenso glatt aussehen wie ein Zufallsverlauf. Diese Anwendungsroutine kann die Modellanpassung erheblich verbessern, da nicht nur zusätzliche Parameter hinzugefügt, sondern auch Random-Walk-Tendenzen erfasst werden können.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	AUTOREG
Erforderliche SAS-Produkte	Base SAS, SAS/ETS
Empfohlene zusätzliche SAS-Produkte	Ohne

Regressionsanalyse mit autoregressiven Fehlern: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Abhängige Variable	gibt die abhängige Variable für die Regressionsanalyse an. Sie müssen dieser Rolle genau eine Variable zuweisen.

Optionsname	Beschreibung
Erklärende Variablen	dienen als unabhängige Regressorvariablen für das Regressionsmodell. Wenn dieser Rolle keine Variablen zugewiesen sind, wird nur der Mittelwert angepasst. (Dies ist eine Möglichkeit, Autokorrelationen einer Reihe zu erhalten.)
Zeit-ID-Variable	gibt die Variable an, mit der den Zeilen Daten zugeordnet werden. Sie müssen dieser Rolle genau eine Variable zuweisen. Die erste Variable, die ein SAS-Datums-, Uhrzeit- oder Datetime-Format enthält, wird dieser Rolle automatisch zugewiesen. Sie können auch eine Zeit-ID-Variable erstellen. Sie müssen dieser Rolle eine Variable zuweisen, wenn Sie grafische Plots der Analysen erzeugen möchten.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits als abhängige Variable oder als unabhängige Variable ausgewählt ist.

Regressionsanalyse mit autoregressiven Fehlern: Festlegen von Modelloptionen

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Methode	gibt die Modellmethoden an. Folgende Optionen können ausgewählt werden: - **Maximum-Likelihood-Schätzer** - **Unbedingte Kleinste-Quadrate-Schätzer** - **Yule-Walker-Schätzer** - **Iterative Yule-Walker-Schätzer**

Optionsname	Beschreibung
Maximale Anzahl Iterationen	gibt die Anzahl Iterationen an. Der Standardwert ist `50`.
Nur erste Folge nichtfehlender Daten verwenden	legt fest, dass die erste zusammenhängende Folge von Daten mit nichtfehlenden Werten verwendet werden soll. Wenn Sie diese Option nicht auswählen, werden alle Beobachtungen verwendet.
Konstantenparameter im Modell einschließen	legt fest, dass der Konstantenparameter im Modell verwendet wird. Dies ist die standardmäßige Option.
Autoregressives Modell anpassen	legt fest, dass das autoregressive Modell angepasst wird. Sie können eine oder mehrere dieser Optionen auswählen. ■ Aktivieren Sie das Kontrollkästchen **Ordnung des autoregress. Prozesses** und geben Sie einen Wert in das Feld ein, oder aktivieren Sie **Lags des autoregress. Prozesses angeben** und geben Sie die Teilmenge der anzupassenden autoregressiven Fehler-Lags an. ■ Aktivieren Sie das Kontrollkästchen **Nichtsignifikante autoregressive Parameter entfernen**, um diese Parameter aus dem Modell in der Reihenfolge der geringsten Signifikanz (Rückwärtsverfahren) zu entfernen. Die Eliminierung wird nur einmal an den Yule-Walker-Schätzern nach der anfänglichen Schätzung der gewöhnlichen kleinsten Quadrate vorgenommen. ■ Geben Sie einen Wert in das Feld **Entfernen, wenn Signifikanzniveau über** ein, um das Signifikanzniveaukriterium, das für das Rückwärtsverfahren verwendet wird, anzugeben.

Regressionsanalyse mit autoregressiven Fehlern: Festlegen von Statistikoptionen

Klicken Sie im Auswahlfenster auf **Statistiken**, um auf diese Optionen zuzugreifen.

Sie können eine oder mehrere dieser Statistiken auswählen.

Optionsname	Beschreibung
Durbin-Watson-Statistiken	zeigt die Durbin-Watson-Statistiken bis zur Ordnung *n* an. Verwenden Sie die Option **Einschließen bis Ordnung**, um den Wert n anzugeben. Gültige Werte sind zwischen 1 und 32.767. Der Standardwert ist 1.
Randwahrscheinlichkeit des Durbin-Watson-Tests	zeigt die *p*-Werte für den verallgemeinerten Durbin-Watson-Test für große Stichprobengrößen an. Wählen Sie die Option **Linearisierte Approximation der Designmatrix verwenden**, um *p*-Werte zu berechnen, die eine lineare Annäherung der Design-Matrix verwenden, wenn das Modell aufgrund eines autoregressiven Fehlerprozesses nicht linear ist.
Durbin's t-Statistik	zeigt die Durbin *t*-Statistik an, die eine Residuen-Autokorrelation bei vorhandenen abhängigen Variablen mit Lag erkennt. **Hinweis:** Wenn Sie der Rolle **Erklärende Variablen** keine Variablen zugewiesen haben, können Sie nur die Durbin *t*-Statistik berechnen.
Durbins h- und t-Statistiken	ermöglicht die Auswahl, ob die Durbin *h*- oder die Durbin *t*-Statistik angezeigt werden soll. Wählen Sie in der Dropdown-Liste **Typ** die zu berechnende Statistik aus. Die Durbin *h*-Statistik sucht nach Autokorrelationen erster Ordnung, wenn die Regressoren die ausgewählte abhängige Variable mit Lag enthalten. Sie können diese Variable in der Dropdown-Liste **Abhängige Variable mit Lag** wählen. (Die Variablen in dieser Dropdown-Liste sind die Variablen, die Sie der Rolle **Erklärende Variablen** zugewiesen haben.) Kann die *h*-Statistik nicht berechnet werden, wird stattdessen die asymptotisch äquivalente *t*-Statistik angezeigt. **Hinweis:** Wenn Sie der Rolle **Erklärende Variablen** keine Variable zugewiesen haben, können Sie wählen, dass die Durbin *h*- oder die Durbin *t*-Statistik berechnet wird.
Godfreys allg. Lagrange-Multiplikator-Test gegen ARMA	zeigt Godfreys allgemeinen Lagrange-Multiplikator-Test gegen ARMA-Fehler bis zur Ordnung *r* an. Verwenden Sie die Option **Einschließen bis zu Ordnung**, um den Wert *r* anzugeben. Gültige Werte sind zwischen 1 und 32.767. Der Standardwert ist 4.
Jarque-Bera-Test auf Normalverteilung	zeigt die Statistik des Jarque-Bera-Tests auf Normalverteilung für Regressionsresiduen an.
Unit-Root-Test nach Phillips-Perron	zeigt den Unit-Root-Test nach Phillips-Perron an, wenn das Modell keine Regressoren enthält. Um die Stutzungspunkte für die gewichteten Varianzschätzer anzugeben, aktivieren Sie das Kontrollkästchen **Stutzungs-Lags einschließen?** und geben Sie die Lag-Werte in das Feld **Stutzungs-Lags** ein. Es gibt drei Typen von

Optionsname	Beschreibung
	Tests nach Phillips-Perron: Typ Zero Mean (ohne Konstante), Typ Single Mean (mit Konstante) und Typ deterministischer Trend. **Hinweis:** Wenn Sie der Rolle **Erklärende Variablen** eine Variable zugewiesen haben, enthält das Modell Regressoren und der Phillips-Ouliaris-Kointegrationstest wird stattdessen angezeigt.
Q- und LM-Statistiken für Abwesenheit von ARCH-Effekten	zeigt eine Tabelle mit Statistiken an, die einen Test auf Ungleichheit der Varianzen durchführen.
Geschätzte Kovarianzen der Parameterschätzer	zeigt die geschätzten Kovarianzen der Parameterschätzer an.
Geschätzte Korrelationen der Parameterschätzer	zeigt die geschätzten Korrelationen der Parameterschätzer an.
Partielle Autokorrelationen	zeigt partielle Autokorrelationen an.
Transformationskoeffizienten	zeigt die Transformationskoeffizienten für die ersten *p* Beobachtungen an. Diese Koeffizienten werden aus einem Skalar gebildet, der mit dem inversen der Cholesky-Wurzel der Toeplitz-Matrix an Autokovarianzen multipliziert wird.
Inverse der Toeplitz-Matrix der Autokovarianzen	zeigt die Inverse der Toeplitz-Matrix der Autokovarianzen für die Yule-Walker-Lösung an.
Log-Likelihood-Wert des Regressionsmodells	zeigt den Log-Likelihood-Wert des Regressionsmodells an, wobei in der Regel von normal verteilten Fehlern ausgegangen wird.
Nichtzentriertes	druckt die nichtzentrierte Regression R^2. Das nichtzentrierte Regressions-R^2 dient zum Berechnen der Lagrange-Multiplikator-Teststatistiken, da die meisten LM-Teststatistiken als T*URSQ

Optionsname	Beschreibung
Regressions-R-Quadrat	berechnet werden, wobei *T* der Anzahl der Beobachtungen entspricht, die in der Schätzung verwendet werden.
Iterationsverlauf der Zielfunktion und Parameterschätzer	zeigt die Zielfunktion und Parameterschätzer bei jeder Iteration an. Die Zielfunktion ist die vollständige Log Likelihood-Funktion für die Maximum Likelihood-Methode. Die Fehlersumme der Quadrate ergibt sich als Zielfunktion der unbedingten kleinsten Quadrate. Für die Maximum-Likelihood-Methode zeigt diese Option den Wert der vollständigen Log-Likelihood-Funktion und nicht die konzentrierte Likelihood an.
Ramseys RESET-Test	zeigt die Ramseys RESET-Teststatistik an. **Hinweis:** Diese Option ist nur verfügbar, wenn Sie der Rolle **Erklärende Variablen** eine Variable zugewiesen haben.

Regressionsanalyse mit autoregressiven Fehlern: Erzeugen von Plots

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Standardmäßig werden alle für die aktuelle Datenauswahl geeigneten Plots in der Ausgabe berücksichtigt. Sie können jedoch durch Auswahl der Option **Benutzerdefinierte Plot-Listen** festlegen, welche Plots Sie in die Ausgabe aufnehmen möchten. Folgende Optionen können ausgewählt werden:

Optionsname	Beschreibung
Autokorrelationsfunktions-Plot	erstellt einen Plot der Autokorrelationsfunktion.
Inverse-Autokorrelationsfunktion-Plot	erstellt einen Plot der inversen Autokorrelationsfunktion.
Partielle-Autokorrelationsfunktion-Plot der Residuen	erstellt einen Plot der partiellen Autokorrelationsfunktion.

Optionsname	Beschreibung
Plot mit vorhergesagten und tatsächlichen Werten	erstellt einen Plot der vorhergesagten gegen beobachtete Werte.
Cook's D-Statistik-Plot	erstellt einen Plot der Cook's *D*-Statistik gegen die Beobachtungszahl. **Hinweis:** Diese Option ist nicht verfügbar, wenn Sie im Fenster **Optionen** das Kontrollkästchen **Autoregressives Modell anpassen** aktiviert haben.
Q-Q-Plot der Residuen	erstellt ein Normal-Quantil-Plot der Residuen.
Plot der Residuen	erstellt einen Plot der Residuen gegen vorhergesagte Werte. Ein Residuum ist die Differenz zwischen dem beobachteten und dem vorhergesagten Wert.
Plot der studentisierten Residuen	erstellt einen Plot der studentisierten Residuen gegen vorhergesagte Werte. **Hinweis:** Diese Option ist nicht verfügbar, wenn Sie im Fenster **Optionen** das Kontrollkästchen **Autoregressives Modell anpassen** aktiviert haben.
Plot der Wahrscheinlichkeiten für weißes Rauschen	erstellt einen Plot der Wahrscheinlichkeiten für weißes Rauschen.
Plot des Histogramms der Residuen	erstellt ein Histogramm der Anpassungsresiduen.

Regressionsanalyse mit autoregressiven Fehlern: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Statistiken einschließen	speichert die Statistiken für Beobachtungsausprägungen. Folgende Optionen können ausgewählt werden: ■ **Vorhergesagte Werte, Residuen und Konfidenzgrenzen für den unbedingten Mittelwert** speichert die vorhergesagten Werte, Residuen, sowie die oberen und unteren Konfidenzwerte, die sich nur aus dem strukturellen Teil des Modells ergeben. Verwenden Sie die Dropdown-Liste **Konfidenzniveaus**, um die Konfidenzgrenzen für die Schätzer des strukturellen Modellteils anzugeben. ■ **Vorhergesagte Werte, Residuen und Konfidenzgrenzen für zukünftige Reihenwerte** speichert vorhergesagte Werte, Residuen sowie die oberen und unteren Konfidenzgrenzen, die durch die strukturellen und autoregressiven Teile des Modells gebildet werden. Verwenden Sie die Dropdown-Liste **Konfidenzniveaus**, um die Konfidenzgrenzen für die Schätzer der zukünftigen Werte der abhängigen Zeitreihe anzugeben. ■ **Kumulative Summe (CUSUM) und Konfidenzniveaus** speichert die CUSUM-Statistik, CUSUMSQ-Statistik und deren obere und untere Konfidenzgrenzen. Verwenden Sie die Dropdown-Liste **Konfidenzniveaus**, um das Signifikanzniveau für die obere und die untere Konfidenzgrenze zu setzen. ■ **Rekursive und BLUS-Residuen** speichert die Werte der Theils BLUS-Residuen, der rekursiven Residuen, die zum Berechnen der CUSUM- und CUSUMSQ-Statistiken verwendet werden, und den Teil der vorhersagenden Fehlervarianz, (vt), die zum Berechnen der rekursiven Residuen verwendet werden.
Parameterschätzer einschließen	erstellt eine Datei mit allen im Modell verwendeten Variablen. Jede Regressorvariable enthält den Schätzer für den entsprechenden Regressionsparameter. Um die Kovarianzmatrix für die Parameterschätzer einzuschließen, aktivieren Sie das Kontrollkästchen **Kovarianzmatrix einschließen**.
Anzeige für gesamte Ergebnisausgabe unterdrücken	zeigt keine Ergebnisse an. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

74

Regressionsanalyse von Paneldaten

Allgemeines über die Anwendungsroutine 'Regressionsanalyse von Paneldaten' **593**
Regressionsanalyse von Paneldaten: Zuweisen von Variablen zu Analyserollen **594**
Regressionsanalyse von Paneldaten: Festlegen von Modelloptionen **595**
Regressionsanalyse von Paneldaten: Festlegen von Ergebnisoptionen **596**
Festlegen von Titeln und Fußnoten **597**
Anzeigen von Eigenschaften **597**

Allgemeines über die Anwendungsroutine 'Regressionsanalyse von Paneldaten'

Die Anwendungsroutine **Regressionsanalyse von Paneldaten** analysiert eine Klasse linearer, ökonometrischer Modelle, die in der Regel vorliegen, wenn Zeitreihen und Querschnittdaten kombiniert werden. Diese Anwendungsroutine verwendet Paneldateien, die aus Zeitreihenbeobachtungen in den verschiedensten Querschnitteinheiten bestehen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	TSCSREG

Name der Anforderung	Prozedur- und Produktnamen
Erforderliche SAS-Produkte	Base SAS, SAS/ETS
Empfohlene zusätzliche SAS-Produkte	Ohne

Regressionsanalyse von Paneldaten: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Abhängige Variable	gibt die Variable für die Regression an. Sie müssen dieser Rolle genau eine Variable zuweisen.
Erklärende Variable	gibt die unabhängigen Variablen für die Regression der abhängigen (Response) Variablen an. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Querschnitts-ID-Variable	gibt die Variable an, welche die Querschnitte definiert. Sie müssen dieser Rolle genau eine Variable zuweisen. Die Daten werden automatisch nach den Werten dieser Variablen sortiert.
Zeit-ID-Variable	gibt die Variable an, mit der den Beobachtungen in jedem Querschnitt Zeitperioden zugewiesen werden. Sie müssen dieser Rolle genau eine Variable zuweisen. Die Daten werden automatisch nach den Werten dieser Variablen sortiert.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert.

Regressionsanalyse von Paneldaten: Festlegen von Modelloptionen

Diese Optionen werden im Auswahlfenster unter der Überschrift **Optionen** angezeigt.

Plot-Name	Beschreibung
Zu verwendende Schätzmethoden	Um Schätzmethoden anzugeben, wählen Sie die Optionsfelder für die gewünschten Methoden aus. Die Performance aller Schätzmethoden für die Modellregressionsparameter hängt von den statistischen Eigenschaften der Fehlerkomponenten im Modell ab. Der Standard ist **Zufällige Zweiwegeffekte**. ■ Ein Modell mit **festen Einwegeffekten** liegt vor, wenn die Spezifikation nur vom Querschnitt abhängt, zu dem die Beobachtung gehört. ■ Ein Modell mit **festen Zweiwegeffekten** liegt vor, wenn die Spezifikation vom Querschnitt und der Zeitreihe abhängt, zu der die Beobachtung gehört. ■ Ein Modell mit **zufälligen Einwegeffekten** liegt vor, wenn die Spezifikation nur vom Querschnitt abhängt, zu dem die Beobachtung gehört. In diesem Fall sind die Effekte zufällig. Für ein ausgeglichenes Panel wird standardmäßig die Fuller-und-Battese-Methode verwendet. Für ein unausgeglichenes Panel wird standardmäßig die Wansbeek-und-Kapteyn-Methode verwendet. ■ Ein Modell mit **zufälligen Zweiwegeffekten (Varianzkomponentenmodell)** liegt vor, wenn die Spezifikation nur vom Querschnitt und dem Zeitraum abhängt, zu dem die Beobachtung gehört. In diesem Fall sind die Effekte zufällig. Für ein ausgeglichenes Panel wird standardmäßig die Fuller-und-Battese-Methode verwendet. Für ein unausgeglichenes Panel wird standardmäßig die Wansbeek-und-Kapteyn-Methode verwendet. ■ Bei einem **autoregressiven Modell (Parks-Methode)** wird von einer autoregressiven Fehlerstruktur erster Ordnung mit kontemporärer Korrelation zwischen den Querschnitten ausgegangen. Wenn Sie die Parks-Methode auswählen, können Sie die fehlenden Werte in der Zeitreihe mit einer kubischen Splinekurve interpolieren. Ein kubischer Spline ist eine segmentierte Funktion, die aus polynomialen Funktionen

Plot-Name	Beschreibung
	dritten Grades (kubisch) besteht, die so gejoint wurden, dass die gesamte Kurve und ihre ersten und zweiten Ableitungen kontinuierlich sind. Eine kubische Splinekurve wird an die nichtfehlenden Werte der Variablen angepasst. Diese Kurve dient zum Bilden der kontinuierlichen Zeit-Approximationen für die Eingabereihen, und die Ausgabezeitreihen werden aus den Spline-Approximationen generiert. Das Kontrollkästchen **Kubische Splineinterpolation der fehlenden Werte** ist standardmäßig aktiviert. **Hinweis:** Vorsicht beim Verwenden der Option **Kubische Splineinterpolation der fehlenden Werte**. Interpolation führt oft zu einer Verzerrung im Modell, die zu schlechten Parameterschätzern und Hypothesentests führen kann. Weitere Informationen zu diesen Methoden, siehe *SAS/ETS: User's Guide*.
Konstantenparameter im Modell einschließen	Um den Konstantenparameter aus dem Modell auszuschließen, deaktivieren Sie das Kontrollkästchen für diese Option. Standardmäßig ist das Kontrollkästchen markiert.

Regressionsanalyse von Paneldaten: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Um die Parameterschätzer in einer Ausgabetabelle zu speichern, markieren Sie das Kontrollkästchen **Parameterschätzer speichern**. Wenn gewünscht, markieren Sie die Kontrollkästchen, um die **Kovarianzmatrix**, **Korrelationsmatrix** oder beides einzuschließen.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Um die Ergebnisse anzuzeigen, aktivieren Sie das Kontrollkästchen **Analyse anzeigen**. Um die Ergebnisse zu unterdrücken, deaktivieren Sie dieses Kontrollkästchen. Die Ergebnisse werden standardmäßig angezeigt.

Wenn Sie Ihre Ergebnisse anzeigen möchten, aktivieren Sie die Kontrollkästchen für die Elemente, die in Ihren Ergebnissen enthalten sein sollen. Einige Elemente sind nur verfügbar, wenn Sie die Parks-Schätzmethode ausgewählt haben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine

- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

75

SAS ODS Statistical Graphics

Allgemeines zu "SAS ODS Statistical Graphics" **599**
Verwenden des ODS Graphics Designer **600**
ODS Statistical Graph anzeigen **600**

Allgemeines zu "SAS ODS Statistical Graphics"

SAS ODS Statistical Graphics oder auch SAS ODS Graphics sind eine Erweiterung von SAS Output Delivery System (ODS). Das ODS verwaltet jegliche Form von Ausgabe, die durch Prozeduren erstellt wird, und ermöglicht es zudem diese Ausgabe auf verschiedenste Weise anzuzeigen, inkl. HTML und PDF.

In vielen analytischen SAS-Prozeduren wird die Funktionalität von ODS-Grafiken verwendet, um gleichzeitig und automatisch sowohl Grafiken als auch Tabellen zu erzeugen, die von diesen Prozeduren generiert werden. ODS-Grafiken basieren auf der Syntax der Graph Template Language (GTL), die es ermöglicht eine große Anzahl komplexer Grafiken zu erstellen. GTL ist eine universelle Sprache zur Erstellung und Definition statistischer Grafiken.

Im SAS Enterprise Guide können Sie mithilfe des ODS Graphics Designer statistische Graphen definieren, ohne dass dazu Kenntnisse der GTL erforderlich wären. Im Anschluss an die Definition eines Graphen wird dieser erstellt. Mit dieser Grafikdefinition können Sie anschließend in SAS Enterprise Guide oder SAS Add-In for Microsoft Office eine ODS Statistical Graph (statistische ODS-Grafik) erzeugen.

Verwenden des ODS Graphics Designer

Hinweis: Der SAS ODS Graphics Designer ist in SAS Add-In for Microsoft Office nicht verfügbar. Sie können jedoch SGD-Dateien, die mit dem SAS ODS Graphics Designer erstellt wurden, im SAS Add-In öffnen.

Der SAS ODS Graphics Designer ist eine interaktive graphische Anwendung für die Erstellung und das Design von benutzerdefinierten Grafiken. Der Designer erstellt Grafiken, die auf der Graph Template Language (GTL) basieren. Dies ist dieselbe Sprache, die auch in analytischen SAS-Prozeduren und SAS ODS Graphics-Prozeduren verwendet wird. Mit der grafischen Benutzeroberfläche des ODS Graphics Designer lassen sich schnell Grafiken erstellen, ohne dass dafür tiefere Kenntnisse über Vorlagen und GTL erforderlich sind.

Hinweis: Ab dem ersten Maintenance Release von SAS 9.4 können Sie den SAS ODS Graphics Designer in SAS Enterprise Guide verwenden, ohne dass dazu eine Installation von Base SAS auf Ihrem lokalen Rechner installiert sein muss.

Mittels Point-and-Click können Sie einfache oder komplexe grafische Ansichten für die Analyse von Daten erstellen. Darüber hinaus stehen im ODS Graphics Designer eine Vielzahl von Plottypen zur Verfügung, mit denen Sie die anspruchsvollsten Grafiken kreieren können. Mehrzellige Grafiken, Klassifizierungs-Panels sowie Streuungsplot-Matrizen lassen sich generieren. Ebenso können Sie die Grafiken mit Titeln, Fußnoten, Legenden und weiteren grafischen Elementen versehen. Ergebnisse können als Bild gespeichert werden, so dass Sie sie in einen Bericht miteinbeziehen oder als ODS Graphics Designer-Datei (SGD) ablegen und später bearbeiten können. Weitere Informationen, siehe “ODS Statistical Graph anzeigen” auf Seite 600.

Um den ODS Graphics Designer in SAS Add-In for Microsoft Office oder SAS Enterprise Guide zu öffnen, doppelklicken Sie im Fensterabschnitt **Anwendungsroutinen** auf **ODS Statistical Graph anzeigen**. Weitere Informationen, siehe *SAS Output Delivery System: User’s Guide*.

ODS Statistical Graph anzeigen

Die Anwendungsroutine **ODS Statistical Graph anzeigen** steht sowohl im SAS Enterprise Guide als auch in SAS Add-In for Microsoft Office zur Verfügung. In dieser Anwendungsroutine wird eine Grafikdefinitionsdatei (SGD-Datei) ausgewählt, die für die ausgewählte Datenquelle verwendet wird. Eine

Grafikdefinition können Sie im SAS Enterprise Guide erstellen. Weitere Informationen, siehe “Verwenden des ODS Graphics Designer” auf Seite 600.

Eine erstellte Grafikdefinition kann in lokalen und Remote-SAS-Sitzungen verwendet werden. Darüber hinaus können Sie diese Grafiken in SAS Stored Processes darstellen.

So verwenden Sie eine Grafikdefinition (SGD-Datei) zur Erstellung einer ODS Statistical Graph (statistischen ODS-Grafik):

1. Doppelklicken Sie im Fensterabschnitt **Anwendungsroutinen** in SAS Add-In for Microsoft Office oder SAS Enterprise Guide auf **ODS Statistical Graph anzeigen**.

   Das Dialogfeld **ODS Statistical Graphics anzeigen** wird geöffnet.

2. Wählen Sie die SGD-Datei aus, in der das gewünschte Design enthalten ist.
3. Wählen Sie die Datenquelle für die Grafik aus.
4. (Optional) Legen Sie Breite und Höhe des Diagramms fest.
5. (Optional) Geben Sie einen Titel und eine Fußnote an.
6. Klicken Sie auf **Ausführen**.

Weitere Informationen zu Diagrammdefinitionen finden Sie in *SAS Output Delivery System: User’s Guide*.

76

SAS Rapid Predictive Modeler

Allgemeines über den SAS Rapid Predictive Modeler ... **603**
Sampling-Strategien für den SAS Rapid Predictive Modeler ... **604**
SAS Rapid Predictive Modeler: Organisieren von Daten ... **605**
SAS Rapid Predictive Modeler: Zuweisen von Variablen zu Analyserollen ... **607**
SAS Rapid Predictive Modeler: Auswählen eines Modells ... **608**
Festlegen von Modelloptionen ... 608
Grundlegendes über die Modelle für den SAS Rapid Predictive Modeler ... 610
SAS Rapid Predictive Modeler: Festlegen von Berichtoptionen ... **613**
Berichtoptionen ... 613
Allgemeines über Standardberichte für den SAS Rapid Predictive Modeler ... 614
SAS Rapid Predictive Modeler: Festlegen von Ausgabeoptionen ... **615**
SAS Rapid Predictive Modeler: Registrieren Ihres Modells ... **616**
Informationen zur Registrierung eines Modells ... 616
Registrieren eines Modells im SAS Metadata Repository ... 617

Allgemeines über den SAS Rapid Predictive Modeler

SAS Rapid Predictive Modeler wurde für die Erstellung von Modellen für die folgenden Arten von Data-Mining-Klassifizierungs- und Regressionsproblemen entwickelt:

- Klassifizierungsmodelle, mit denen der Wert einer diskreten Variable prognostiziert wird. Beispiele sind Klassifizierungsmodelle, die den Wert einer Variablen prognostizieren, etwa Wahr oder Falsch, Zuwachs oder Rückgang, Hoch, Mittel oder Niedrig und Abbruch oder Fortsetzen.
- Regressionsmodelle, mit denen der Wert einer reellwertigen Variable prognostiziert wird. Beispiele sind Regressionsmodelle, mit denen eine

Quantität prognostiziert wird, etwa ein Ertrag, ein Umsatz oder eine Erfolgsquote.

Um ein Modell mit dem SAS Rapid Predicitve Modeler zu erstellen, ist lediglich eine Datei erforderlich, in der jede Zeile eine Menge unabhängiger Prädiktorvariablen (auch Eingänge genannt) und mindestens eine abhängige Variable (auch Ziel genannt) enthält. SAS Rapid Predictive Modeler entscheidet automatisch, ob es sich um eine kontinuierliche oder eine kategoriale Variable handelt, und wählt die Eingangsvariablen aus, die im Modell verwendet werden sollen. Weitere Informationen über das Sampling von Daten mit dieser Anwendungsroutine finden Sie unter Sampling-Strategien für den SAS Rapid Predictive Modeler.

Das Modell kann als SAS-Code gespeichert und anschließend in einer SAS-Umgebung bereitgestellt werden. Mit dem SAS-Code für das Modell können Sie Scorewerte für neue Daten berechnen und diese Ergebnisse anschließend in eine fundierte Geschäftsprozessentscheidung einfließen lassen. Dieser Prozess wird als Modell-Scoring bezeichnet. Es gibt eine Vielzahl von Anwendungsbeispielen für diese Scoring-Daten. Mit gescorten Daten können Sie entscheiden, für welche Kunden eine bestimmtes Angebot attraktiv ist, Sie können Bedarfe in der Produktion, im Einkauf oder in der Logistik vorhersagen, Sie können das Abwanderungsverhalten von Kunden ermitteln oder die Daten auch für die Betrugserkennung einsetzen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	
Erforderliche SAS-Produkte	SAS Enterprise Miner
Empfohlene zusätzliche SAS-Produkte	

Sampling-Strategien für den SAS Rapid Predictive Modeler

Im SAS Rapid Predictive Modeler kommt ein zusammengesetzter Sampling-Ansatz zum Einsatz. Die Anzahl der Beobachtungen, die der Stichprobe enthalten sind, ist von folgenden Faktoren abhängig:

- Anzahl der Eingangsvariablen
- in der Datenquelle
- Vorkommen von seltenen Zielereignissen in den Daten

- Anzahl der Ereignisse in den Daten

Die Anzahl der zu verarbeitenden Beobachtungen wird im SAS Rapid Predictive Modeler mit folgenden Kriterien bestimmt:

Anzahl der Eingangsvariablen	Anzahl der verarbeiteten Beobachtungen
< 100	80.000
100 – 200	40.000
> 200	20.000

Bedingung	Seltenes Ereignis: Ja	Seltenes Ereignis: Nein
< Anzahl der verarbeiteten Beobachtungen ODER Gesamtanzahl der Ereignisse < (0,10*Anzahl der verarbeiteten Beobachtungen)	Sampling der Daten, so dass das Verhältnis von Nicht-Ereignissen zu Ereignissen 10:1 beträgt	Kein Sampling
Gesamtanzahl der Ereignisse > (0.10*Anzahl der verarbeiteten Beobachtungen)	Sampling des folgenden Verhältnisses der seltenen Ereignisse: 10*[(0,10*(Anzahl der verarbeiteten Beobachtungen))/ Anzahl Ereignisse]	Geschichtete Stichproben

SAS Rapid Predictive Modeler: Organisieren von Daten

Bevor Sie ein Modell erstellen können, benötigen Sie Eingabedaten, die historische Ereignisse und Eigenschaften aufweisen, die für eine Prognose verwendet werden können. Darüber hinaus sind Zieldaten erforderlich, die das Ereignis oder den Wert darstellen, der prognostiziert werden soll. Häufig werden die Eingabedaten aus einem Zeitraum und die Zieldaten aus einem späteren Zeitraum abgeleitet. Die Kombination aus Eingabe- und Zieldaten, die Sie in Ihrem Modell verwenden, wird als Trainingsdaten bezeichnet.

Beispiel: Sie möchten aus den Umsätzen des vergangenen Jahres die Umsatzerlöse für das nächste Jahr prognostizieren oder etwa vorhersagen, welche Kunden

Interesse an einem bestimmtes Angebot haben könnten. Die Verwendung von historischen Daten mit zurückliegenden Ereignissen zur Prognose von künftigen Ereignissen wird als Modelltraining bezeichnet.

Die besten Modellergebnisse lassen sich erzielen, wenn die Modelltrainingdaten eine große Anzahl an Beobachtungen enthalten, die in Zeilen gespeichert vorliegen. Für Kunden im Einzelhandelsbereich werden beispielsweise häufig Eingabedaten verwendet, die mehrere Zehntausend Beobachtungen enthalten.

Wenn die gewählte Zielvariable ein seltenes Ereignis enthält (etwa ein Angebot, das nur für ca. 1% Ihrer Kunden interessant ist), muss sichergestellt sein, dass die Trainingsdaten eine signifikante Anzahl dieser Kunden enthalten. Durch Oversampling der Trainingsdaten können alle Kunden, die das Angebot angenommen haben ausgewählt werden und darüber hinaus die gleiche Anzahl an Kunden angegeben werden, die das Angebot nicht angenommen haben. Oversampling ermöglicht es Ihnen, eine stabile Lösung für ein Modell mit einem seltenen Zielereignis zu finden.

Durch das Oversampling von seltenen Ereignisse in Ihren Trainingsdaten, wird die Anzahl der Ziel-Ereignisse in Ihren Trainingsdaten im Vergleich zu deren tatsächlichem Vorkommen künstlich erhöht. Um diese Differenz zwischen Vorkommen in den Trainingsdaten und in der Grundgesamtheit ausgleichen zu können, verfügt der SAS Rapid Predictive Modeler über eine Einstellung für A-Priori-Wahrscheinlichkeiten. Mit diesen A-Priori-Wahrscheinlichkeiten wird das tatsächliche proportionale Verhältnis des Zielereignisses in der Grundgesamtheit festgelegt.

Die mit dem SAS Rapid Predictive Modeler untersuchten Daten sollten in Zeilen (Beobachtungen) und Spalten (Variablen) angeordnet vorliegen. Eine der Spalten sollte eine Zielvariable darstellen. Dazu folgendes Beispiel:

Name	Alter	Geschlecht	Einkommen	Behandlung	Kauf
Richard	29	M	33.000	J	J
Susanne	35	W	51.000	J	N
Andreas	49	M	110.000	N	J

Name
: ist eine Spalte, die ID-Werte für jede einzelne Beobachtung enthält. ID-Variablenspalten werden vom SAS Rapid Predictive Modeler nicht hinsichtlich ihres analytischen Inhalts verarbeitet.

Alter, Geschlecht, Einkommen und Behandlung
: sind Eingabespalten. Sie werden vom SAS Rapid Predictive Modeler verwendet.

Kauf
: ist eine Zielspalte.

Beim Konfigurieren der Tabelle mit Eingangsdaten können Sie eine Häufigkeitszählungs-Spalte (oder auch Häufigkeitsspalte) bestimmen. Die sogenannten Häufigkeitsvariablen stellen die relative Gewichtung dar, die einer Zeile (Beobachtung) zugewiesen werden soll. Beispielsweise ist es möglich, dass in bestimmten Dateien, eine einzelne Zeile mehr als nur eine Beobachtung darstellt.

Es ist auch möglich, Spalten auszuwählen, die bei der Analyse mit dem SAS Rapid Predictive Modeler nicht beachtet werden.

Trainingsdaten müssen immer Werte für Eingangs- und Zielvariablen enthalten. Daten, die Sie lediglich zum Scoring einsetzen, müssen nur Werte für Eingangsvariablen enthalten; eine Zielspalte ist hierfür optional. Wird das Modell zum Erstellen von Prognosen auf Basis neuer Daten verwendet, ist eine Zielspalte nicht erforderlich. Wird das Modell verwendet, um die Leistungsfähigkeit zu überwachen, ist eine Zielspalte erforderlich. Darüber hinaus enthalten Daten, die zum Scoring eingesetzt werden, in der Regel eine ID-Spalte.

SAS Rapid Predictive Modeler: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Welche Analyserollen zur Verfügung stehen, ist vom zu analysierenden Datentyp abhängig.

Optionsname	Beschreibung
Abhängige Variable	legt den Wert fest, für den Sie eine Prognose oder Klassifizierung erstellen möchten. Die abhängige Variable wird auch als Zielvariable bezeichnet. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Häufigkeitszählung	legt die Variable fest, die verwendet wird, um die Gewichtungen darzustellen, die Sie auf die Datenzeilen anwenden. Sie können dieser Rolle nur eine Variable zuweisen.
ID	legt die Variablen fest, die für Auswertungen und Scoring-Auswahl von Nutzen sein können. Diese Variablen werden nicht in die Analyse einbezogen.
Exklusive	legt die Variablen fest, die nicht in der Analyse berücksichtigt werden sollen.

SAS Rapid Predictive Modeler: Auswählen eines Modells

Festlegen von Modelloptionen

Klicken Sie im Auswahlfenster auf **Modell**, um auf diese Optionen zuzugreifen.

Mit diesen Optionen können Sie den Grad der Komplexität für das zu erstellende Modell angeben. Diese Modellierungsgrade bauen hierarchisch aufeinander auf: Die mittlere Methode enthält neben den mittleren Modellen auch die einfachen Modelle. Erweitert schließt sowohl einfache als auch mittlere Modelle ein.

Nach dem oben beschriebenen einfachen Verfahren erstellte Modelle werden in der Regel in kürzerer Zeit erstellt und ausgeführt als Modelle, die nach dem mittleren Modellierungsverfahren erstellt wurden; sie sind jedoch gewöhnlich auch weniger genau. Dies gilt üblicherweise auch, wenn man nach dem mittleren Modellierungsverfahren erstellte Modelle mit erweiterten Modellen vergleicht. Weitere Informationen über die Data Mining-Operationen für die einzelnen Modelle, siehe “Grundlegendes über die Modelle für den SAS Rapid Predictive Modeler” auf Seite 610.

Der SAS Rapid Predictive Modeler führt die Modellierungsfunktionen des SAS Enterprise Miner aus. Welche Funktionen ausgeführt werden, ist abhängig von der gewählten Modellierungsmethodik.

Klicken Sie für die abhängige Variable auf **Entscheidungen und A-Priori**, um folgende Informationen anzugeben:

Tabelle 76.1 *Entscheidungen und A-Priori-Optionen für die abhängige Variable*

Optionsname	Beschreibung
Ereignisausprägung	gibt das zu modellierende Klassenzielereignis an. SAS Rapid Predictive Modeler erstellt ein Modell, das für jedes Zielereignis Wahrscheinlichkeiten angibt. Die Ergebnisse des Rapid Predictive Modeler lassen sich darüber hinaus noch verbessern, wenn die gewünschte Zielausprägung bekannt ist.
Priori-Wahrscheinlichkeiten	zeigt die Anzahl sowie die Anteile der einzelnen Zielvariablenausprägungen an, die in den Trainingsdaten für das Modell enthalten sind. Sie können diese Werte anpassen, wenn Ihre Zielvariable eine kategorische Variable ist und die Trainingsdaten sowie die Daten der Grundgesamtheit unterschiedliche Zielverteilungen aufweisen. Beispiel: Sie

Optionsname	Beschreibung
	haben ein Modell untersucht, das auf Oversampling-Daten beruht. Bei 50% der Beobachtungen handelt es sich um Kunden, die reagiert haben, und bei 50% der Beobachtungen um Kunden, die nicht reagiert haben. In der Grundgesamtheit, für die das Modell Scoring-Ergebnisse berechnet, betrug der Anteil der Reagierenden lediglich 10% und der der Nicht-Reagierenden 90%. Mit den A-Priori-Wahrscheinlichkeiten können Sie dem Modell das historisch zu erwartende Verhältnis von Reagierenden zu Nicht-Reagierenden "mitteilen". ■ Wenn Sie keine A-Priori-Wahrscheinlichkeiten angeben möchten, lassen Sie die Standardeinstellung **Keine** unverändert. ■ Wenn die Wahrscheinlichkeiten für alle Ausprägungen der Zielvariable gleich groß sein sollen, wählen Sie **Gleich**. ■ Um Ihre eigenen benutzerdefinierten A-Priori-Wahrscheinlichkeiten für die Ausprägungen der Zielvariable in den gescorten Daten anzugeben, wählen Sie **Benutzerdefiniert** und geben Sie die gewünschten Wahrscheinlichkeiten ein. Achten Sie darauf, dass die Werte in der Summe stets 1 ergeben.
Entscheidungsfunktion	legt die Kosten, Profite oder Gewichte fest, die Sie den prognostizierten Ergebnissen zuordnen möchten. Die Wertetabelle wird als Entscheidungsmatrix bezeichnet. Mit einer Entscheidungsmatrix können Sie einen Wert jedem möglichen Entscheidungsergebnis zuordnen. ■ Wenn für Ihr Modell keine Entscheidungsmatrix erforderlich ist, lassen Sie die Standardeinstellung auf **Keine**. ■ Wenn Sie Ihr Modell zur Maximierung eines Gewinns verwenden möchten, wählen Sie **Maximum** und geben Sie ggf. einen höheren Wert für die Gewichtung in der Richtig-Positiv-Zelle der Matrix ein. ■ Wenn Sie Ihr Modell zur Minimierung von Kosten verwenden möchten, wählen Sie **Minimum** und geben Sie ggf. einen höheren Wert für die Gewichtung in der Richtig-Negativ-Zelle der Matrix ein. ■ Wenn Sie mit Ihrem Modell Prognosen von seltenen Ereignissen erstellen möchten, wählen Sie **Inverse**, um Richtig-Positive und Richtig-Negative-Prognosen zu identifizieren. Falsch-Positive- und Falsch-Negative-Prognosen werden dabei ggf. nicht korrekt geschätzt.

Tabelle 76.2 Modellerstellungsoptionen

Optionsname	Beschreibung
Einfach	legt den Wert fest, für den Sie eine Prognose oder Klassifizierung erstellen möchten. Die abhängige Variable wird auch als Zielvariable bezeichnet. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Mittel	legt die Variable fest, die verwendet wird, um die Gewichtungen darzustellen, die Sie auf die Datenzeilen anwenden. Sie können dieser Rolle nur eine Variable zuweisen.
Erweitert	legt die Variablen fest, die für Auswertungen und Scoring-Auswahl von Nutzen sein können. Diese Variablen werden nicht in die Analyse einbezogen.

Grundlegendes über die Modelle für den SAS Rapid Predictive Modeler

In SAS Rapid Predictive Modeler stehen drei unterschiedliche Methodiken zur Modellerstellung zur Verfügung: einfache, mittlere und erweiterte Modelle. Durch die unterschiedliche, ansteigende Komplexität der drei verschiedenen Ansätze ist es möglich, das jeweils am besten geeignete Modellierungsverfahren für eine bestimmte Situation auszuwählen.

- Das einfache Modellierungsverfahren stellt eine einfache Regressionsanalyse dar.
- Bei den mittleren Modellen werden neben der einfachen Analyse auch eine komplexere Analyse durchgeführt. Am Ende wird das am besten geeignete Modell aus beiden Ansätzen ausgewählt.
- Erweiterte Modelle enthalten die beiden zuvor genannten Modellierungsmethodiken und bieten darüber hinaus eine nochmal komplexere Analyse als dritten Ansatz.

Einfaches Modell

Bei einfachen Modellen werden drei aufeinanderfolgende Data Mining-Operationen durchgeführt.

- Variablenauswahl: Beim einfachen Modell werden die Top-100-Variablen zur Modellierung ausgewählt.

- Transformation: Bei einfachen Modellen wird für die zur Modellierung ausgewählten Top-100-Variablen eine sogenannte Optimal-Binning-Transformation durchgeführt. Bei der Optimal-Binning-Transformation werden fehlende Variablenwerte ausgeglichen, sprich es wird keine Ersetzung fehlender Werte vorgenommen.
- Modellerstellung: Bei einfachen Modellen wird eine schrittweise Vorwärtsregression eingesetzt. Dabei wird zunächst jede Variable schrittweise nacheinander ausgewählt und zur linearen Gleichung hinzugefügt, bis der Variablenbeitrag nicht mehr signifikant ist. Die Vorwärtsregression versucht somit, Variablen aus der Analyse auszuschließen, die keine Bedeutung für die Prognose aufweisen (oder Variablen, die mit anderen Prädiktorvariablen stark korrelieren).

Mittleres Modell

Bei mittleren Modellen werden sieben aufeinanderfolgende Data Mining-Operationen durchgeführt.

- Variablenauswahl: Beim mittleren Modell werden die Top-200-Variablen zur Modellierung ausgewählt.
- Transformation: Bei mittleren Modellen wird für die zur Modellierung ausgewählten Top-200-Variablen eine sogenannte Best-Power-Transformation durchgeführt. Best-Power-Transformationen zählen zu den Box-Cox-Transformationen. Die Best-Power-Transformation bewertet eine Teilmenge der exponentiellen Power-Transformationen und wählt die Transformation aus, die das beste Ergebnis für das angegebene Kriterium liefert.
- Imputation: Bei mittleren Modellen wird eine Imputation vorgenommen, um fehlende Variablen durch durchschnittliche Variablenwerte zu ersetzen. Dabei werden auch Indikatorvariablen erstellt, mit deren Hilfe Beobachtungen erkannt werden können, die ersetzte Variablenwerte enthalten.
- Variablenauswahl: Bei mittleren Modellen werden Chi-Quadrat- und R-Quadrat-Tests verwendet, um Variablen zu entfernen, die keine Beziehung zur Zielvariablen aufweisen.
- Kombination von Variablenselektionstechniken: Bei mittleren Modellen werden die durch die Chi-Quadrat- und R-Quadrat-Tests ausgewählten Variablen zusammengeführt.
- Modellerstellung: Bei mittleren Modellen werden die Trainingsdaten an drei konkurrierende Modellierungsalgorithmen weitergeleitet. Bei den Modellen handelt es sich um einen Enscheidungsbaum, eine logistische Regression und eine schrittweise Regression. Bei der logistischen Regression werden die Trainingsdaten zunächst an einen Entscheidungsbaum übergeben, in dem die Variable NODE_ID erstellt wird, die wiederum als Eingabe an das Regressionsmodell weitergegeben wird. Die Variable NODE_ID ermöglicht es, Modelle mit Variablen-Interaktionen zu erstellen.
- Auswahl des Champion-Modells: Bei mittleren Modellen wird eine analytische Bewertung der Prognose- oder Klassifizierungs-Performance der verschiedenen Modelle durchgeführt. Das Modell, das die beste Prognose- oder

Klassifizierungs-Performance aufweist, wird zur Durchführung der Modellierungsanalyse herangezogen. Bei der Selektion des besten Modells werden nicht nur ausschließlich die mittleren Modelle untersucht, sondern auch die Modelle, die nach der einfachen Modellierungsmethode erstellt wurden.

Zunächst wird das beste mittlere Modell ausgesucht. Anschließend wird die Prognose-Performance dieses Modells mit der des einfachen Modells verglichen. Am Ende wird das bessere der beiden Modell verwendet.

Erweitertes Modell

Bei erweiterten Modellen werden sieben aufeinanderfolgende Data Mining-Operationen durchgeführt.

- Variablenauswahl: Beim erweiterten Modell werden die Top-400-Variablen zur Modellierung ausgewählt.
- Transformation: Bei erweiterten Modellen wird für die 400 Variablen, die zur Modellierung ausgewählt wurden, der Transformationsalgorithmus "Multiple" durchgeführt. Dabei werden verschiedene Variablentransformationen erstellt, die in nachfolgenden Variablenselektionen verwendet werden. Die Zahl der Eingangsvariablen erhöht sich durch die Verwendung von "Multiple"-Transformationen. Aus diesem Grund werden die besten 400 Eingangsvariablen aus dem Ergebnis des Multiple-Transformationsalgorithmus ausgewählt.
- Imputation: Bei erweiterten Modellen wird eine Imputation vorgenommen, um fehlende Variablen durch durchschnittliche Variablenwerte zu ersetzen. Dabei werden auch Indikatorvariablen erstellt, mit deren Hilfe Beobachtungen erkannt werden können, die ersetzte Variablenwerte enthalten.
- Variablenauswahl: Bei erweiterten Modellen werden Chi-Quadrat- und R-Quadrat-Tests verwendet, um Variablen zu entfernen, die keine Beziehung zur Zielvariable aufweisen. Während der R-Quadrat-Analyse werden AOV16-Variablen erstellt.
- Kombination von Variablenselektionstechniken: Bei erweiterten Modellen werden die durch die Chi-Quadrat- und R-Quadrat-Tests ausgewählten Variablen zusammengeführt.
- Modellerstellung: Bei erweiterten Modellen werden die Trainingsdaten an vier konkurrierende Modellierungsalgorithmen weitergeleitet. Dabei handelt es sich um einen Entscheidungsbaum, ein neuronales Modell, eine Rückwärtsregression sowie ein Ensemble-Modell. Im neuronalen Netz-Modell werden begrenzte Suchen ausgeführt, um ein optimales Feed-Forward-Netz zu finden. Die Rückwärtsregression ist ein lineares Regressionsmodell, bei dem Variablen eliminiert werden, indem eine Variable nach der anderen entfernt wird, bis die R-Quadrat-Werte signifikant niedriger sind. Beim Ensemble-Modell werden neue Modelle erstellt, indem A-Posteriori-Wahrscheinlichkeiten (für Klassifizierungsziele) oder prognostizierte Werte (für Intervallziele) aus unterschiedlichen Eingangsmodellen kombiniert werden. Mit dem neuen Ensemble-Modell wird anschließend das Scoring von neuen Daten durchgeführt. Das bei der erweiterten Methodik verwendete Ensemble-Modell wird aus dem Ausgabemodell eines einfachen Modells, dem Champion-Modell eines mittleren Modells und dem Champion-Modell eines erweiterten Modells erstellt.

- Auswahl des Champion-Modells: Bei erweiterten Modellen wird eine analytische Bewertung der Prognose- oder Klassifizierungs-Performance des Entscheidungsbaums sowie des "neuronalen" und des Regressionsmodells durchgeführt. Das Modell mit der besten Prognose- oder Klassifizierungs-Performance wird als Eingabe zur Erstellung eines Ensemble-Modells verwendet. Ebenso werden dafür die Champion-Modelle des einfachen und des mittleren Modells herangezogen. Anschließend wird das neu erstellte erweiterte Ensemble-Modell mit dem Entscheidungsbaummodell, dem "neuronalen" Modell und der Rückwärtsregression verglichen, um das Champion-Modell mit der besten Performance herauszufiltern.

Nachdem das Champion-Modell ausgewählt wurde, wird es ausgeführt und die Prognose-Performance des erweiterten Modells wird mit den Champion-Modells der mittleren und einfachen Modelle verglichen, bis am Ende das Champion-Modell mit der besten Performance feststeht.

SAS Rapid Predictive Modeler: Festlegen von Berichtoptionen

Berichtoptionen

Klicken Sie im Auswahlbereich auf **Bericht**, um auf diese Optionen zuzugreifen.

Aus den Berichten lassen sich signifikante Terme im Modell erkennen. Darüber hinaus werden übliche Diagramme, wie z.B. Lift-Diagramme erzeugt. In den Ergebnissen sind Statistiken für Trainingsdaten und Validierungsdaten enthalten. Bei der Verarbeitung der Daten im SAS Rapid Predictive Modeler werden die Eingangsdaten nach Trainingsdaten und Validierungsdaten unterschieden. Die Trainingsdaten werden zur Berechnung der Parameter für die einzelnen Modelle verwendet. Mit ihnen werden letztlich die Gütestatistiken auf den Trainingsdaten berechnet. Anschließend werden die Validierungsdaten mittels der einzelnen Modelle gescort, was die zugehörigen Validierungs-Gütestatistiken liefert. Die Validierungs-Gütestatistiken werden verwendet, um Modelle miteinander zu vergleichen und Überanpassungen zu erkennen. Wenn die Trainingstatistiken signifikant besser sind als die Validierungsstatistiken, kann von einer Überanpassung ausgegangen werden, was bedeutet, dass im Modell zufällige Signale in den Daten erlernt wurden. Üblicherweise werden Modelle mit den besseren Validierungstatistiken präferiert.

Der Bericht enthält auf kurze und dennoch übersichtliche Art Informationen über die Datenquelle sowie Variablen, die bei der Modellierung verwendet wurden, eine Rangordnung der wichtigen Prädiktorvariablen, mehrere Gütestatistiken, mit denen die Genauigkeit des Modells eingeschätzt werden kann, sowie eine Modell-Scorecard. Weitere Informationen finden Sie unter Allgemeines über die Standardberichte für den SAS Rapid Predictive Modeler.

Darüber hinaus können Sie diese Informationen zu Ihren Berichten hinzufügen:

- **Modellübersicht**
- **Variablenrangfolge**
- **Kreuztabellen**
- **Klassifizierungsmatrix**
- **Anpassungsstatistiken**
- **Lift-Plot**
- **Modellvergleich**

Hinweis: Diese Option ist für das einfache Modell nicht verfügbar, da dieses nur einen Algorithmus zur Modellerstellung verwendet.

Allgemeines über Standardberichte für den SAS Rapid Predictive Modeler

Die folgenden Berichte werden standardmäßig im SAS Rapid Predictive Modeler generiert:

Gewinn-Diagramm
: Gewinn-Diagramme sind nur für Modelle verfügbar, die Klassifizierungs-Zielvariablen enthalten. In diesem Diagramm werden die Perzentile der Daten, angeordnet nach prognostiziertem Wert dargestellt. "Lift" ist ein Maß für den Quotienten aus der Anzahl der vom Modell identifizierten Zielereignisse und der Anzahl der Zielereignisse, die durch eine Zufallsauswahl erkannt wurden.

ROC-Plot
: In diesem Plot wird die maximale prädiktive Leistungsfähigkeit eines Modells für die gesamte Stichprobe dargestellt (nicht für nur ein einzelnes Dezil). Die Daten werden dargestellt als Sensitivität gegen (1 – Spezifizität). Die Trennung zwischen der Modellkurve und der diagonalen Linie (die ein Zufallsselektionsmodell darstellt) wird als Kolmogorow-Smirnow-(KS)-Wert bezeichnet. Je höher der KS-Wert, desto leistungsfähiger das Modell.

Scorecard
: Aus der Scorecard können die Eigenschaften des Modells abgelesen werden können. Bei der Erstellung der Scorecard wird jede Intervallvariable in Wertebereiche mit Einzelwerten unterteilt. Anschließend wird jede Variable nach ihrer Bedeutung für das Modell angeordnet und auf ein Maximum von 1.000 Punkten skaliert. Dem Einzelwert jeder Variable wird daraufhin sein Anteil an dem skalierten Gesamtwert zugeschrieben.

Projektinformationen
: Aus den Projektinformationen ist ersichtlich, welcher Benutzer das Modell erstellt hat, wann das Modell erstellt wurde und wo die Dateien für die Modellkomponenten abgelegt sind.

SAS Rapid Predictive Modeler: Festlegen von Ausgabeoptionen

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Enterprise Miner-Projektdaten speichern	legt fest, ob die SAS Enterprise Miner-Daten aus dieser Anwendungsroutine gespeichert werden sollen. Ein SAS Rapid Predictive Modeler-Modell ist ein Beispiel für ein SAS Enterprise Miner-Projekts. Wenn Sie SAS Enterprise Miner-Daten speichern, können Sie die Benutzeroberfläche des SAS Enterprise Miner verwenden, um das erstellte SAS Rapid Predictive Modeler-Modell zu öffnen und bearbeiten. Im SAS Enterprise Miner können Sie die Analyse speichern und exportieren, um sie so außerhalb des SAS Enterprise Miner wiederzuverwenden. Darüber hinaus ist es möglich, das Modell in einem SAS Metadata Repository zu registrieren. Das Kontrollkästchen **Enterprise Miner-Projektdaten speichern** ist standardmäßig aktiviert. ■ Bei einer Installation auf einer Workstation, wird das Projekt lokal in dem Workstation-Verzeichnis gespeichert, das im Feld **SAS App Server-Ordner** angegeben ist. Klicken Sie auf **Durchsuchen**, um diesen Speicherort zu ändern. ■ Bei einer Client-/Server-Installation des SAS Enterprise Miner wird das Projekt im Stammverzeichnis des Benutzers auf dem Server gespeichert (z.B. `C:\users\username\Documents`). Wenn Sie mithilfe der SAS Management Console ein Systemstammverzeichnis auf dem Server erstellt haben (Beispiel: `C:\projects`), ist das Feld **SAS App Server-Ordner** leer. Klicken Sie in diesem Fall auf **Durchsuchen**, um den Speicherort für Ihr Projekt festzulegen. Wenn das Feld **SAS App Server-Ordner** nicht ausgefüllt und kein Speicherort angegeben wird, bevor Sie ein Projekt speichern, wird eine Fehlermeldung angezeigt, die Sie zur Angabe eines gültigen Pfads auffordert.
Score für Datei	legt den Namen der Datei fest, die durch das vom SAS Rapid Predictive Modeler erstellten Modell gescort wird.

SAS Rapid Predictive Modeler: Registrieren Ihres Modells

Klicken Sie im Auswahlbereich auf **Registrierung**, um auf diese Optionen zuzugreifen.

Informationen zur Registrierung eines Modells

Modelle, die mit dem SAS Rapid Predictive Modeler erstellt wurden, können im Handumdrehen bereitgestellt werden. Sie werden als SAS-Code gespeichert, so dass Sie ein Modell auf jeder beliebigen Installation von SAS 9.2 (oder später) ausführen können.

Hinweis: Sie müssen den SAS Rapid Predictive Modeler mindestens einmal ausführen oder speichern, bevor Sie Ihr Modell im SAS Metadata Repository registrieren können. Wenn Sie das Modell erstellt haben, muss die Option **Enterprise Miner-Projektdaten speichern** aktiviert sein, damit die Registrierungsoptionen zur Verfügung stehen.

Ein einmal registriertes Modell kann auch von anderen SAS-Benutzern verwendet werden.

- Somit können andere Benutzer das registrierte Modell verwenden, um ihre eigenen Scoring-Prozesse in **SAS Enterprise Guide** oder **SAS Add-In for Microsoft Office** durchzuführen. Beispielsweise kann ein registriertes Modell in der Anwendungsroutine **Model Scoring** verwendet werden, um damit neue Daten zu scoren.
- Die Modelle können auch als Projekte und Diagramme im SAS Enterprise Miner geöffnet und zudem bearbeitet werden, um etwa weitere Analysen durchzuführen oder das Verhalten unter anderen Parametern alternativ zu testen.
- Darüber hinaus kann das Modell in den **SAS Model Manager** importiert werden, wo weitere Modellierungsfunktionen zur Verfügung stehen. Ebenso ist es dort möglich, die Modell-Performance hinsichtlich der Verschlechterung im zeitlichen Verlauf zu überwachen.
- Benutzer von SAS Data Integration Studio können Modelle importieren, um verwaltete und zeitlich geplante Scoring-Prozesse zu erstellen.

Es ist auch möglich, SAS Rapid Predictive Modeler-Modelle mithilfe des SAS Scoring Accelerator for Teradata, DB2 oder Netezza zu veröffentlichen.

Registrieren eines Modells im SAS Metadata Repository

1. Geben Sie den Namen, den Pfad und die Beschreibung für das Modell an.
2. Klicken Sie auf **Registrieren**.

Hinweis: Die Schaltfläche **Registrieren** ist nicht verfügbar, wenn keine Verbindung zu einem SAS Workspace Server hergestellt wurde.

77

Streuungsdiagramm

Allgemeines über die Anwendungsroutine 'Streuungsdiagramm' ... **619**

Streuungsdiagramm: Auswählen eines Diagrammtyps ... **620**

Streuungsdiagramm: Zuweisen von Variablen zu Analyserollen ... **620**

Streuungsdiagramm: Festlegen von Darstellungsoptionen ... **622**
- Optionen für 3D Streuungsdiagramme festlegen ... 622
- Weitere Plotoptionen festlegen ... 623
- Weitere Darstellungsoptionen ... 624

Festlegen von Titeln und Fußnoten ... **624**

Anzeigen von Eigenschaften ... **624**

Allgemeines über die Anwendungsroutine 'Streuungsdiagramm'

Die Anwendungsroutine **Streuungsdiagramm** erstellt zweidimensionale Streuungsdiagramme, dreidimensionale Streuungsdiagramme oder dreidimensionale Nadeldiagramme, in denen die Beziehungen zwischen zwei oder drei Variablen anhand von Mustern oder Datenpunktkonzentrationen aufgezeigt werden.

Beispielsweise kann ein zweidimensionales Streuungsdiagramm Gewicht und Alter aller Patienten anzeigen, die an einer klinischen Studie teilnehmen.

Weitere Informationen, siehe “Auswählen des Grafikausgabeformats” auf Seite 6.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GPLOT, G3D
Erforderliche SAS-Produkte	Base SAS, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	Ohne

Streuungsdiagramm: Auswählen eines Diagrammtyps

Klicken Sie im Auswahlfenster auf **Streuungsdiagramm**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
2D Streuungsdiagramm	erstellt ein zweidimensionales Diagramm, in dem die Werte der beiden Variablen als Datenpunkte auf einer horizontalen Achse (x) und einer vertikalen Achse (y) angezeigt werden. Jedes Paar mit x- und y-Werten bildet einen Datenpunkt.
3D Streuungsdiagramm	erstellt ein dreidimensionales Diagramm, in dem die Werte von drei numerischen Variablen als Datenpunkte angezeigt werden. Jede Gruppe mit x-, y- und z-Werten bildet einen Datenpunkt.
3D Streuungsdiagramm mit Nadeln	erstellt ein dreidimensionales Diagramm, in dem die Werte von drei numerischen Variablen als Datenpunkte angezeigt werden. Jede Gruppe mit x-, y- und z-Werten bildet einen Datenpunkt. Jeder Datenpunkt ist mit sogenannten Nadeln mit der horizontalen Ebene verbunden.

Streuungsdiagramm: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Horizontal	Die Spalte, die Sie dieser Rolle zuweisen, ist die horizontale oder X-Achsenvariable für das Diagramm.
Vertikal	Die Spalte, die Sie dieser Rolle zuweisen, ist die vertikale oder Y-Achsenvariable für das Diagramm. So erstellen Sie eine statistische Zusammenfassung der eindeutigen Werte der horizontalen Variablen. Wählen Sie die Variable unter der Rolle **Vertikal** aus. Wählen Sie **Zusammenfassung der einzelnen horizontalen Werte** und anschließend die Zusammenfassungsfunktion. Wenn z. B. Ihre Eingabedaten die Umsätze und Ausgaben jedes Landes nach Regionen aufführen. Weisen Sie der Variablen 'Country' die Rolle **Horizontal** und der Variablen 'Sales' die Rolle **Vertikal** zu. Wenn Sie die Summe der Umsätze für jedes Land darstellen möchten, aktivieren Sie das Kontrollkästchen und wählen aus der Dropdown-Liste **Summe** aus. Die Anwendungsroutine fügt die regionalen Umsätze für ein Land hinzu und stellt in der Endausgabe einen einzelnen Umsatzwert für jedes Land dar.
Vertikal (Rechts)	Die Spalte, die Sie dieser Rolle zuweisen, ist die vertikale bzw. Y-Achsenvariable für die rechte Seite der Grafik. Diese Variable wird gegen die X-Achsenvariable dargestellt, sodass ein anderer Plot auf der Grafik erstellt wird. So erstellen Sie eine statistische Zusammenfassung der eindeutigen Werte der horizontalen Variablen. Wählen Sie die Variable unter der Rolle **Vertikal** aus. Wählen Sie **Zusammenfassung der einzelnen horizontalen Werte** und anschließend die Zusammenfassungsfunktion. Wenn z. B. Ihre Eingabedaten die Umsätze und Ausgaben jedes Landes nach Regionen aufführen. Weisen Sie der Variablen 'Country' die Rolle **Horizontal** und der Variablen 'Expenses' die Rolle **Vertikal (Rechts)** zu. Wenn Sie die Summe der Ausgaben für jedes Land darstellen möchten, aktivieren Sie das Kontrollkästchen und wählen aus der Dropdown-Liste **Summe** aus. Die Anwendungsroutine fügt die regionalen Ausgaben für ein Land hinzu und stellt in der Endausgabe einen einzelnen Ausgabenwert für jedes Land dar.

Optionsname	Beschreibung
	Hinweis: Diese Rolle ist nur für ein zweidimensionales Streuungsdiagramm verfügbar.
Tiefe	Die Spalte, die Sie dieser Rolle zuweisen, ist die Variable der dritten Dimension bzw. Z-Achse für das Diagramm. **Hinweis:** Diese Option ist nur für dreidimensionale Plots verfügbar.
Diagramme gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Diagramme, die erstellt werden. Für jeden eindeutigen Wert wird eine separate Grafik erzeugt. Aktivieren Sie das Kontrollkästchen **Gleiche Achsenskalierung für alle Diagramme**, wenn Sie möchten, dass für alle erzeugten Diagramme die gleiche Achsenskalierung verwendet wird. Standardmäßig basiert der Wertebereich für jede Achse auf den Minimal- und Maximalwerten der dargestellten Daten und könnte daher in den Diagrammen variieren. Wenn Sie diese Option auswählen haben alle Diagramme den gleichen Bereich für jede Achse.

Streuungsdiagramm: Festlegen von Darstellungsoptionen

Optionen für 3D Streuungsdiagramme festlegen

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **3D Streuung**, um auf diese Optionen zuzugreifen.

Wählen Sie einen Symboltyp, Farbe und Größe aus und bestimmen Sie, ob für jeden Datenpunkt Nadeln angezeigt werden sollen.

Stellen Sie über die Schieberegler oder die Steuerelemente **Drehen** und **Neigen** den Ansichtswinkel ein. Standardmäßig betragen Dreh- und Neigungswinkel 70 Grad. Klicken Sie auf **Zurücksetzen**, um den Betrachtungswinkel auf die Standardeinstellung zurückzusetzen.

Hinweis: Wenn Sie das Kontrollkästchen **Werte für Drehen und Neigen angeben** deaktivieren, wird das Drehen und Neigen der Grafik durch das

Grafikausgabeformat bestimmt. Standardmäßig ist diese Option aktiviert und Sie können den Sichtwinkel festlegen.

Weitere Plotoptionen festlegen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Plots**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Optionen sind nur für zweidimensionale Streuungsdiagramme verfügbar.

So steuern Sie die Darstellung einer bestimmten Variablen im Plot:

1. Wählen Sie eine Variable aus der Liste aus. In der Liste werden die Variablen angezeigt, die Sie den Rollen **Vertikal** und **Vertikal (Rechts)** zugewiesen haben. K
2. Geben Sie im Bereich **Linie** den Linienstil, die Linienbreite- und -farbe an.
3. Wählen Sie eine Umrissfarbe aus. Diese Option legt die Farben für folgende Elemente an:
   - den Umriss sämtlicher eingeschlossener Flächen
   - Konfidenzgrenzen in einer Regressionsanalyse
   - die Linien und Umrisse für Standardabweichungsplots
4. Geben Sie im Bereich **Datenpunktmarker** den Typ, das zu verwendende Symbol, die Höhe für jeden Marker und die Farbe der Datenpunktmarker an.

   Hinweis: Manche Symbole werden nicht von allen Grafikausgabeformaten unterstützt. Wenn Sie ein nicht durch das Ausgabeformat unterstütztes Symbol auswählen, wird stattdessen ein anderes Symbol (ein unterstütztes Symbol) angezeigt.
5. Legen Sie fest, ob die Plotlinien bei fehlenden Werten unterbrochen werden soll.

   Hinweis: Das Kontrollkästchen **Zweite Plotlinie bei fehlenden Werten unterbrechen** ist nur verfügbar, wenn Sie der Rolle **Vertikal (Rechts)** eine Variable zuweisen.

Weitere Darstellungsoptionen

Einige Darstellungsoptionen gelten für alle Grafik-Anwendungsroutinen. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Auswählen der Interpolationsmethode” auf Seite 754
- “Festlegen von Achsenoptionen” auf Seite 745
- “Hinzufügen von Referenzlinien” auf Seite 748
- “Anpassen des Diagrammbereichs” auf Seite 750

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

78

Streuungsplot-Matrix

Allgemeines über die Anwendungsroutine 'Streuungsplot-Matrix' 627
Streuungsplot-Matrix: Erstellen einer Matrix mit Streuungsplots 628

Allgemeines über die Anwendungsroutine 'Streuungsplot-Matrix'

Mit der Anwendungsroutine **Streuungsplot-Matrix** wird ein Diagramm mit Streuungsdiagrammen für mehrere Kombinationen von Variablen erstellt. Mit dieser Option können Sie Anpassungsplots und Ellipsen in Ihren Streuungsdiagrammen überlagern.

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SGSCATTER
Erforderliche SAS-Produkte	Base SAS, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	Ohne

Streuungsplot-Matrix: Erstellen einer Matrix mit Streuungsplots

Optionsname	Beschreibung
Etikett	zur Identifizierung der Anwendungsroutine im Projektbaum und Prozessfluss. Geben Sie im Feld **Etikett** den Namen für die Anwendungsroutine ein.
Datenquelle	Ermöglicht die Bearbeitung der Eingabedatenquelle für die Anwendungsroutine. Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Anwendungsroutinenfilter** angezeigt.
Matrixvariablen	Legt die im Diagramm gegenübergestellten Variablen fest, die zur Erstellung des Streuungsdiagramms verwendet werden. Sie müssen dieser Rolle mindestens zwei Variablen zuweisen.
Gruppierungsvariable	Legt die Klassifizierungsvariable fest, nach der die Werte in Gruppen eingeteilt werden. Dieser Rolle muss keine Variable zugewiesen werden.
Diagrammtitel	Legt den Titel fest, der in den Ergebnissen angezeigt wird. Standardmäßig ist der Titel für die Ergebnisse auch der Titel für die Streuungsplot-Matrix. Sie können diesen Titel jedoch selbst anpassen. Klicken Sie auf den Abwärtspfeil, um den Titeltext zu bearbeiten. Sie können auch Makrovariablen im Titel verwenden. Die Werte für die Makrovariablen werden bei der Ausführung der Anwendungsroutine ermittelt.
Diagonal	Fügt Grafiken zu den Diagonalzellen der Plotmatrix hinzu. Sie können Histogramme in die Diagonalzellen platzieren und diese Histogramme mit Normaldichtekurven oder Kerndichteschätzungen überlagern. Darüber hinaus können Sie festlegen, ob die Diagonale links oben oder links unten beginnen soll. Wenn Sie keine Grafik für die Diagonale auswählen, enthalten die Diagonalzellen die Variablennamen.

Optionsname	Beschreibung
Ellipse	Fügt eine Konfidenz- oder Prognoseellipse zu jeder Zelle hinzu, in der ein Streuungsdiagramm enthalten ist. Bei beiden Ellipsenarten wird eine bivariate Normalverteilung vorausgesetzt. Sie können zwischen folgenden Arten von Ellipsen wählen: ■ **Ohne** - Es wird keine Ellipse hinzugefügt. ■ **Mittelwert** - Fügt eine Konfidenzellipse zur mittleren Grundgesamtheit hinzu. ■ **Vorhergesagt** - Fügt eine Prognoseellipse für eine neue Beobachtung hinzu. Sie können auch das Konfidenzniveau für die Ellipse angeben.
Legende	Legt fest, ob eine Legende für die Streuungsplot-Matrix angezeigt werden soll. Legenden sind nur verfügbar, wenn eine Gruppierungsvariable zugewiesen wurde. In dem Fall wird standardmäßig eine Legende angezeigt. Um die Legende zu entfernen, aktivieren Sie das Kontrollkästchen **Legende anzeigen**. Geben Sie einen Titel für die Legende an und wählen Sie anschließend eine Position für die Legende aus der Dropdown-Liste **Position** aus. Sie haben auch die Möglichkeit, einen Rahmen um die Legende anzuzeigen.
Diagrammbereich	Legt die Höhe und die Breite eines Diagramms fest. Sie können auch angeben, ob beim Bewegen des Mauszeigers über einen Datenpunkt im Streuungsplot Datentipps angezeigt werden. Der Standarddiagrammbereich ist 800 x 800 Pixel groß und es werden keine Datentipps angezeigt.
Fußnote für Diagramm	Legt die Fußnote fest, die in den Ergebnissen angezeigt wird. Standardmäßig werden in der Fußnote des Diagramms das Datum und die Uhrzeit der Ausführung der Anwendungsroutine angezeigt. (Im Textfeld **Fußnote für Diagramm** sind SAS-Makrovariablen für das aktuelle Datum und die Uhrzeit enthalten.) Sie können die Fußnoten für Ihr Diagramm jedoch individuell anpassen. Klicken Sie auf den Abwärtspfeil, um den Text in der Fußnote zu bearbeiten. Sie können auch Makrovariablen in der Fußnote verwenden. Um eine Fußnote zu löschen, löschen Sie den gesamten Text im Textfeld und klicken Sie auf **OK**. In der Ergebnissen wird daraufhin keine Fußnote angezeigt.

79

Daten sortieren

Allgemeines über die Anwendungsroutine 'Daten sortieren' 631
Daten sortieren: Zuweisen von Variablen zu Analyserollen 632
Daten sortieren: Optionen für "Daten sortieren" festlegen 632
Daten sortieren: Festlegen von Ergebnisoptionen 635
Anzeigen von Eigenschaften 635

Allgemeines über die Anwendungsroutine 'Daten sortieren'

Mit der Anwendungsroutine **Daten sortieren** können Sie eine Datei in Ihrem Projekt nach einer beliebigen Spalte sortieren. Sie können die Daten nach einer oder mehreren Spalten sortieren, und Sie können für jede Spalte eine aufsteigende oder absteigende Reihenfolge auswählen. Die Anwendungsroutine **Daten sortieren** erstellt die sortierte Datei als neues Datenelement in Ihrem Projekt.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SORT
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	Ohne

Daten sortieren: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Sortieren nach	Wenn Sie dieser Rolle mindestens eine Variable zuweisen, wird die Tabelle nach der bzw. den ausgewählten Variablen sortiert. Die Reihenfolge, in der die Variablen in dieser Rolle angezeigt werden, bestimmt die Variable, die als primärer Sortierschlüssel, sekundärer Sortierschlüssel usw. verwendet wird. Der primäre Sortierschlüssel ist stets die erste Variable, die innerhalb der Rolle **Sortieren nach** aufgelistet ist.
Zu entfernende Spalten	Wenn Sie dieser Rolle mindestens eine Variable zuweisen, enthält die generierte Ausgabe nicht die angegebenen Variablen. Sie können dieser Rolle maximal (*n* - 1) Variablen zuweisen, wobei *n* die Gesamtzahl der zu sortierenden Variablen in der Datei ist.

Daten sortieren: Optionen für "Daten sortieren" festlegen

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Ausgabereihenfolge	Sie können die Reihenfolge der Daten bestimmen, indem Sie folgende Optionen modifizieren: ■ **Sortierreihenfolge** gibt an, welche Reihenfolge beim Sortieren der alphanumerischen Variablen verwendet werden soll. Sie können die auf dem Server definerte Sortierreihenfolge verwenden. Sie können aber auch ASCII oder EBCDIC verwenden, oder auch die nationalen Standards Dänisch, Finnisch, Norwegisch oder Schwedisch. Sie können alphanumerische Variablen auch mit einer benutzerdefinierten Sortierreihenfolge sortieren, die durch Ihre Installations-Site definiert ist. **Hinweis:** Wenn Sie **Benutzerdefiniert** wählen, muss Ihr Site eine benutzerdefinierte nationale Sortierfolge aufweisen, die definiert wurde. Ihr Administrator kann festlegen, ob eine benutzerdefinierte nationale Sortierreihenfolge verfügbar ist. Wenn keine Sortierfolge verfügbar ist, wird die für den Server definierte Standardreihenfolge verwendet. ■ **Originaldatenreihenfolge innerhalb "Sortieren nach"-Gruppierungen beibehalten** gruppiert die Daten in der Reihenfolge, die Sie für die Analyserollen festlegen. Wenn diese Option nicht ausgewählt ist, wird die Ausgabedatei in unbestimmter Reihenfolge innerhalb der sortierten Schlüsselgruppen gruppiert.
Doppelte Datensätze	Sie können den Umgang mit doppelten Datensätzen in der Ausgabe mit den folgenden Optionen festlegen: ■ **Alle Datensätze beibehalten** behält alle Datensätze in der Ausgabedatei bei, einschließlich aller Datensatzduplikate. ■ **Nur den ersten Datensatz für jede "Sortieren nach"-Gruppe beibehalten** entfernt alle doppelten Beobachtungen, die dieselben Werte für die Gruppe **Sortieren nach** aufweisen. Wenn die Option **Reihenfolge der Originaldaten innerhalb der "Sortieren nach"-Gruppierungen beibehalten** ausgewählt ist, wird die Beobachtung, die für jede **Sortieren nach**-Gruppe beibehalten wird, die erste Beobachtung sein, die aus der Originaldatei gelesen wird. Ist jedoch die Option **Reihenfolge der Originaldaten innerhalb der "Sortieren nach"-Gruppierungen beibehalten** nicht ausgewählt, kann die Beobachtung, die für jede **Sortieren nach**-Gruppe beibehalten wird, nicht vorab festgelegt werden. ■ **Benachbarte doppelte Datensätze nicht beibehalten** vergleicht jeden Datensatz mit dem vorherigen Datensatz in der Ausgabedatendatei. Wenn eine exakte Übereinstimmung gefunden wird, wird der doppelt

Optionsname	Beschreibung
	vorhandene Datensatz nicht in die Ausgabedatei geschrieben. **Hinweis:** Wenn nicht alle Variablen zur Rolle **Sortieren nach** zugewiesen werden, können einige doppelte Datensätze ggf. nicht entfernt werden.
Erweiterte Sortieroptionen	Sie können für die Art und Weise, nach der die Daten in der Ausgabe sortiert werden sollen, erweiterte Optionen festlegen. Sie können folgende Optionen ändern: ■ **Speicher für Sortierung** gibt die maximale Speichermenge an, die für die Anwendungsroutine Daten sortieren verwendet werden kann. Sie können die Speichermenge in Byte (B), Kilobyte (KB), Megabyte (MB) oder Gigabyte (GB) angeben. Sie können auch angeben, dass der gesamte verfügbare Speicher oder die auf dem Server zugeordnete Standardspeichermenge verwendet werden kann. ■ **Temporäre Anforderungen an den Plattenplatz reduzieren** weist darauf hin, dass während der Sortierung der Daten nur die **Sortieren nach**-Variablen und die Beobachtungsnummern in temporären Dateien gespeichert werden, um die erforderliche Speichermenge zum Durchführen der Sortierung zu reduzieren. In der Endphase der Sortierung wird die temporäre Datei als Index zum Zugreifen auf die Originaldatei und zum Senden der Daten an die Ergebnisdatei mit der richtigen Sortierfolge gesendet. ■ **Sortierung indizierter Daten erzwingen** gibt an, dass Sie alle Dateien sortieren möchten, einschließlich der Dateien, die eventuell bereits in der gewünschten Reihenfolge sortiert sind oder die einen vom Benutzer erstellten Index mit Schlüsselbegriffen enthalten, die denen über die Rolle **Sortieren nach** angegebenen Schlüsselbegriffen entsprechen. Wenn Sie diese Option angeben, wird die Datei unabhängig von der aktuellen Reihenfolge der Datei und unabhängig davon sortiert, ob sie einen Index enthält oder nicht.

Daten sortieren: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Sie können der Ausgabedatei einen anderen Namen oder Speicherort zuweisen. Außerdem können Sie doppelte Datensätze in einer Datei speichern, wenn Sie doppelte Beobachtungen entfernen möchten, die die gleichen Werte für die Gruppe **Sortieren nach** haben, oder wenn Sie nur einen Datensatz derjenigen Datensätze behalten möchten, deren Werte genau übereinstimmen. Weitere Informationen finden Sie unter den Optionen **Doppelte Datensätze**.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten

- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

80

Teilen von Spalten

Allgemeines über die Anwendungsroutine 'Spalten teilen' 637
Spalten teilen: Zuweisen von Variablen zu Analyserollen 639
Spalten teilen: Festlegen von Ergebnisoptionen 640
Anzeigen von Eigenschaften 640

Allgemeines über die Anwendungsroutine 'Spalten teilen'

Die Anwendungsroutine **Spalten teilen** erstellt eine Ausgabedatei durch Teilen einer eindeutigen Wertekombination aus den gewählten Spalten der Eingabedatei in mehrere Spalten. Sie können die Ausgabedatei verwenden, um individuelle Analysen der Spalten vorzunehmen, in denen mehrere Spalten der Eingabedatei enthalten sind.

Diese Anwendungsroutine ist nützlich, wenn Sie über eine Datei verfügen, in der eine Spalte mehrere Beobachtungen für verschiedene Untergruppen enthält und Sie die Untergruppenmessungen in separate Spalten aufteilen möchten.

Beispiel: In der folgenden Datei ist eine einzelne Spalte mit den monatlichen Temperaturmessungen für verschiedene Messstellen im US-Bundesstaat North Carolina enthalten:

	Location	Month	Temperature
1	Asheville	January	46.9
2	Asheville	February	50.6
3	Asheville	March	59
4	Asheville	April	68
5	Asheville	May	74.9
6	Asheville	June	81
7	Charlotte	January	51
8	Charlotte	February	52.8
9	Charlotte	March	61.7
10	Charlotte	April	70.5
11	Charlotte	May	79.6
12	Charlotte	June	86.5
13	Raleigh	January	51
14	Raleigh	February	54.1
15	Raleigh	March	62.1
16	Raleigh	April	71.6
17	Raleigh	May	79
18	Raleigh	June	85.8
19	Wilmington	January	56
20	Wilmington	February	59

Die Anwendungsroutine **Spalten teilen** ermöglicht es Ihnen, die folgende Datei zu erstellen. Dabei werden die monatlichen Messwerte für jede Messstelle in eine Spalte für jeden Monat transponiert:

	City	January	February	March	April	May
1	Raleigh	51	54.1	62.1	71.6	79
2	Wilmington	56	59	65.8	74.4	81.1
3	Asheville	46.9	50.6	59	68	74.9
4	Charlotte	51	52.8	61.7	70.5	79.6

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	TRANSPOSE
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

Spalten teilen: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Zu teilende Spalte	Die Spalte, die Sie dieser Rolle zuweisen, enthält die Werte, die in die neuen Spalten transponiert werden. Sie können dieser Rolle nur eine Variable zuweisen.
Werte-Identifier-Spalte	Die Spalte, die Sie dieser Rolle zuweisen, liefert die Namen für die neuen Spalten, in denen die transponierten Werte aus den **zu teilenden Spalten** enthalten sind. Die Folge der eindeutigen Werte in dieser Spalte gibt an, welche Zeilen in die neuen Spalten transponiert werden. Wenn Sie keine Gruppierungsvariablen zuweisen, muss diese Spalte nur eine Menge eindeutiger Werte enthalten. Die Werte in **Zu teilende Spalten** werden in eine Zeile der Ausgabedatei transponiert. Sie müssen der Rolle **Werte-Identifier-Spalte** eine Spalte zuweisen. Um eine Gruppierungsvariable auszuwählen, weisen Sie der Rolle **Analyse gruppieren nach** eine Spalte zu.
Werteetiketten-Spalte	Die Spalte, die Sie dieser Rolle zuweisen, stellt die Etiketten für die neuen Spalten zur Verfügung, die die transponierten Werte der **zu teilenden Spalten** enthalten. **Hinweis:** Welche Wirkung diese Spalte hat, können Sie überprüfen, indem Sie für die neuen Spalten in der Ausgabedatei die Eigenschaften anzeigen.
Analyse gruppieren nach	Jede Variable, die Sie dieser Rolle zuweisen, wird verwendet, um die Zeilen für **Zu teilende Spalte** in Untergruppen zu unterteilen, die in neue Spalten transponiert werden. Jede Untergruppe, definiert durch eine eindeutige Wertekombination der

Optionsname	Beschreibung
	Gruppierungsvariablen, ergibt eine Zeile in der Ausgabedatei. Die Anzahl der Zeilen, die eine Untergruppe bilden, wird durch die Anzahl eindeutiger Werte in der **Werte-Identifier-Spalte** bestimmt.

Spalten teilen: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Sie können der Ausgabedatei einen anderen Namen zuweisen.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Sie können den Spaltennamen in der Ausgabedatei ein Präfix hinzufügen. Standardmäßig werden die Namen der Spalten aus den eindeutigen Werten in der Spalte **Werte-Identifer-Spalte** erstellt. Um ein Präfix anzugeben, aktivieren Sie das Kontrollkästchen **Präfix verwenden** und geben das Präfix in das Textfeld ein. Das Standardpräfix ist `Spalte`.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

81

Spalten stapeln

Allgemeines über die Anwendungsroutine 'Spalten stapeln' **643**
Spalten stapeln: Zuweisen von Variablen zu Analyserollen **645**
Spalten stapeln: Festlegen von Optionen für Spaltenbenennungen **646**
Spalten stapeln: Festlegen von Ergebnisoptionen **647**
Anzeigen von Eigenschaften **647**

Allgemeines über die Anwendungsroutine 'Spalten stapeln'

Die Anwendungsroutine **Spalten stapeln** erstellt eine Ausgabedatei durch Neustrukturieren ausgewählter Spalten in der Einlesedatei, so dass diese Spalten in Beobachtungen transponiert werden. Sie können die Ausgabedatei verwenden, um Werteanalysen über mehrere Spalten der Einlesedatei vorzunehmen. Wenn Sie die Beobachtungen gruppieren, werden die ausgewählten Spalten in Untergruppen unterteilt, die auf den eindeutigen Kombinationen der Gruppierungswerte basieren. Jede Untergruppe ergibt eine Zeile in der Ausgabedatei.

Diese Anwendungsroutine ist hilfreich, wenn eine Datei vorliegt, in der jede Beobachtung denselben Datentyp in mehreren Spalten enthält und Sie die Daten in mehreren Spalten analysieren möchten. Beispiel: In der folgenden Datei sind monatliche Temperaturmessungen für verschiedene Messstellen im US-Bundesstaat North Carolina enthalten:

	City	January	February	March	April	May
1	Raleigh	51	54.1	62.1	71.6	79
2	Wilmington	56	59	65.8	74.4	81.1
3	Asheville	46.9	50.6	59	68	74.9
4	Charlotte	51	52.8	61.7	70.5	79.6

Die Struktur der Daten ermöglicht es Ihnen, beschreibende Statistiken für jeden Monat zu berechnen, basierend auf allen Messstellen in dem Bundesstaat. Es ist jedoch nicht möglich, beschreibende Statistiken für eine bestimmte Messstelle zu berechnen, da die Statistiken jeweils über eine Beobachtung berechnet werden müssen.

Um Statistiken für alle Messwerte jeder Messstelle zu berechnen, können Sie die Spalten mit den monatlichen Temperaturwerten nach Messstelle stapeln und so die folgende Ausgabedatei erstellen:

	Location	Month	Temperature
1	Asheville	January	46.9
2	Asheville	February	50.6
3	Asheville	March	59
4	Asheville	April	68
5	Asheville	May	74.9
6	Asheville	June	81
7	Charlotte	January	51
8	Charlotte	February	52.8
9	Charlotte	March	61.7
10	Charlotte	April	70.5
11	Charlotte	May	79.6
12	Charlotte	June	86.5
13	Raleigh	January	51
14	Raleigh	February	54.1
15	Raleigh	March	62.1
16	Raleigh	April	71.6
17	Raleigh	May	79
18	Raleigh	June	85.8
19	Wilmington	January	56
20	Wilmington	February	59

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	TRANSPOSE
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

Spalten stapeln: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Zu stapelnde Spalten	Jede Spalte, der Sie diese Rolle zuweisen, wird in der Ausgabedatei zu einer oder mehreren Zeilen. Wenn Sie eine oder mehrere Gruppierungsvariablen auswählen, werden mit den Gruppierungsvariablen alle Spalten in Unterspalten segmentiert, von denen jede in eine Zeile umgewandelt wird. In diesem Fall gehen aus einer Spalte soviele Zeilen hervor wie von den Gruppierungsvariablen Gruppen definiert werden. Sie müssen mindestens eine Spalte der Rolle **Zu stapelnde Spalten** zuweisen. Um eine Gruppierungsvariable auszuwählen, weisen Sie der Rolle **Analyse gruppieren nach** eine Spalte zu. **Hinweis:** Wenn Sie keine Gruppierungsvariablen zuweisen, darf die Einlesedatei nur eine Beobachtung enthalten. Jede Spalte ergibt eine Zeile in der Ausgabedatei.
Analyse gruppieren nach	Jede Variable, die Sie dieser Rolle zuweisen, wird verwendet, um die zu stapelnden Spalten in Untergruppen zu unterteilen, die separat transponiert werden. Jede Untergruppe, die durch eine eindeutige Wertekombination der Gruppierungsvariablen definiert wurde,

Optionsname	Beschreibung
	ergibt eine Zeile in der Ausgabedatei. Jede eindeutige Wertekombination der Gruppierungsvariable ist nur einmal vorhanden.

Spalten stapeln: Festlegen von Optionen für Spaltenbenennungen

Klicken Sie im Auswahlfenster auf **Spaltenbenennung**, um Namen und Etiketten für die neuen Spalten in der Ausgabedatei anzugeben.

Optionsname	Beschreibung
Name und Etikett für die Neue Wertespalte angeben	Durch das Transponieren der Werte einer Beobachtung in der Einlesedatei in Werte einer Variablen in der Ausgabedatei wird eine neue Spalte erstellt. Um eine Überschrift für die Ausgabespalte zu erstellen, welche die transponierten Werte enthält, geben Sie die Überschrift im Feld **Name** des Bereichs **Neue Wertespalte** ein. Zunächst enthält das Feld **Name** die Standardüberschrift `Gestapelte Werte`. Um eine Beschreibung für die Spalte anzugeben, die in der Ausgabedatei gespeichert wird, geben Sie im Feld **Etikett** im Bereich **Neue Wertespalte** einen Namen für das Etikett ein. Die alphanumerische Zeichenkette für die Beschreibung kann bis zu 256 Zeichen lang sein. Wenn Sie kein Etikett angeben möchten, lassen Sie dieses Feld leer.
Name und Etikett für die Abgeleitete Spalte angeben	Jede Zeile der Ausgabedatei enthält den Namen der Variablen in der Einlesedatei, die transponiert wird, um die aktuelle Beobachtung zu erstellen. Um eine Überschrift für die Ausgabespalte anzugeben, welche diese Variablennamen enthält, geben Sie die Überschrift in das Feld **Name** im Bereich **Abgeleitete Spalte** ein. Zunächst enthält das Feld **Name** die Standardüberschrift `Wertquelle`. Um eine Beschreibung für die Spalte anzugeben, die in der Ausgabedatei gespeichert wird, geben Sie im Feld **Etikett** im Bereich **Abgeleitete Spalte**einen Namen für das Etikett ein. Die alphanumerische Zeichenkette für die Beschreibung kann bis zu 256 Zeichen lang sein. Wenn Sie kein Etikett angeben möchten, lassen Sie dieses Feld leer.
Name und Etikett für die	Jede Zeile der Ausgabedatei enthält das Etikett der Variablen in der Eingabedatei, zu der die Werte in dieser Ausgabezeile gehören. Um

Optionsname	Beschreibung
Spalte für Beschreibung angeben	eine Überschrift für die Ausgabespalte anzugeben, welche diese Variablennamen enthält, geben Sie die Überschrift in das Feld **Name** im Bereich **Beschreibungsspalte** ein. Zunächst enthält das Feld **Name** die Standardüberschrift `Wertbeschreibung`. Um eine Beschreibung für die Spalte anzugeben, die in der Ausgabedatei gespeichert wird, geben Sie im Feld **Etikett** im Bereich **Spalte für Beschreibung** einen Namen für das Etikett ein. Die alphanumerische Zeichenkette für die Beschreibung kann bis zu 256 Zeichen lang sein. Wenn Sie kein Etikett angeben möchten, lassen Sie dieses Feld leer. **Hinweis:** Um die Auswirkungen zu überprüfen, die das Angeben eines Etiketts für eine Spalte hat, können Sie für die neue Spalte in der Ausgabedatei die Eigenschaften anzeigen.

Spalten stapeln: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Sie können der Ausgabedatei einen anderen Namen zuweisen.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

82

Standardisieren von Daten

Allgemeines über die Anwendungsroutine 'Daten standardisieren' 649
Daten standardisieren: Zuweisen von Variablen zu Analyserollen 650
Daten standardisieren: Standardisierungsoptionen festlegen 651
Daten standardisieren: Festlegen von Ergebnisoptionen 651
Anzeigen von Eigenschaften 652

Allgemeines über die Anwendungsroutine 'Daten standardisieren'

Die Anwendungsroutine **Daten standardisieren** sorgt für die Standardisierung von Variablen in einer SAS-Datei auf einen bestimmten Mittelwert und eine Standardabweichung. Außerdem erstellt sie eine neue SAS-Datei mit den standardisierten Werten.

Sie können diese Anwendungsroutine beispielsweise zum Standardisieren von Testscorewerten mithilfe eines angegebenen Mittelwerts und einer Standardabweichung verwenden.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	STANDARD
Erforderliche SAS-Produkte	Base SAS

Name der Anforderung	Prozedur- und Produktnamen
Empfohlene zusätzliche SAS-Produkte	Ohne

Daten standardisieren: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Analysevariablen	führt Standardisierungen für jede Variable durch, die Sie dieser Rolle zuweisen. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	legt die zur Berechnung eines gewichteten Mittelwerts und einer gewichteten Varianz für die Analysevariable verwendeten Werte fest. Sie können dieser Rolle maximal eine Variable zuweisen.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Sie können die Tabelle nicht nach einer Variablen gruppieren, die bereits einer anderen Rolle zugewiesen ist.

Daten standardisieren: Standardisierungsoptionen festlegen

Klicken Sie im Auswahlfenster auf **Standardisieren**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Standardisieren auf	Standardmäßig standardisiert diese Anwendungsroutine die Daten auf einen Mittelwert von 0 und eine Standardabweichung von 1. Sie können einen neuen Mittelwert und eine neue Standardabweichung für die analysierte(n) Variable(n) angeben oder den derzeitigen Mittelwert bzw. die derzeitige Standardabweichung beibehalten. **Hinweis:** Der gleiche Mittelwert und die gleiche Standardabweichung werden auf alle Analysevariablen angewendet. Sie können den Mittelwert bzw. die Standardabweichung jeder Variablen beibehalten, indem Sie das Kontrollkästchen **Neuer Mittelwert** bzw. **Neue Standardabweichung** deaktivieren.
Fehlende Werte durch Spaltenmittelwert ersetzen	Standardmäßig werden fehlende Werte in einer Eingabe-Analysevariable aus dem Standardisierungsprozess ausgeschlossen. Sie können die fehlenden Eingabewerte mit dem Mittelwert der nichtfehlenden Werte für diese Variable ersetzen. Markieren Sie hierzu das Kontrollkästchen **Fehlende Werte durch Spaltenmittelwert ersetzen**.

Daten standardisieren: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Die Option **Neue standardisierte Variablen hinzufügen** ist standardmäßig ausgewählt, so dass die standardisierten Variablen mit den ursprünglichen Variablen in die Ausgabedatei eingeschlossen werden. Um die Originalvariablen mit den standardisierten Variablen zu ersetzen, wählen Sie **Vorhandene Variablen ersetzen**.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

83

Beschreibende Statistiken

Allgemeines über die Anwendungsroutine 'Beschreibende Statistiken' ... **653**

Beschreibende Statistiken: Zuweisen von Variablen zu Analyserollen ... **654**

Beschreibende Statistiken: Auswählen beschreibender Statistiken ... **655**

Auswählen von Basisstatistiken ... 655

Auswählen von Perzentilstatistiken ... 656

Auswählen weiterer beschreibender Statistiken ... 656

Beschreibende Statistiken: Festlegen der Plotoptionen ... **657**

Beschreibende Statistiken: Festlegen von Ergebnisoptionen ... **658**

Festlegen von Titeln und Fußnoten ... **660**

Anzeigen von Eigenschaften ... **661**

Allgemeines über die Anwendungsroutine 'Beschreibende Statistiken'

Die Anwendungsroutine **Beschreibende Statistiken** stellt Tools zur Verdichtung von Daten zur Verfügung, mit denen beschreibende Statistiken für Variablen in allen Beobachtungen und innerhalb von Beobachtungsgruppen berechnet werden können. Sie können die Daten auch in einer grafischen Anzeige verdichten, wie z.B. in einem Histogramm.

Sie könnten z.B. mit dieser Anwendungsroutine einen Bericht über die Anzahl der Neuverkäufe nach Produkttyp und Land sortiert erstellen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	MEANS, UNIVARIATE
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

Beschreibende Statistiken: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Analysevariablen	gibt die Variable an, für die Statistiken generiert werden sollen. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Klassifizierungsvariablen	gibt die Variablen an, mittels derer die Eingabedaten in Kategorien oder Untergruppen aufgeteilt werden sollen. Diese Variablen können alphanumerisch oder numerisch sein. Die Statistiken werden für alle ausgewählten Variablen für jede eindeutige Kombination der Klassifizierungsvariablen berechnet. Sie können die Option **Mehrere Etiketten zulassen** aktivieren, so dass primäre und sekundäre Formatetiketten für eine beliebige Spannweite oder überlappende Spannweiten verwendet werden, um Untergruppenkombinationen zu erstellen, wenn einer Variablen ein Format mit mehreren Etiketten zugewiesen wird.

Optionsname	Beschreibung
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Relative Gewichtung	gibt die Berechnung der gewichteten Mittelwerte, Varianzen und Summen verwendete Variable an.
Variablen kopieren	Legt die Variablen für die Ausgabedatei fest. Sie können mithilfe der Option **Für die Rolle Variablen kopieren zu kopierender Wert** entweder den maximalen oder minimalen Wert, den die Variable in den entsprechenden Beobachtungen der Eingabedatei aufweist, in die Ausgabedatei kopieren.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert.

Beschreibende Statistiken: Auswählen beschreibender Statistiken

Auswählen von Basisstatistiken

Klicken Sie im Auswahlfenster unter der Überschrift **Statistiken** auf **Basis**, um auf diese Optionen zuzugreifen.

Standardmäßig enthält der Bericht **Beschreibende Statistiken** die folgenden Statistiken: Mittelwert, Standardabweichung, Minimalwert, Maximalwert und Anzahl der Beobachtungen. Sie können die Statistiken, die berechnet werden, die Anzahl der Dezimalstellen, die im Bericht angezeigt werden, sowie die Methode, die zum Berechnen der Quantilstatistiken verwendet wird, festlegen.

Optionsname	Beschreibung
Basisstatistiken	legt die einfachen Statistiken fest, die in die Analyse einbezogen werden sollen.

Optionsname	Beschreibung
Max. Dezimalstellen	legt die maximale Anzahl der Dezimalstellen in den berechneten Statistiken aus. Standardmäßig wird eine Statistik mit der Darstellung Optimal angezeigt, im Allgemeinen mit sieben Dezimalstellen.
Teiler für Standardabweichung und Varianz:	bestimmt den Teiler, der bei der Berechnung der Standardabweichungen und Varianzen verwendet wird. Standardmäßig entspricht der Teilers den Freiheitsgraden.

Auswählen von Perzentilstatistiken

Klicken Sie im Auswahlfenster unter der Überschrift **Statistiken** auf **Perzentile**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Perzentilstatistiken	gibt die zu berechnenden Perzentile an. Sie können auch das untere Quartil (25. Perzentil), den Median (50. Perzentil) und das obere Quartil (75. Perzentil) auswählen. Standardmäßig sind keine Perzentile ausgewählt.
Quantilmethode	gibt die Methode an, die zum Berechnen der Quantile, des Medians und der Perzentile verwendet wird. Sie können diese beiden Methoden auswählen: ■ **Ordnungstatistiken**, um eine Methode zu verwenden, die alle Daten in den Speicher einliest und nach den eindeutigen Werten sortiert. ■ **Stückweise parabolischer Algorithmus**, um eine weniger speicherintensive Methode zu verwenden, die das Quantil approximiert.

Auswählen weiterer beschreibender Statistiken

Klicken Sie im Auswahlfenster unter der Überschrift 'Statistiken' auf **Weitere**, um auf diese Optionen zuzugreifen.

Sie können auch folgende Statistiken berechnen:

- **Konfidenzgrenzen des Mittelwerts**

- **t-Statistik und Prob > |t|**
- **Variationskoeffizient**
- **Korrigierte Summe der Quadrate**
- **Unkorrigierte Summe der Quadrate**

Sie können auch das Konfidenzniveau für die Konfidenzgrenzen des Mittelwerts berechnen.

Beschreibende Statistiken: Festlegen der Plotoptionen

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Hinweis: Wenn Sie der Rolle **Klassifizierungsvariable** eine Rolle zugewiesen haben, werden die Plots nur für die *n*-fache Kombination der Klassifizierungsvariablen generiert.

Plot-Name	Beschreibung
Histogramm	Ein Histogramm ist ein Diagramm, mit dem die Verteilung eines Satzes an Daten bestimmt wird. Sie können Midpoints für das Histogramm festlegen. Die in der Dropdown-Liste verfügbaren Optionen hängen davon ab, ob Sie ein Histogramm oder ein Vergleichshistogramm erstellen. Wenn Sie der Rolle **Klassifizierungsvariablen** eine Variable zuweisen, erstellen Sie ein Vergleichshistogramm. Für Vergleichshistogramme sind die folgenden Optionen verfügbar: ■ **Midpoints angeben** - Ermöglicht das Angeben der Start- und End-Midpoints. Sie können auch das Schrittintervall festlegen, das für die Berechnung der anderen Midpoints verwendet werden soll. Der Midpoint ist der Mittelwert des Wertebereichs, der von den einzelnen Balken dargestellt wird. ■ **Schlüssel** - Bestimmt die Midpoints für die Daten in der Schlüsselzelle. Standardmäßig werden die Ausprägungen der Klassifizierungsvariablen von oben nach unten (links nach rechts) in aufsteigender Reihenfolge der internen (nicht formatierten) Werte der ersten Klassifizierungsvariablen angezeigt. Wenn Sie nur eine Klassifizierungsvariable angegeben haben, ist die Schlüsselzelle die Ausprägung, die als erstes in dieser Reihenfolge erscheint. Wenn Sie zwei Klassifizierungsvariablen angegeben haben, ist die Schlüsselzelle die Kombination der Ausprägungen von

Plot-Name	Beschreibung
	Variable 1 und Variable 2, die als erstes in dieser Reihenfolge erscheint. Die Auswahl der Schlüsselzelle bestimmt die einheitliche Horizontalachse, die für alle Zellen verwendet wird. Die Midpoint-Liste für die Schlüsselzelle wird dann wie erforderlich in beide Richtungen erweitert, bis sie die Daten in den übrigen Zellen abdeckt. ■ **Einheitlich** - Bestimmt die Anzahl der Midpoints abhängig von der Gesamt-Stichprobengröße. Wenn Sie der Rolle **Klassifizierungsvariablen** keine Variable zuweisen, erstellen Sie ein Histogramm. Für Histogramme stehen diese Optionen zur Verfügung: ■ **Midpoints angeben** - Ermöglicht das Angeben der Start- und End-Midpoints. Sie können auch das Schrittintervall festlegen, das für die Berechnung der anderen Midpoints verwendet werden soll. Der Midpoint ist der Mittelwert des Wertebereichs, der von den einzelnen Balken dargestellt wird. ■ **Anzahl der Intervalle angeben** - Ermöglicht, die Anzahl der Intervalle (auch Histogrammintervalle genannt) für die Daten anzugeben. Standardmäßig basiert die Anzahl der angezeigten Intervalle auf dem Datenbereich. Sie können jedoch eine Maximalzahl anzuzeigender Intervalle angeben.
Box-and-Whisker	Box-Plots zeigen ein Maß des zentralen Orts (der Median), zwei statistische Maße der Streuung (der Bereich und der Quartilsabstand), die Schiefe (die Ausrichtung des Medians auf die Quartilen) und potentielle Ausreißer. Box-Plots eignen sich besonders zum Vergleichen zweier oder mehrerer Datensätze.

Beschreibende Statistiken: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Statistiken in Datei speichern	legt fest, ob die Statistiken dauerhaft in einer Datentabelle gespeichert werden sollen. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und

Optionsname	Beschreibung
	speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Auswählen der in die Ergebnisse einzubeziehende n Statistiken	gibt die in die Ergebnisse einzubeziehenden Statistiken an. Folgende Optionen können ausgewählt werden: ■ **Für die Rolle "Variable kopieren" zu kopierender Wert** - Wenn Sie der Rolle Variable kopieren eine Variable zuweisen, können Sie entweder den maximalen oder minimalen Wert, den die Variable in den entsprechenden Beobachtungen der Eingabedatei aufweist, in die Ausgabedatei kopieren. ■ **Kombinationen der Klassifizierungsvariablen** - Mit der Rolle **Klassifizierungsvariable** können Sie die beschreibenden Statistiken für die Klassifizierungsvariablen in Tabellen gruppieren. Sie können bis zu n-fache Tabellen erstellen, wobei n der Anzahl der Klassifizierungsvariablen entspricht. Die Gruppen werden durch die eindeutigen Kombinationen der Klassifizierungsvariablen festgelegt. Die Reihenfolge der Variablen in der Tabelle hängt von der Reihenfolge der Variablen in der Rolle **Klassifizierungsvariablen** ab. Mit den folgenden Optionen können Sie die Ausgabedaten in Untergruppen einteilen: □ **Nur N-fach** erstellt eine einzelne Tabelle, die die beschreibenden Statistiken für die n-fache Kombination der Klassifizierungsvariablen anzeigt. Angenommen, Sie möchten die Umsätze aus zwei verschiedenen Marketingkampagnen vergleichen. Sie weisen die Variable Profit der Rolle **Analysevariablen** und die Variablen Month, Marketing Campaign und Region der Rolle **Klassifizierungsvariablen** zu. Wenn Sie nun die Option **Nur N-fach** auswählen, erhalten Sie eine Tabelle Month*Campaign*Region. □ **Vollständig** erstellt eine separate Tabelle mit beschreibenden Statistiken für jede eindeutige Kombination der Klassifizierungsvariablen. Wenn Sie z. B. die Variablen 'Month', 'Campaign' und 'Region' verwenden und dann die Option **Vollständig** auswählen, werden folgende Tabellen erstellt: ▪ eine Tabelle mit den Gesamtumsätzen. Das ist eine sogenannte nullfache Tabelle, und die Klassifizierungsvariablen sind nicht eingeschlossen. ▪ eine einfache Tabelle, welche die beschreibenden Statistiken für jeden Monat (z.B. Mai, Juni und Juli) anzeigt

Optionsname	Beschreibung
	▪ eine einfache Tabelle, welche die beschreibenden Statistiken für jede Region (z.B. Norden, Süden, Osten und Westen) anzeigt ▪ eine einfache Tabelle, welche die beschreibenden Statistiken für jede Art von Marketingkampagne (z.B. Werbespot, Zeitungsannonce) anzeigt ▪ eine zweifache Tabelle, welche die beschreibenden Statistiken für Month*Region anzeigt ▪ eine zweifache Tabelle, welche die beschreibenden Statistiken für Month*Campaign anzeigt ▪ eine zweifache Tabelle, welche die beschreibenden Statistiken für Region*Campaign anzeigt ▪ eine dreifache Tabelle, welche die beschreibenden Statistiken für Month*Campaign*Region anzeigt □ **Ordnung n angeben** erstellt eine separate Tabelle mit beschreibenden Statistiken für jede eindeutige Kombination der von Ihnen angegebenen Klassifizierungsvariablen. Wenn Sie beispielsweise eine `1`, `2` oder `3` in das Feld **Ordnung n angeben** eingeben, enthalten die Ergebnisse alle möglichen einfachen, zweifachen oder dreifachen Tabellen. Wenn Sie in diesem Feld `0` eingeben, enthalten die Ergebnisse eine nullfache Tabelle, die keine Klassifizierungsvariablen einschließt.
Analyseetiketten anzeigen	zeigt die Etiketten für die Analysevariablen an.
Anzeige für gesamte Ergebnisausgabe unterdrücken	zeigt keine Ergebnisse an. Wenn Sie die Anzeige der Ausgabe unterdrücken, müssen Sie die Ergebnisse in eine Ausgabedatei speichern.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

84

Assistent für beschreibende Statistiken

Allgemeines über den Assistenten für beschreibende Statistiken **663**

Assistent für beschreibende Statistiken: Auswählen der Daten **663**

Assistent für beschreibende Statistiken: Zuweisen von Variablen zu Rollen **664**

Assistent für beschreibende Statistiken: Festlegen von Berichtsstatistiken und Ausgabergebnissen **665**

Assistent für beschreibende Statistiken: Festlegen von Titeln und Fußnoten **666**

Allgemeines über den Assistenten für beschreibende Statistiken

Der **Assistent für beschreibende Statistiken** unterstützt Sie bei der Erstellung von Tabellen beschreibender Statistiken, Histogrammen und Box-Plots mit optionalen Gruppierungsvariablen.

Assistent für beschreibende Statistiken: Auswählen der Daten

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im

Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Klicken Sie auf **Weiter**, um Variablen Rollen zuzuweisen.

Assistent für beschreibende Statistiken: Zuweisen von Variablen zu Rollen

1. Bevor Sie Variablen zu Rollen zuweisen können, müssen Sie die Daten auswählen.
2. Weisen Sie der Rolle **Beschreibende Statistiken von (Analysevariable)** mindestens eine Variable zu. Die Variablen, die Sie dieser Rolle zuweisen, sind die numerischen Variablen, zu denen Sie Statistiken generieren möchten.

   Hinweis: Sie müssen dieser Rolle mindestens eine Variable zuweisen. Alle anderen Rollen sind optional.
3. Weisen Sie der Rolle **Für jeden Wert von: (Klassifizierungsvariable)** mindestens eine Variable zu. Die Variablen, die Sie dieser Rolle zuweisen, sind alphanumerische oder diskrete numerische Variablen, die zum Unterteilen der Eingabedaten in Kategorien oder Untergruppen verwendet werden. Die Statistiken werden für alle ausgewählten Variablen für jede eindeutige Kombination der Klassifizierungsvariablen berechnet.
4. Weisen Sie der Rolle **Separate Tabellen für Werte von (Gruppierungsvariable)** mindestens eine Variable zu. Die Variablen, die Sie dieser Rolle zuweisen, werden für die Berechnung separater Statistiken für jeden Einzelwert oder jede Kombination von Werten der **Analyse gruppieren nach-**Variablen verwendet. Die Daten werden automatisch nach der Variablen in dieser Rolle sortiert, bevor Statistiken berechnet werden. Verwenden Sie die Nach-oben- und Nach-unten-Pfeile, um die Sortierfolge für jede Variable festzulegen.
5. (Optional) Klicken Sie auf **Erweitert**, um Variablen zu diesen Rollen zuzuweisen:
   - **Häufigkeit angeben nach (Häufigkeitszählung)** - Wenn Sie dieser Rolle eine Variable zuweisen, wird angenommen, dass jede Beobachtung in der Tabelle n Beobachtungen darstellt, wobei n der Wert der Häufigkeitszählung für diese Zeile ist. Statistiken werden entsprechend berechnet. Sie können dieser Rolle maximal eine Variable zuweisen.
   - **Gewichtungsanalyse nach (Relative Gewichtung)** - Wenn Sie dieser Rolle eine Variable zuweisen, werden mit dem Wert der Variablen für jede Beobachtung gewichtete Mittelwerte, Varianzen und Summen berechnet. Sie können dieser Rolle maximal eine Variable zuweisen.

- **Etiketten einschließen von (Variable kopieren)** - Variablen, die dieser Rolle zugewiesen sind, werden als weitere Variablen in der Ausgabedatei eingeschlossen.

6 Klicken Sie auf **Weiter**, um die Statistiken und Ergebnisse anzugeben.

Assistent für beschreibende Statistiken: Festlegen von Berichtsstatistiken und Ausgabergebnissen

1 Bevor Sie die Berichtsstatistiken und die Ausgabeergebnisse angeben können, müssen Sie den Rollen Variablen zuweisen.

2 Wählen Sie die Statistiken aus, die in die Ergebnisse einbezogen werden sollen. Standardmäßig wird der Mittelwert, die Standardabweichung, Minimum, Maximum und Anzahl der Beobachtungen einbezogen. Sie können zusätzliche Statistiken hinzufügen, indem Sie die Optionen auf diesen Registerkarten verwenden:

Basis
schließt die beschreibenden Statistiken ein. Sie können auch die maximale Anzahl Dezimalstellen für die berechneten Statistiken festlegen. Standardmäßig wird eine Statistik mit der Darstellung Optimal angezeigt, im Allgemeinen mit sieben Dezimalstellen.

Perzentil
gibt die zu berechnenden Perzentile an. Sie können auch das untere Quartil (25. Perzentil), den Median (50. Perzentil) und das obere Quartil (75. Perzentil) auswählen. Standardmäßig sind keine Perzentile ausgewählt.

Verwenden Sie die Option **Quantilmethode**, um die Methode zur Berechnung von Quantilen, Median und Perzentilen auszuwählen. Sie können diese beiden Methoden auswählen:

- **Ordnungstatistiken**, um eine Methode zu verwenden, die alle Daten in den Speicher einliest und nach den eindeutigen Werten sortiert.
- **Stückweise parabolischer Algorithmus**, um eine weniger speicherintensive Methode zu verwenden, die das Quantil approximiert.

Weitere
beinhaltet diese Statistiken:

- **Konfidenzgrenzen des Mittelwerts**
- **t-Statistik und Prob > |t|**
- **Variationskoeffizient**
- **Korrigierte Summe der Quadrate**

- **Unkorrigierte Summe der Quadrate**

  Sie können auch das Konfidenzniveau für die Konfidenzgrenzen des Mittelwerts berechnen.

3 Wählen Sie die Option **Statistiken anzeigen**, um die Ergebnisse anzuzeigen. Sie müssen die Statistiken entweder in den Ergebnissen anzeigen oder sie permanent in einer Datentabelle speichern oder beides.

4 Aktivieren Sie das Kontrollkästchen **Histogramm** oder **Box-and-Whisker**, um den gewünschten Plottyp auszuwählen. Sie können einen oder beide auswählen. Standardmäßig sind keine Plots ausgewählt.

Hinweis: Wenn Sie der Rolle **Klassifizierungsvariable** eine Rolle zugewiesen haben, werden die Plots nur für die *n*-fache Kombination der Klassifizierungsvariablen generiert.

5 Aktivieren Sie zum Speichern der Ausgabe das Kontrollkästchen **Statistiken in Datei speichern**. Sie müssen die Statistiken entweder in den Ergebnissen anzeigen oder sie permanent in einer Datentabelle speichern oder beides.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

6 Klicken Sie auf **Weiter**, um die Titel und Fußnoten anzugeben.

Assistent für beschreibende Statistiken: Festlegen von Titeln und Fußnoten

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

1 Bevor Sie die Titel und Fußnoten festlegen können, müssen Sie die Berichtsstatistiken und Ausgabeergebnisse angeben.

2 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld. Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Hinweis: Einige Titel sind möglicherweise nicht verfügbar. Das hängt von den Plots ab, die Sie zum Einbeziehen in die Ergebnisse ausgewählt haben.

3 Klicken Sie auf **Fertig stellen**, um die Anwendungsroutine auszuführen.

85

Zusammenfassungstabellen

Allgemeines über die Anwendungsroutine 'Zusammenfassungstabellen' **669**

Zusammenfassungstabellen: Zuweisen von Variablen zu Analyserollen **670**

Zusammenfassungstabellen: Einrichten der Zusammenfassungstabelle **671**

Erstellen the Zusammenfassungstabelle 671

Anpassen der Darstellung der Tabelle 672

Zusammenfassungstabellen: Festlegen von Ergebnisoptionen **674**

Festlegen von Titeln und Fußnoten **674**

Anzeigen von Eigenschaften **675**

Allgemeines über die Anwendungsroutine 'Zusammenfassungstabellen'

Die Anwendungsroutine **Zusammenfassungstabellen** zeigt beschreibende Statistiken in einem tabellarischen Format an. Dabei werden einige oder alle Variablen in einer Datei verwendet. Sie können verschiedene Tabellen erstellen, die von einfachen Tabellen bis hin zu individuell angepassten Tabellen reichen.

Angenommen, es liegt eine SAS-Datei mit den Umsatzdaten eines Produkts für zwei Kundentypen - Privat- und Geschäftskunden - in verschiedenen amerikanischen Bundesstaaten im Nordosten und im Westen der USA vor. Mithilfe der Anwendungsroutine "Zusammenfassungstabelle" können Sie die Verkäufe für jeden Kundentyp in jedem Staat einer geographischen Region zusammenfassen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	TABULATE
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	Ohne

Zusammenfassungstabellen: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Analysevariablen	gibt die Variablen für die Analyse an. Sie müssen dieser Rolle mindestens eine Variable oder die Rolle **Klassifizierungsvariablen** zuweisen.
Klassifizierungsvariablen	gibt die alphanumerischen oder diskreten numerischen Variablen an, die zum Unterteilen der Eingabedaten in Kategorien oder Untergruppen verwendet werden sollen. Die Statistiken werden für alle ausgewählten Variablen für jede eindeutige Kombination der Klassifizierungsvariablen berechnet. Sie können die Option **Mehrere Etiketten** aktivieren, so dass primäre und sekundäre Formatetiketten für eine beliebige Spannweite oder überlappende Spannweiten verwendet werden, um Untergruppenkombinationen zu erstellen, wenn einer Variablen ein Format mit mehreren Etiketten zugewiesen wird.
Seiten	erzeugt eine separate Zusammenfassungstabelle für jeden Einzelwert dieser Variable.

Optionsname	Beschreibung
Relative Gewichtung	gibt die Berechnung der gewichteten Mittelwerte, Varianzen und Summen verwendete Variable an.
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.

Zusammenfassungstabellen: Einrichten der Zusammenfassungstabelle

Erstellen the Zusammenfassungstabelle

Klicken Sie im Auswahlfenster auf **Zusammenfassungstabellen**, um auf diese Optionen zuzugreifen.

Um eine Zusammenfassungstabelle einzurichten, müssen Sie festlegen, welche Informationen enthalten sein sollen und wie diese erscheinen sollen. Die Anwendungsroutine **Zusammenfassungstabellen** ist sehr flexibel. Sie können Klassifizierungs- und Analysevariablen beliebig anordnen.

So verwenden Sie eine Variable:

1. Wählen Sie die Variable im Feld **Verfügbare Variablen** aus.
2. Ziehen Sie die Variable in den oberen oder den linken Teil des Felds **Vorschau**. Wenn Sie die neue Variable auf einer bestehenden Variablen ablegen, ersetzt die neue Variable die alte. Um die Variable zu den bestehenden Variablen hinzuzufügen, legen Sie sie auf einem Zellenrand ab. Der Rand wird blau, wenn der Mauszeiger an der richtigen Stelle ist.

Hinweis: Standardmäßig ist die Statistik für die verwendeten Analysevariablen die Anzahl der Beobachtungen (N). Sie können mithilfe der oben beschriebenen Methoden Statistiken hinzufügen oder ersetzen.

Anpassen der Darstellung der Tabelle

Klicken Sie mit der rechten Maustaste auf die Vorschautabelle, um die Optionen einzurichten, mit denen die Darstellung der Zusammenfassungstabelle angepasst wird. Mit Ausnahme von **Zellen entfernen** öffnen die Elemente im Kontextmenü jeweils ein Fenster, in dem Sie verschiedene Optionen festlegen können.

Plot-Name	Beschreibung
Zellen entfernen	Dieses Element ist aktiviert, wenn Sie mit der rechten Maustaste auf eine Analysevariable, eine Klassifizierungsvariable oder eine Statistik in der Vorschautabelle klicken. Mit diesem Element können Sie die Variable oder Statistik aus der Zusammenfassungstabelle entfernen.
Tabelleneigenschaften	Dieses Element ist immer aktiviert. Optionen, die im Dialogfeld **Tabelleneigenschaften** eingestellt werden, gelten für die gesamte Zusammenfassungstabelle. ■ Reiter **Allgemein**: Geben Sie ein Etikett für fehlende Werte an. Der Standard ist ein Punkt ('.'). Mit Hilfe der Kontrollkästchen können Sie fehlende Werte als zulässige Werte für Klassifizierungsvariable einbeziehen und die Überschriften (Etiketten) für die Zeilenklassifizierungsvariablen unterdrücken. ■ Reiter **Formatieren**: Weisen Sie allen Datenwerten in der Zusammenfassungstabelle ein Standard-SAS-Format oder ein benutzerdefiniertes Format zu. Wählen Sie das Format, die Feldbreite und die Anzahl der Dezimalstellen aus. ■ Reiter **Schriftart**: Geben Sie die Schriftart, Schriftgröße, den Schriftschnitt, die Vordergrund- und Hintergrundfarbe, die Ausrichtung und die vertikale Ausrichtung für alle Texte in der Tabelle an.
Rahmenbereicheigenschaften	Dieses Element ist immer aktiviert. Legen Sie mit diesem Element die Optionen für die obere linke Ecke der Tabelle fest. Diese Optionen überschreiben alle Optionen, die im Dialogfeld **Tabelleneigenschaften** ausgewählt wurden. ■ Reiter **Allgemein**: Geben Sie den Text ein, der im Rahmenbereich verwendet werden soll. Wenn Sie der Rolle Seiten mindestens eine Variable zugewiesen haben, können Sie die Seitenüberschriften im Rahmenbereich anzeigen lassen. Sie können auch auswählen, ob der Text umgebrochen oder gekürzt werden soll, wenn er für den Rahmenbereich zu lang ist. ■ Reiter **Formatieren**: Dieser Reiter gilt nicht für den Rahmenbereich.

Plot-Name	Beschreibung
	■ Reiter **Schriftart**: Geben Sie die Schriftart, Schriftgröße, den Schriftschnitt, die Vordergrund- und Hintergrundfarbe, die Ausrichtung und die vertikale Ausrichtung für den Text im Rahmenbereich an.
Überschriftenei genschaften	Sie müssen mit der rechten Maustaste auf eine Spalten- oder Zeilenüberschrift klicken, um dieses Element zu aktivieren. Mit diesem Element legen Sie die Optionen für einzelne Spalten- oder Zeilenüberschriften fest. Diese Optionen überschreiben alle Optionen, die im Dialogfeld **Tabelleneigenschaften** ausgewählt wurden. ■ Reiter **Allgemein**: Geben Sie ein Etikett für die Zeilen- oder Spaltenüberschrift ein. Sie können auch auswählen, ob der Text umgebrochen oder gekürzt werden soll, wenn er für die Überschrift zu lang ist. ■ Reiter **Formatieren**: Dieser Reiter gilt nicht für Klassifizierungsvariablenwerte. Um einer Klassifizierungsvariablen ein Überschriftformat zuzuweisen, klicken Sie im Auswahlfenster auf **Daten** und wählen die Klassifizierungsvariable aus. Wählen Sie die Option Überschriftformat, um das Format zuzuweisen. ■ Reiter **Schriftart**: Geben Sie die Schriftart, Schriftgröße, den Schriftschnitt, die Vordergrund- und Hintergrundfarbe, die Ausrichtung und die vertikale Ausrichtung für den Text der Überschrift an.
Eigenschaften der Klassifizierung sausprägungen	Dieses Element ist nur dann aktiviert, wenn Sie mit der rechten Maustaste in der Vorschautabelle auf eine Klassifizierungsvariable klicken. Legen Sie mit diesem Element die Optionen für die Klassifizierungsvariablenwerte fest. ■ Reiter **Allgemein**: Legen Sie fest, ob der Text für den Klassifizierungsvariablenwert umgebrochen werden soll. ■ Reiter **Formatieren**: Dieser Reiter gilt nicht für Klassifizierungsvariablenwerte. Um einer Klassifizierungsvariablen ein Überschriftformat zuzuweisen, klicken Sie im Auswahlfenster auf **Daten** und wählen die Klassifizierungsvariable aus. Wählen Sie die Option **Überschriftformat**, um das Format zuzuweisen. ■ Reiter **Schriftart**: Geben Sie die Schriftart, Schriftgröße, den Schriftschnitt, die Vordergrund- und Hintergrundfarbe, die Ausrichtung und die vertikale Ausrichtung für den Text des Klassifizierungsvariablenwertes an.
Datenwerteige nschaften	Sie aktivieren dieses Element, indem Sie mit der rechten Maustaste in der Vorschautabelle auf eine Variable oder Statistik klicken. Mit diesem Element können Sie die Optionen für die Datenwerte in der ausgewählten Zeile oder Spalte der Zusammenfassungstabelle

Plot-Name	Beschreibung
	festlegen. Diese Optionen überschreiben alle Optionen, die im Dialogfeld **Tabelleneigenschaften** ausgewählt wurden. ■ Reiter **Allgemein**: Legen Sie fest, ob der Text für den Datenwert umgebrochen werden soll. ■ Reiter **Formatieren**: Weisen Sie allen Datenwerten in der ausgewählten Zeile oder Spalte der Zusammenfassungstabelle ein Standard-SAS-Format oder ein benutzerdefiniertes Format zu. Wählen Sie das Format, die Feldbreite und die Anzahl der Dezimalstellen aus. ■ Reiter **Schriftart**: Geben Sie die Schriftart, Schriftgröße, den Schriftschnitt, die Vordergrund- und Hintergrundfarbe, die Ausrichtung und die vertikale Ausrichtung für den Text der Datenwerte an.

Zusammenfassungstabellen: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Um Ihre Ergebnisse in eine Ausgabedatei zu speichern, aktivieren Sie das Kontrollkästchen **Ergebnisse in Datei speichern**.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

86

Zusammenfassungstabellen-Assistent

Allgemeines über den Assistenten für Zusammenfassungstabellen 677
Zusammenfassungstabellen-Assistent: Auswählen der Daten 677
Zusammenfassungstabellen-Assistent: Zuweisen von Variablen zu Rollen 678
Zusammenfassungstabellen-Assistent: Auswählen der Klassifizierungsvariablen 679
Zusammenfassungstabellen-Assistent: Angeben der Summen 679
Zusammenfassungstabellen-Assistent: Auswählen zusätzlicher Ausgaben 680
Zusammenfassungstabellen-Assistent: Festlegen von Titeln und Fußnoten 680

Allgemeines über den Assistenten für Zusammenfassungstabellen

Der Assistent **Zusammenfassungstabellen** hilft Ihnen bei der Erstellung von ganz unterschiedlichen Tabellen, inkl. Kreuztabellen. Es stehen auch Optionen für beschreibende Statistiken zur Verfügung.

Zusammenfassungstabellen-Assistent: Auswählen der Daten

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine

ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Klicken Sie auf **Weiter**, um Analysevariablen und Statistiken zuzuweisen.

Zusammenfassungstabellen-Assistent: Zuweisen von Variablen zu Rollen

1. Bevor Sie Variablen Rollen zuweisen können, müssen Sie die Daten auswählen.
2. Wenn Sie dem Feld **Analysevariablen** eine Variable hinzufügen möchten, klicken Sie auf **Hinzufügen**. Wählen Sie aus dem Popup-Menü die Variable aus, die dieser Rolle hinzugefügt werden soll. Die Variablen, die Sie dieser Rolle zuweisen, sind die numerischen Variablen für die Berechnungen.
3. Wählen Sie im Feld **Analysevariablen** die Statistik aus, die Sie für jede Analysevariable berechnen möchten. Standardmäßig wird die Statistik **Summe** berechnet. Wenn Sie eine andere Statistik berechnen möchten, markieren Sie die Zelle der Statistik in der Spalte **Statistik**, die Sie ändern möchten. Wählen Sie in der Dropdown-Liste die zu berechnende Statistik aus.

   Hinweis: Analysevariablen können nicht zugewiesen werden, wenn die Daten keine numerischen Variablen enthalten. Ist dies der Fall, so wird die Häufigkeitsstatistik anstelle einer Analysevariablen verwendet.

4. Wählen Sie aus der Dropdown-Liste **Etiketten für Analysevariablen** den Speicherort der Etiketten für die Analysevariablen aus. Es stehen die Optionen **in Spalten**, **in Zeilen**, oder **in Seiten** oder **verborgen** zur Verfügung.
5. Wählen Sie aus der Dropdown-Liste **Etiketten für Statistiken** den Speicherort der Etiketten für die Statistiken aus. Es stehen die Optionen **in Spalten**, **in Zeilen**, oder **in Seiten** oder **verborgen** zur Verfügung.
6. Klicken Sie auf **Weiter**, um die Statistiken und Ergebnisse anzugeben.

Zusammenfassungstabellen-Assistent: Auswählen der Klassifizierungsvariablen

1. Vor der Auswahl der Klassifizierungsvariablen können Sie die Analysevariablen und Statistiken zuweisen.
2. Um eine Klassifizierungsvariable den Spalten, Zeilen oder Seiten hinzuzufügen, klicken Sie im entsprechenden Feld auf **Hinzufügen**. Wählen Sie aus dem Popup-Menü die Variable aus, die hinzugefügt werden soll. Die Variablen, die Sie als Klassifizierungsvariablen zuweisen, sind alphanumerische oder diskrete numerische Variablen, mit denen die Eingabedaten in Kategorien unterteilt werden. Die Statistiken werden für alle ausgewählten Variablen für jede eindeutige Kombination der Klassifizierungsvariablen berechnet. Für jede Variable, die Sie dem Feld **Seiten** zuweisen wird eine separate Zusammenfassungstabelle generiert.
3. Klicken Sie auf **Weitere Optionen**, um die Tabellenstruktur anzupassen. Im Dialogfeld **Tabellenstruktur** können Sie die Überschriften für die Spalten, Zeilen und Seiten verbergen. Sie können ebenso die Spalten, Zeilen und Tabellen verschachteln oder zusammensetzen.
4. Klicken Sie auf **Weiter**, um die Titel und Fußnoten anzugeben.

Zusammenfassungstabellen-Assistent: Angeben der Summen

1. Bevor Sie die Summen angeben können, müssen Sie die Klassifizierungsvariablen festlegen.
2. Wählen Sie in der Dropdown-Liste **Spalten**, **Zeilen** und **Seiten** aus, und ob Sie die Gesamtsumme für jede Spalte, Zeile oder Seite anzeigen möchten. Wenn Sie die Summe anzeigen möchten, wählen Sie **Nur Gesamtsumme**. Wenn Sie die Summe nicht anzeigen möchten, wählen Sie **Ohne**. Bei Seiten können Sie auch festlegen, ob die Seitensummen am Anfang oder am Ende der Seite angezeigt werden sollen.

Hinweis: Wenn Sie für die Spalten, Zeilen oder Seiten keine Klassifizierungsvariable festgelegt haben, ist die entsprechende Option nicht verfügbar.

3 Geben Sie im Feld **Summenetikett** das Etikett für die Summe ein.

4 Klicken Sie auf **Weiter**, um die Titel und Fußnoten anzugeben.

Zusammenfassungstabellen-Assistent: Auswählen zusätzlicher Ausgaben

1 Vor der Auswahl zusätzlicher Ausgabe können Sie die Summen angeben.

2 Aktivieren Sie das Kontrollkästchen **Ergebnisse in Datei speichern**.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

3 Klicken Sie auf **Weiter**, um die Titel und Fußnoten anzugeben.

Zusammenfassungstabellen-Assistent: Festlegen von Titeln und Fußnoten

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

1 Bevor Sie den Titel und die Fußnoten angeben, können Sie angeben, ob die Ergebnisse in einer Ausgabedatei gespeichert werden sollen.

2 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld. Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Hinweis: Einige Titel sind möglicherweise nicht verfügbar. Das hängt von den Plots ab, die Sie zum Einbeziehen in die Ergebnisse ausgewählt haben.

3 Klicken Sie auf **Fertig stellen**, um die Anwendungsroutine auszuführen.

87

Oberflächenplot

Allgemeines über die Anwendungsroutine 'Oberflächenplot' ... **683**

Oberflächenplot: Auswählen eines Diagrammtyps ... **684**

Oberflächenplot: Zuweisen von Variablen zu Analyserollen ... **685**

Oberflächenplot: Festlegen von Darstellungsoptionen ... **686**
- Festlegen von Optionen für Oberflächenplots ... 686
- Weitere Optionen ... 686
- Glätten von Daten ... 687

Festlegen von Titeln und Fußnoten ... **687**

Anzeigen von Eigenschaften ... **688**

Allgemeines über die Anwendungsroutine 'Oberflächenplot'

Die Anwendungsroutine **Oberflächenplot** erstellt dreidimensionale Drahtrahmenplots, dreidimensionale Glättungsplots oder dreidimensionale Gradientenplots, die die mathematischen Beziehungen zwischen drei numerischen Variablen aufzeigen. Im dreidimensionalen Plot wird eine vertikale Variable (z) für eine Position auf einer Ebene dargestellt, die durch zwei horizontale Variablen (x und y) angegeben wird. Die Koordinaten jedes Punkts entsprechen den Werten der drei numerischen Variablenwerte in einer Beobachtung der Eingabedatei. Die Beobachtung kann Werte in der Form z = f(x,y) oder unabhängige Werte enthalten, wie die Höhe bei einer gegebenen Länge und Breite.

Weitere Informationen, siehe “Auswählen des Grafikausgabeformats” auf Seite 6.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	G3D
Erforderliche SAS-Produkte	Base SAS, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	Ohne

Oberflächenplot: Auswählen eines Diagrammtyps

Klicken Sie im Auswahlfenster auf **Oberflächenplot**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
3D Drahtrahmen-Oberflächenplot	erstellt einen dreidimensionalen Plot, in dem das Drahtrahmengitter die mathematische Beziehung zwischen drei Variablen anzeigt.
3D Glättungsplot	erstellt einen dreidimensionalen Plot, in dem die gesamte Oberfläche in einer Farbe dargestellt wird.
3D Gradientenplot mit Bändern	erstellt einen dreidimensionalen Plot, der mithilfe von Farben eine Wertänderung angibt. Die Bereiche sind klar definiert.
3D Gradientenplot mit Glättung	erstellt einen dreidimensionalen Plot, der mithilfe von Farben eine Wertänderung angibt. Die Bereiche sind nicht so klar definiert wie im Gradientenplot mit Bändern.

Wie das Oberflächenplot angezeigt wird, hängt vom ausgewählten Grafikausgabeformat ab.

ActiveX und ActiveX image
: rendert alle Oberflächenplottypen.

Java und Java-Bild
: rendert das dreidimensionale Drahtrahmen-Oberflächenplot. Das dreidimensionale Glättungsplot, das dreidimensionale Gradientenplot mit Bändern und das dreidimensionale Gradientenplot mit Glättung werden als dreidimensionale Gradientenplots mit Glättung gerendert.

GIF, JPEG, PNG und SAS EMF
: rendert alle Oberflächenplottypen als dreidimensionale Drahtrahmen-Oberflächenplots.

Oberflächenplot: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Horizontal	Die Spalte, die Sie dieser Rolle zuweisen, ist die horizontale oder X-Achsenvariable für das Diagramm.
Vertikal	Die Spalte, die Sie dieser Rolle zuweisen, ist die vertikale oder Y-Achsenvariable für das Diagramm.
Tiefe	Die Spalte, die Sie dieser Rolle zuweisen, ist die Variable der dritten Dimension bzw. Z-Achse für das Diagramm. **Hinweis:** Diese Option ist nur für dreidimensionale Plots verfügbar.
Diagramme gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Diagramme, die erstellt werden. Für jeden eindeutigen Wert wird eine separate Grafik erzeugt.

Oberflächenplot: Festlegen von Darstellungsoptionen

Festlegen von Optionen für Oberflächenplots

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Oberfläche**, um auf diese Optionen zuzugreifen.

Wählen Sie eine Farbe für die Diagrammoberseite (positive Z-Richtung) und -unterseite (negative Z-Richtung) aus.

Wählen Sie aus, ob auch die Diagrammseiten gezeichnet werden sollen.

Stellen Sie über die Schieberegler oder die Steuerelemente **Drehen** und **Neigen** den Ansichtswinkel ein. Standardmäßig betragen Dreh- und Neigungswinkel 70 Grad. Klicken Sie auf **Zurücksetzen**, um den Betrachtungswinkel auf die Standardeinstellung zurückzusetzen. Hinweis

Hinweis: Wenn Sie das Kontrollkästchen **Werte für Drehen und Neigen angeben** deaktivieren, wird das Drehen und Neigen der Grafik durch das Grafikausgabeformat bestimmt. Standardmäßig ist diese Option aktiviert und Sie können den Sichtwinkel festlegen.

Bei 3D Gradienten mit Bändern und 3D-Gradientenplots mit Glättung können Sie auswählen, ob zusammen mit der Oberfläche eine Kontur der Daten gezeichnet werden soll. Sie können die Kontur ober- oder unterhalb der Oberfläche anzeigen.

Weitere Optionen

Einige Darstellungsoptionen gelten für alle Grafik-Anwendungsroutinen. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Festlegen von Achsenoptionen” auf Seite 745
- “Anpassen des Diagrammbereichs” auf Seite 750

Glätten von Daten

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Gitterwerte**, um auf diese Optionen zuzugreifen.

Sie möchten ein rechteckiges Gitter interpolierter oder geglätteter Werte erstellen, wenn:

- die Datenpunkte in der Eingabetabelle ungleichmäßig verteilt sind
- die Eingabetabelle nicht genügend Zeilen aufweist, um ein dreidimensionales Oberflächenplot zu erzeugen
- die Daten in der Eingabetabelle "rauschen".

Standardmäßig wird keine Interpolationsmethode verwendet. Sie können jedoch aus folgenden Interpolationsmethoden auswählen:

Joinen
verwendet eine lineare Interpolation innerhalb eines Satzes dreieckiger Bereiche, die aus der Eingabedatei gebildet werden. Diese Interpolationsmethode erstellt Werte im Bereich der Anfangswerte der vertikalen Variablen. Die resultierende interpolierte Oberfläche ist in der Regel jedoch nicht geglättet.

Ohne Skalierung
gibt an, dass die X- und Y-Variablen vor der Interpolation nicht auf denselben Bereich skaliert werden. Standardmäßig wird die Interpolation durchgeführt, nachdem beide Variablen auf denselben Bereich skaliert wurden, weil die Interpolationsmethoden davon ausgehen, dass die X- und Y-Skala vergleichbar sind.

Partiell
gibt an, dass ein Spline zum Schätzen der Ableitungen für die biquintische Polynominterpolation verwendet wird. Ein bivariater Spline wird an die nächsten Nachbarn angepasst. Er wird verwendet, um die benötigten Ableitungen zu schätzen.

Spline
gibt an, dass ein bivariater Spline für die Interpolation verwendet wird. Nach dieser Option muss ein Algorithmus der Ordnung n3 verwendet werden, wobei n die Anzahl der Eingabedatenpunkte darstellt. Diese Methode kann sehr zeitaufwändig sein.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

88

t-Test

Allgemeines über die Anwendungsroutine 't-Test' 689
t-Test: Auswahl des t-Test-Typs 690
t-Test: Zuweisen von Variablen zu Analyserollen 690
t-Test: Festlegen der Analyseoptionen 692
t Test: Erzeugen von Plots 692
Festlegen von Titeln und Fußnoten 693
Anzeigen von Eigenschaften 694

Allgemeines über die Anwendungsroutine 't-Test'

Die Anwendungsroutine *t*-Test führt *t*-Tests für eine Stichprobe, zwei Stichproben und paarweise Beobachtungen aus. Beim *t*-Test mit einer Stichprobe wird der Mittelwert der Stichprobe mit einer bestimmten Zahl verglichen. Beim *t*-Test mit zwei Stichproben wird der Mittelwert der ersten Stichprobe abzüglich des Mittelwerts der zweiten Stichprobe mit einer bestimmten Zahl verglichen. Beim *t*-Test mit paarweisen Beobachtungen wird der Mittelwert der Beobachtungsdifferenzen mit einer bestimmten Zahl verglichen.

Manchmal kann nicht von der Annahme ausgegangen werden, dass zwei Datengruppen unabhängig sind und dass eine natürliche Paarbildung der Daten vorliegt. In diesem Fall ist es von Vorteil, einen paarweisen *t*-Test zu verwenden. Angenommen, es wird ein Stimulus untersucht, um dessen Auswirkung auf den systolischen Blutdruck zu ermitteln. Zwölf Männer nehmen an dieser Studie teil. Dabei wird ihr systolischer Blutdruck vor und nach der Anwendung des Stimulus gemessen. Verwenden Sie die Anwendungsroutine *t*-Test, um festzustellen, ob die mittlere Änderung des systolischen Blutdrucks erheblich von Null abweicht.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	TTEST
Erforderliche SAS-Produkte	Base SAS, SAS/STAT
Empfohlene zusätzliche SAS-Produkte	SAS/GRAPH

t-Test: Auswahl des t-Test-Typs

Klicken Sie im Auswahlfenster auf **t-Test-Typ**, um auf diese Optionen zuzugreifen.

Name der Anforderung	Prozedur- und Produktnamen
Zwei Stichproben	vergleicht den Mittelwert der ersten Stichprobe minus den Mittelwert der zweiten Stichprobe mit einer bestimmten Zahl.
Paarweise	vergleicht den Mittelwert der Differenzen in den Beobachtungen mit einer bestimmten Zahl.
Eine Stichprobe	vergleicht den Mittelwert der Stichprobe mit einer bestimmten Zahl.

t-Test: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Klassifizierungsvariablen	Verfügbar für *t*-Tests mit zwei Stichproben. Weisen Sie dieser Rolle eine Klassifizierungsvariable zu, um zwei unabhängige Stichproben zu analysieren. Die Variable muss genau zwei Ausprägungen haben, eine für jede Stichprobe. Sie können dieser Rolle maximal eine Variable zuweisen.
Analysevariablen	Verfügbar für *t*-Test mit einer und zwei Stichproben. *t*-Tests werden für jede Variable in dieser Liste durchgeführt. Sie müssen dieser Rolle mindestens eine Variable zuweisen.
Paarweise Variablen	Verfügbar für paarweise *t*-Tests. Um einen paarweisen *t*-Test durchzuführen, weisen Sie die beiden zu analysierenden Variablen dieser Rolle zu. Da paarweise *t*-Tests durch Subtrahieren jedes Werts der zweiten Variablen vom entsprechenden Wert der ersten durchgeführt werden, ist die Reihenfolge der Variablen relevant. **Hinweis:** Sie müssen dieser Rolle genau zwei Variablen zuweisen.
Analyse gruppieren nach	Wenn Sie dieser Rolle mindestens eine Variable zuweisen, wird die Tabelle nach der/den ausgewählten Variablen sortiert und für jede Gruppe werden separate t-Tests durchgeführt. **Hinweis:** Sie können *t*-Tests nicht nach einer Variablen gruppieren, die bereits für **Analysevariablen** oder **Paarweise Variablen** ausgewählt ist.
Häufigkeitszählung	Wenn Sie dieser Rolle eine Variable zuweisen, wird angenommen, dass jede Beobachtung in der Tabelle *n* Beobachtungen darstellt, wobei *n* der Wert der Häufigkeitszählung für diese Zeile ist. Sie können dieser Rolle maximal eine Variable zuweisen.
Relative Gewichtung	Wenn Sie dieser Rolle eine Variable zuweisen, werden mit dem Wert der Variablen in jeder Beobachtung gewichtete Statistiken berechnet. Sie können dieser Rolle maximal eine Variable zuweisen. **Hinweis:** Wenn die Eingabedatei für diese Anwendungsroutine nur beschreibende Statistiken enthält, ist die Rolle **Relative Gewichtung** nicht verfügbar. Die Gewichtungsoption ist nicht wirksam, wenn die Eingabedatei beschreibende Statistiken enthält.

t-Test: Festlegen der Analyseoptionen

Klicken Sie im Auswahlfenster auf **Analyse**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Nullhypothese angeben	Standardmäßig hat die Nullhypothese den Wert 0. Um einen neuen Nullhypothesewert anzugeben, geben Sie den neuen Wert im Feld $\mathbf{H}_0$= ein.
Optionen für das Standardabweichungskonfidenzintervall festlegen	Sie legen Sie Konfidenzintervalloptionen fest: ■ Das Kontrollkästchen **Gleichlastig** ist standardmäßig markiert. Das gleichlastige Konfidenzintervall sorgt für eine gleiche Flächenverteilung an jedem Ende der Chi-Quadrat-Verteilung. Diese Option bewirkt, dass ein gleichlastiges Konfidenzintervall angezeigt wird. ■ Aktivieren Sie das Kontrollkästchen **UMPU**, um ein vom UMPU-Test (Uniformly Most Powerful Unbiased Test) abgeleitetes Konfidenzintervall zu erhalten. **Hinweis:** Um kein Konfidenzintervall anzuzeigen, deaktivieren Sie die Kontrollkästchen **Gleichlastig** und **UMPU**.

t Test: Erzeugen von Plots

Klicken Sie im Auswahlfenster auf **Plots**, um auf diese Optionen zuzugreifen.

Plot-Name	Beschreibung
Zusammenfassungs-Plot	erzeugt einen Zusammenfassungs-Plot. Der Inhalt des Zusammenfassungs-Plots hängt vom Typ des *t*-Tests ab, den Sie ausführen möchten. ■ Bei einem *t*-Test mit einer Stichprobe enthält der Zusammenfassungs-Plot ein Histogramm mit überlagerten Normal- und Kerndichten, einen Box-Plot und ein Konfidenzintervallband.

Plot-Name	Beschreibung
	■ Bei einem *t*-Test mit zwei Stichproben enthält der Zusammenfassungs-Plot Vergleichshistogramme mit überlagerten Dichten und Box-Plots. ■ Bei einem paarweisen *t*-Test enthält der Zusammenfassungs-Plot ein Histogramm, Dichten, einen Box-Plot und ein Konfidenzintervall der Differenz oder der Ratio.
Histogramm	erzeugt ein Histogramm bzw. Vergleichshistogramme mit überlagerten Normal- und Kerndichten. Bei *t*-Tests mit einer Stichprobe und paarweiser *t*-Tests basieren die Histogramme und Dichten auf dem Testkriterium (das die Mittelwertdifferenz oder die Ratio für eine paarweise Anordnung ist). Bei *t*-Tests mit zwei Stichproben werden Vergleichshistogramme (eines für jede Klasse) gezeigt.
Box-Plot	generiert einen Box-Plot oder vergleichende Box-Plots. Die Box wird aus dem 25. Perzentil (unteres Quartil) bis zum 75. Perzentil (oberes Quartil) bezogen. Die vertikale Linie in der Box zeigt den Standort des Medians. Bei paarweisen *t*-Tests und t-Tests mit einer Stichprobe wird ein Konfidenzintervall für den Mittelwert als Band im Hintergrund anzeigt. Bei *t*-Tests mit zwei Stichproben werden vergleichende Box-Plots (eines für jede Klasse) angezeigt.
Konfidenzintervall-Plot	generiert Plots des Konfidenzintervalls für die Mittelwerte.
Normaler Quantil-Quantil-(Q-Q)-Plot	erzeugt einen normalen Quantil-Quantil-Plot (Q-Q-Plot). Bei *t*-Tests mit zwei Stichproben werden getrennte Plots für jede Klasse in einem einzelnen Panel angezeigt.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

89

Tabellenanalyse

Allgemeines über die Anwendungsroutine 'Tabellenanalyse' ... **695**

Tabellenanalyse: Zuweisen von Variablen zu Analyserollen ... **696**

Tabellenanalyse: Definieren von Tabellen ... **697**

Tabellenanalyse: Auswählen von Zellenstatistiken ... **698**

Tabellenanalyse: Angeben von Tabellenstatistiken ... **699**
- Durchführen von Assoziationstests ... 699
- Durchführen von Übereinstimmungstests ... 701
- Durchführen geordneter Differenzentests ... 701
- Durchführen von Tests auf Trend ... 702
- Festlegen von Berechnungsoptionen ... 702

Tabellenanalyse: Festlegen von Ergebnisoptionen ... **704**
- Speichern von Zellenstatistiken ... 704
- Speichern von Tabellenstatistiken ... 705

Festlegen von Titeln und Fußnoten ... **706**

Anzeigen von Eigenschaften ... **706**

Allgemeines über die Anwendungsroutine 'Tabellenanalyse'

Mithilfe der Anwendungsroutine **Tabellenanalyse** können Sie aus Ihren Daten Kreuztabellen, auch Kontingenztabellen genannt, generieren.

Angenommen, Sie möchten die Wahrscheinlichkeit einer Erkrankung der Herzkranzgefäße für zwei verschiedene Ernährungstypen untersuchen. In einer SAS-Datei liegen Daten einer Fallkontrollstudie zu fettreicher Ernährung und dem Risiko der Erkrankung der Herzkranzgefäße vor. Die Daten werden als Zellenanzahl aufgezeichnet, wobei die Variablenanzahl die Häufigkeiten für die verschiedenen Kombinationen aus Ernährungsweise und Erkrankungsrisiko enthält. Mithilfe der Anwendungsroutine **Tabellenanalyse** können Sie eine 2x2-Kontingenztabelle

erstellen, die die Häufigkeitswerte für die einzelnen Kombinationen aus Ernährungsweise (fettreich und fettarm) und Risiko (hohes Risiko und geringes Risiko) anzeigt.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	FREQ
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	Ohne

Tabellenanalyse: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Häufigkeitszählung	gibt die Variable an, die die Häufigkeit für jede Zeile in der Tabelle enthält. Für jede der Rollen wird angenommen, dass sie *n* Beobachtungen repräsentiert, wobei *n* dem Wert der Häufigkeitszählung für diese Zeile entspricht.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Sie können Analysen nicht nach einer Variable gruppieren, die Sie bereits anderen Rollen zugewiesen haben.

Optionsname	Beschreibung
Tabellenvariablen	verwendet diese Variablen, um zweifache bis n-fache Häufigkeits- und Kreuztabellen zu erstellen. Es werden Statistiken für die Tabellenvariable(n) berechnet. Sie müssen dieser Rolle mindestens zwei Variablen zuweisen.

Tabellenanalyse: Definieren von Tabellen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

So generieren Sie eine neue Tabelle:

1. Wählen Sie im Feld **In Tabelle zulässige Variablen** die Variable aus, die Sie der Tabelle hinzufügen möchten. Die in dieser Liste enthaltenen Variablen wurden der Rolle **Tabellenvariablen** zugewiesen.
2. Ziehen Sie die ausgewählten Variablen auf das Feld **Vorschau**. Die Werte der ersten Variablen, die Sie hinüberziehen, bilden die Variablen der Kreuztabelle. Die Werte der zweiten Variablen, die Sie hinüberziehen, bilden die Zeilen der Kreuztabelle.

   Wenn Sie mehr als zwei Variablen in das Feld **Vorschau** ziehen, werden mehrere Kreuztabellen erzeugt, eine für jede Kombination der Werte für die dritte (und nachfolgende) Variable(n), die Sie hinüberziehen.

   Hinweis: Nachdem Sie der Tabelle eine Variable hinzugefügt haben, wird diese im Feld **In Tabelle zulässige Variablen** gelöscht.

Um eine Variable aus der Tabelle zu entfernen, wählen Sie die Variable im Feld **Vorschau** aus und ziehen Sie sie auf das Feld **In Tabelle zulässige Variablen**.

Um eine neue Tabelle zu beginnen, wählen Sie in der Liste der definierten Tabellen **< auswählen, um neue Tabelle zu definieren >** aus.

Tabellenanalyse: Auswählen von Zellenstatistiken

Klicken Sie im Auswahlfenster auf **Zellenstatistiken**, um auf diese Optionen zuzugreifen.

Die ausgewählten Statistiken werden in jeder Zelle der Kreuztabelle angezeigt.

Optionsname	Beschreibung
Kumulative Spaltenprozentwerte	Zeigt den kumulativen Spaltenprozentwert in jeder Zelle an.
Zeilenprozentwerte	Zeigt den Zeilenprozentwert für jede Zelle an.
Spaltenprozentwerte	Zeigt den Spaltenprozentwert für jede Zelle an. Standardmäßig werden diese Statistiken in die Ausgabe eingeschlossen.
Zellenhäufigkeiten	Zeigt die Häufigkeitszählung für jede Zelle an. Standardmäßig werden diese Statistiken in die Ausgabe eingeschlossen.
Zellenprozentwerte	Zeigt den Prozentwert, Zeilenprozentwert und Spaltenprozentwert in Kreuztabellen oder Prozentwerte und kumulative Prozentwerte in einfachen Häufigkeitstabellen an.
Häufigkeiten fehlender Werte	Zeigt die Häufigkeiten für die fehlenden Werte an.
Zellenverteilung zu Pearsonschem Chi-Quadrat	Zeigt die Zellenverteilung für die gesamte Pearson Chi-Square-Statistik an.
Abweichung der Zellenhäufigkeit vom erwarteten Wert	Zeigt die Abweichung der Zellenhäufigkeit vom erwarteten Wert für jede Zelle an.
Erwartete Zellenhäufigkeit	Zeigt die erwartete Zellenhäufigkeit für jede Zelle an.

Optionsname	Beschreibung
Prozentwert der Gesamthäufigkeit	Zeigt den Prozentwert der Gesamthäufigkeit bei *n*-fachen Tabellen an, wenn *n*> gleich 2 ist.
Prozentwerte in Datei einschließen	zeigt die Prozentwerte für die Spalten und Zeilen in der Ausgabedatei an.

Tabellenanalyse: Angeben von Tabellenstatistiken

Durchführen von Assoziationstests

Klicken Sie im Auswahlfenster unter der Überschrift 'Tabellenstatistiken' auf **Assoziation**, um auf diese Optionen zuzugreifen.

Hinweis: Wenn Sie **Exakte p-Werte** auswählen, können die Berechnungen des exakten *p*-Wertes sehr speicher- und zeitintensiv sein. Geben Sie mit der Option **Max. Berechnungszeit** eine Zeitbegrenzung für jede Berechnung des exakten *p*-Wertes an.

Optionsname	Beschreibung
Assoziationstests	Folgende Optionen können angegeben werden: ■ **Chi-Quadrat-Test** berechnet die Chi-Quadrat-Tests auf Homogenität oder Unabhängigkeit und Assoziationsmaße basierend auf Chi-Quadrat. Die Tests umfassen Pearson-Chi-Quadrat, Likelihood-Ratio-Chi-Quadrat und Mantel-Haenszel-Chi-Quadrat. Für 2×2-Tabellen umfasst dieser Test Fishers exakten Test und das Chi-Quadrat mit angepasster Kontinuität. ■ **Exakte p-Werte** berechnet die exakten *p*-Werte für die folgenden Statistiken: Chi-Quadrat-Test auf Güte der Anpassung für einfache Tabellen, Pearson-Chi-Quadrat-, Likelihood-Ratio-Chi-Quadrat- und Mantel-Haenszel-Chi-Quadrat-Tests für zweifache Tabellen. ■ **Fishers exakter Test** berechnet Fishers exakten Test für Tabellen, die größer sind als 2×2. Dieser Test wird auch Freeman-Halton-Test genannt.

Optionsname	Beschreibung
Assoziationsmaße	Folgende Optionen können angegeben werden: ■ **Maße** - Berechnet mehrere Assoziationsmaße und deren asymptotische Standardfehler (ASE). Die Maße umfassen Gamma, Kendalls Tau-*b*, Stuarts Tau-*c*, Somers *D (C|R)*, Somers *D (R|C)*, den Pearsonschen und Spearman-Korrelationskoeffizienten, Lambda (symmetrisch und asymmetrisch) und Unsicherheitskoeffizienten (symmetrisch und asymmetrisch). Für 2×2-Tabellen bietet dieses Maß auch das Odds Ratio, das relative Risiko in Spalte 1, das relative Risiko in Spalte 2 und die entsprechenden Konfidenzgrenzen. ■ **Exakter p-Wert und Konfidenzgrenzen für Odds Ratio** berechnet die exakten *p*-Werte für die Pearson-Korrelation und die Spearman-Korrelation sowie die Odds-Ratio-Konfidenzgrenzen für 2×2-Tabellen. ■ **Test, ob das Maß gleich null ist** berechnet asymptotische Tests für Gamma, Kendalls Tau-*b*, Stuarts Tau-*c*, Somers *D (C|R)*, Somers *D (R|C)*, die Pearson-Korrelation und die Spearman-Korrelation. ■ **Risikodifferenzen für 2x2-Tabellen** stellt Risikoschätzer (oder Binomialverhältnisse) und Risikodifferenzen für 2×2-Tabellen zur Verfügung. Diese Analyse kann angebracht sein, wenn Sie das Verhältnis einiger Merkmale für zwei Gruppen vergleichen, wobei Zeile 1 und Zeile 2 den beiden Gruppen entsprechen, während die Spalten den beiden möglichen Merkmalen oder Ergebnissen entsprechen. ■ **Relatives Risiko für 2x2-Tabellen** berechnet die Maße für das relative Risiko und deren Konfidenzgrenzen für 2×2-Tabellen. Diese Maße schließen das Odds-Ratio und die relativen Risiken der Spalte 1 und 2 ein.
Cochran-Mantel-Haenszel-Statistiken	Um alle Cochran-Mantel-Haenszel-Statistiken zu berechnen, aktivieren Sie das Kontrollkästchen **CMH-Statistiken**. Diese Statistiken schließen die CMH-Korrelationsstatistik, die Zeilenmittelwert-Scorewerte (ANOVA) und die angepassten relativen Risiken und Odds Ratios ein. Der Standardtyp für den Scorewert ist `Tabelle`. Um den Scorewert-Typ zu ändern, verwenden Sie die Option **Scorewert-Typ**.

Durchführen von Übereinstimmungstests

Klicken Sie im Auswahlfenster unter der Überschrift 'Tabellenstatistiken' auf **Übereinstimmung**, um auf diese Optionen zuzugreifen.

Sie können folgende Tests und Maße der Übereinstimmung für *n*x*n*-Tabellen wählen.

Optionsname	Beschreibung
Maße	Berechnet Tests und Maße der Klassifizierungsübereinstimmung für Quadrattabellen. Diese Option stellt den McNemars Test für 2×2-Tabellen und den Bowkers Test auf Symmetrie für Tabellen mit mehr als zwei Abhängigkeitskategorien zur Verfügung. Die Option erstellt auch den einfachen Kappa-Koeffizienten, den gewichteten Kappa-Koeffizienten, die asymptotischen Standardfehler für die einfachen und gewichteten Kappas und die entsprechenden Konfidenzgrenzen. Wenn multiple Schichten und zwei Abhängigkeitskategorien vorliegen, berechnet diese Option auch Cochrans Q-Test.
Exakte p-Werte	Berechnet die exakten *p*-Werte für den McNemars Test für 2×2-Tabellen, den einfachen Kappa-Koeffizienten und den gewichteten Kappa-Koeffizienten. **Hinweis:** Wenn Sie **Exakte p-Werte** auswählen, können die Berechnungen des exakten *p*-Wertes sehr speicher- und zeitintensiv sein. Geben Sie mit der Option **Max. Berechnungszeit** eine Zeitbegrenzung für jede Berechnung des exakten *p*-Wertes an.
Test, ob das Maß gleich null ist	Berechnet asymptotische Tests für den einfachen Kappa-Koeffizienten und den gewichteten Kappa-Koeffizienten.

Durchführen geordneter Differenzentests

Klicken Sie im Auswahlfenster unter der Überschrift 'Tabellenstatistiken' auf **Geordnete Differenzen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Jonckheere-Terpstra-Test	Berechnet den Jonckheere-Terpstra-Test, der ein nicht parametrischer Test für geordnete Differenzen zwischen Klassen

Optionsname	Beschreibung
	ist. Er überprüft die Nullhypothese, dass die abhängige Variable zwischen den Klassen nicht abweicht.
Exakte p-Werte	Berechnet den exakten *p*-Wert für den Jonckheere-Terpstra-Test. **Hinweis:** Wenn Sie **Exakte p-Werte** auswählen, können die Berechnungen des exakten *p*-Wertes sehr speicher- und zeitintensiv sein. Geben Sie mit der Option **Max. Berechnungszeit** eine Zeitbegrenzung für jede Berechnung des exakten *p*-Wertes an.

Durchführen von Tests auf Trend

Klicken Sie im Auswahlfenster unter der Überschrift 'Tabellenstatistiken' auf **Test auf Trend**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Cochran-Armitage-Test	Berechnet den Cochran-Armitage-Test auf Trend, der in Binomialverhältnissen übergreifend über Ausprägungen eines einzelnen Faktors oder einer einzelnen Kovariate auf Trends testet. Dieser Test eignet sich für eine Kontingenztabelle, in der eine Variable zwei Ausprägungen aufweist, während die andere Variable eine erklärende Variable mit geordneten Abstufungen darstellt. Die Variable mit zwei Ausprägungen stellt die abhängige (Response) Variable und die andere Variable stellt die erklärende Variable mit geordneten Abstufungen dar.
Exakte p-Werte	Berechnet den exakten *p*-Wert für den Cochran-Armitage-Test auf Trends. **Hinweis:** Wenn Sie **Exakte p-Werte** auswählen, können die Berechnungen des exakten *p*-Wertes sehr speicher- und zeitintensiv sein. Geben Sie mit der Option **Max. Berechnungszeit** eine Zeitbegrenzung für jede Berechnung des exakten *p*-Wertes an.

Festlegen von Berechnungsoptionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Tabellenstatistiken' auf **Berechnungsoptionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Scorewert-Typ	Wenn Sie Assoziationstests, Übereinstimmungstests oder Tests auf Trend durchführen, wählen Sie einen Scorewerttyp aus. Der Scorewerttyp gibt den Zeilentyp und den Typ der Spaltenscorewerte an, der mit der Mantel-Haenszel-Chi-Quadrat-Statistik, den Cochran-Mantel-Haenszel-Statistiken, der Pearsonschen Korrelation, dem Cochran-Armitage-Test auf Trends und den gewichteten Kappa-Koeffizienten verwendet wird. Folgende Optionen können ausgewählt werden: ■ **Tabelle** - Für numerische Variablen sind Tabellen-Scorewerte die Werte der Zeilen- und Spaltenausprägungen. Wenn die Zeilen- oder Spaltenvariablen formatiert sind, ist der Tabellen-Scorewert der interne numerische Wert, der dieser Ausprägung entspricht. Wenn zwei oder mehrere numerische Werte in derselben formatierten Ausprägung klassifiziert sind, ist der interne numerische Wert für diese Ausprägung der kleinste dieser Werte. Für alphanumerische Variablen werden Tabellen-Scorewerte als die Zeilen- und Spaltennummern definiert (d. h. 1 für die erste Zeile, 2 für die zweite Zeile usw.). ■ **Rang** - Rang-Scorewerte werden auf Basis von Zeilen- und Spalten-Scorewerten berechnet. ■ **RIDIT** - RIDIT-Scorewerte werden aus Rang-Scorewerten abgeleitet. ■ **Modifizierter RIDIT** - Modifizierte RIDIT-Scorewerte (MODRIDIT) zeigen die erwarteten Werte von Ordnungsstatistiken der einheitlichen Verteilung auf (0,1). Modifizierte RIDIT-Scorewerte werden aus Rang-Scorewerten abgeleitet.
Exakte Berechnungen	Wenn Sie in einem Ihrer Tests exakte *p*-Werte berechnen, beachten Sie, dass die Berechnung exakter *p*-Werte viel Arbeitsspeicher und Verarbeitungszeit in Anspruch nehmen kann. Aktivieren Sie im Bereich **Exakte Berechnungen** das Kontrollkästchen **Max. Berechnungszeit** und legen Sie eine Zeitgrenze (in Sekunden) für die Berechnung jedes *p*-Werts für jede Tabelle fest. Der Standardwert beträgt `900` Sekunden. Zur Berechnung der Monte-Carlo-Schätzer für die exakten *p*-Werte anstelle der direkten Berechnung der exakten *p*-Werte, aktivieren Sie das Kontrollkästchen **Monte-Carlo-Schätzung**. Die Monte-Carlo-Schätzung kann bei großen Problemen hilfreich sein, bei denen die exakte Berechnung zeit- und speicherintensiv ist, asymptotische Annäherungen jedoch eventuell nicht ausreichend sind.
Fehlende Werte in	Sie können auch fehlende Werte bei den Berechnungen von Prozentwerten und anderen Statistiken einschließen.

Optionsname	Beschreibung
Berechnungen einbeziehen	
Konfidenzniveau	Sie können für die Berechnungen ein Konfidenzniveau angeben. Das Konfidenzniveau ist standardmäßig eine Konfidenzgrenze von 95%.
Werte sortieren nach	Sie können die Reihenfolge auswählen, in der die Werte der Variablen in den Kreuztabellen aufgelistet werden. **Werte anordnen nach** kann zum Festlegen folgender Optionen verwendet werden: ■ **Dateireihenfolge** sortiert die Werte entsprechend ihres Auftretens in der Datei ■ **Formatierte Werte** sortiert die Werte in aufsteigender Reihenfolge der formatierten Werte ■ **Absteigende Häufigkeiten** ordnet die Werte so, dass die Ebenen mit den höchsten Häufigkeitszählungen zuerst angezeigt werden ■ **Unformatierte Werte** sortiert die Werte nach den unformatierten Werten. Hierbei handelt es sich um die Standardeinstellung.

Tabellenanalyse: Festlegen von Ergebnisoptionen

Speichern von Zellenstatistiken

Klicken Sie im Auswahlfenster unter der Überschrift 'Ergebnisse' auf **Ergebnisse der Zellenstatistiken**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Speichern von Ausgabedaten	Aktivieren Sie das gewünschte bzw. die gewünschten Kontrollkästchen für die entsprechende Kreuztabelle. Die aufgeführte Tabelle entspricht der Variable, die der Rolle **Tabellenvariablen** zugewiesen ist. Geben Sie im Feld **Zellenstatistikendatei** die Datei an, in der die Zellenstatistiken für die ausgewählte Tabelle gespeichert werden sollen. Der Name der Datei ändert sich für jede ausgewählte Tabelle.

Optionsname	Beschreibung
	Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Anzeige für gesamte Ergebnisausgabe unterdrücken	Bewirkt, dass keine Ergebnisse angezeigt werden.

Speichern von Tabellenstatistiken

Klicken Sie im Auswahlfenster unter der Überschrift 'Ergebnisse' auf **Ergebnisse der Tabellenstatistiken**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Statistiken für die folgende Tabelle speichern	speichert die Tabellenstatistiken für eine Kreuztabelle. Geben Sie im Feld **Tabellenstatistikendatei** die Datei an, in der die Tabellenstatistiken für die ausgewählte Tabelle gespeichert werden sollen. Der Name der Datei ändert sich für jede ausgewählte Tabelle. **Hinweis:** SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste. Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.
Tabellen als Listen anzeigen	zeigt die gesamte Mehrfachtabelle in einer Tabelle an anstatt eine separate Zweifachtabelle für jede Schicht anzuzeigen. **Hinweis:** Die Option **Tabellen als Listen anzeigen** wird ignoriert, wenn Sie eine Option auswählen, mit der eine Tabellenstatistik erzeugt wird.
Anzeige für gesamte Ergebnisausga	zeigt keine Ergebnisse an.

Optionsname	Beschreibung
be unterdrücken	

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1. Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.
2. Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.
3. Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten

- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

90

Kacheldiagramm

Allgemeines über die Anwendungsroutine 'Kacheldiagramm' ... **709**

Kacheldiagramm: Zuweisen von Variablen zu Analyserollen ... **710**

Kacheldiagramm: Festlegen von Darstellungsoptionen ... **711**
Festlegen der Darstellung von Kacheldiagrammen ... 711
Festlegen der Legendenoptionen ... 713

Kacheldiagramm: Festlegen erweiterter Optionen ... **714**

Festlegen von Titeln und Fußnoten ... **715**

Allgemeines über die Anwendungsroutine 'Kacheldiagramm'

Mit der Anwendungsroutine **Kacheldiagramm** können Sie ein Kacheldiagramm erstellen, sodass Sie umfangreiche hierarchische Daten auf einer begrenzten Fläche anzeigen können. Jede Kategoriekombination wird durch eine rechteckige Kachel dargestellt, deren Größe und Farbe durch abhängige Variablen bestimmt wird. Diese Kacheln werden in einer hierarchischen Struktur angeordnet.

Bei einem Kacheldiagramm handelt es sich um eine rechteckige Tree Map. Da das Kacheldiagramm in einer Baumstruktur dargestellt wird, ist es besonders in einer interaktiven Umgebung, wie dem Internet, hilfreich. Dies sind die Vorteile von Kacheldiagrammen in einer interaktiven Umgebung:

- Um kleine Abschnitte des Originaldiagramms ausführlicher anzuzeigen, können Sie die Drilldown-Funktion (oder auch Zoomfunktion) verwenden.
- Wenn Sie den Mauszeiger über eine Kachel bewegen, werden in Datatips die Messwerte für eine bestimmte Kachel angezeigt.
- Sie können die Kacheln in einer bestimmten Ebene mit einem Bericht oder einer Webseite verknüpfen. Wenn Sie in dieser Ebene auf eine Kachel klicken, wird der Bericht oder die Webseite, der bzw. die mit dieser Kachel verknüpft ist, angezeigt.

Nur ActiveX, Java, ActiveX Image und Java Image sind für die Anwendungsroutine **Kacheldiagramm** gültige Datenformate. Besitzt das aktuelle Grafikformat keines dieser Formate, verwendet die Anwendungsroutine Kacheldiagramm das Grafikformat ActiveX.

Bei HTML-Ergebnissen sind einige Funktionen nur für bestimmte Grafikformate verfügbar.

- Die Drilldown-Funktion ist verfügbar, wenn Sie ActiveX oder Java als Format für die Grafikausgabe gewählt haben.
- Der Link von Ihrem Kacheldiagramm zu einer Webseite ist verfügbar, wenn Sie Java als Format für die Grafikausgabe gewählt haben.

Im Dialogfeld **Optionen** können Sie das Grafikformat und das Format für Ihre Ergebnisse festlegen.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	GTILE
Erforderliche SAS-Produkte	Base SAS, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	Ohne

Kacheldiagramm: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Rollen für Anwendungsroutine**, um auf diese Optionen zuzugreifen.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Hinweis: Die jeweils verfügbaren Rollen hängen vom ausgewählten Diagrammtyp ab.

Optionsname	Beschreibung
Klassifizierungsvariablen	gibt die hierarchischen Kategorie-Variablen (bzw. ID-Variablen) an. Eine Klassifizierungsvariable ist erforderlich und Sie weisen dieser Rolle bis zu sechs Variable zu.

Optionsname	Beschreibung
	Wenn Sie dieser Rolle mehrere Variable zuweisen, wird die Hierarchie anhand der Reihenfolge der Variablen bestimmt. Wenn beispielsweise die Reihenfolge der Variablen in der Rolle Klassifizierungsvariable Land, Region und Verwaltungsbezirk ist, lautet die entsprechende Hierarchie Land > Region > Verwaltungsbezirk.
Variable Farbanalyse	gibt die Variablenwerte an, die die Farben der Kacheln bestimmen. Der kleinsten abhängigen Farbvariablen wird der Minimalwert und der größten abhängigen Variablen der Maximalwert der Farbliste zugewiesen. Andere abhängige Farbvariablen werden linear in einem Gradienten zwischen den Farben in der Liste zugewiesen. Sie können dieser Rolle nur eine Variable zuweisen. **Hinweis:** Wenn keine Variable als Variable Farbanalyse zugewiesen ist, dann wird die Variable als Variable Farbanalyse verwendet, die Sie der Rolle **Variable Größenanalyse** zugewiesen haben.
Variable Größenanalyse	gibt die Variablenwerte an, die die relative Größe der Kacheln bestimmen. Die Gesamtgröße der Kacheln bedeckt die gesamte Kachelfläche. Bei fehlenden, negativen und Nullwerten werden die entsprechenden Kacheln ignoriert. Eine Variable Größenanalyse ist erforderlich. Sie können dieser Rolle nur eine Variable zuweisen.
Diagramme gruppieren nach	Die Werte der Spalte, die Sie dieser Rolle zuweisen, bestimmen die Anzahl der Diagramme, die erstellt werden. Für jeden eindeutigen Wert wird eine separate Grafik erzeugt. Sie können dieser Rolle mehr als eine Variable zuweisen.

Kacheldiagramm: Festlegen von Darstellungsoptionen

Festlegen der Darstellung von Kacheldiagrammen

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Kacheldiagramm**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Kachellayout	Legen Sie das Kachellayout im Diagramm fest. Folgende Optionen können ausgewählt werden: ■ **Fluss** - Unterteilt das Kachellayout in mehrere ausgerichtete Zeilen. Im Diagramm werden die angeordneten Kacheln von links nach rechts und dann von oben nach unten gelesen. ■ **Quadrat** - Ist die klassische rechteckige Tree Map. Normalerweise werden die Kacheln von der größten bis zur kleinsten von unten links bis oben rechts im Diagramm dargestellt. Die Sortierung für Datenvariable wird ignoriert und die Form der Kacheln ist so quadratisch wie möglich. ■ **Umschalten** - Unterteilt die Richtung für jede Hierarchieebene, die sich aus den Variablen ergeben, die Sie der Rolle **Klassifizierungsvariable** zugewiesen haben. Die höchste Hierarchieebene wird im verfügbaren Platz in vertikale Segmente aufgeteilt und dann in sortierter Reihenfolge entlang einer einzigen Achse angezeigt. In den unteren Ebenen wird die Aufteilungsrichtung geändert (oder gewechselt), sodass die zweite Ebene horizontal aufgeteilt wird.
Klassifizierungsausprägung angeben	geben sie die Anzahl der in den Ergebnissen angezeigten Ausprägungen an. Wenn beispielsweise die Reihenfolge der Variablen in der Rolle **Klassifizierungsvariable** Land, Region und Verwaltungsbezirk ist, lautet die entsprechende Hierarchie Land > Region > Verwaltungsbezirk. Die Anzahl der Ausprägungen, die in dem Kacheldiagramm angezeigt werden, hängt vom Wert der Option **Klassifizierungsausprägung angeben** ab. ■ Wenn diese Option auf 1 gesetzt wird, dann wird in den Ergebnissen eine Hierarchieebene angezeigt. Im Kacheldiagramm werden die Kacheln für jedes einzelne Land angezeigt (wie z. B. Kanada, Deutschland und die USA). Wenn Sie einen Drilldown zur Kachel "Kanada" vornehmen, werden im Kacheldiagramm die Kacheln für die beiden Regionen (Westen und Osten) in Kanada angezeigt. ■ Wenn diese Option auf 2 gesetzt wird, dann werden in den Ergebnissen zwei Hierarchieebenen angezeigt. Im Kacheldiagramm werden die Kacheln für jedes Land und jede Region angezeigt (wie z. B. Westkanada, Ostkanada, Westdeutschland, Ostdeutschland, Westen der USA und Osten der USA). Wenn Sie einen Drilldown zur Kachel "Kanada" vornehmen, werden die Kacheln für alle Verwaltungsbezirke in den Regionen Westkanada und Ostkanada angezeigt. Standardmäßig werden alle Hierarchieebenen angezeigt.
Farben angeben	passt den Farbbereich für das Kacheldiagramm an. In einem Farbbereich erstrecken sich die Farben über verschiedene

Optionsname	Beschreibung
	Datenwerte. Sie müssen mindestens zwei Farben angeben. Diese Farben sind Start- und Endfarbe für den Farbbereich. Die Farben im Farbverlauf werden standardmäßig gleichmäßig verlaufend verteilt. Durch die Verwendung von Farbpunkten können Sie festlegen, wie die Farben auf die Daten verteilt werden. Klicken Sie auf **Bearbeiten**, um Farbpunkte festzulegen. Im Dialogfeld **Farbpunkte bearbeiten** können die Werte von 0 bis 1 reichen; sie müssen in aufsteigender Reihenfolge angegeben werden. Beispiel: Wenn Sie drei Farben für den Farbbereich ausgewählt haben, könnten Sie 0, 0,75 und 1 als Farbpunkte angeben. Es ist möglich, denselben Wert für verschiedene Farbpunkte festzulegen. Beispiel: Wenn Sie fünf Farben für den Farbbereich ausgewählt haben, könnten Sie 0, 0,2, 0,4, 0,4 und 0,6 als Farbpunkte angeben. Es müssen immer genauso viele Farbpunkte angegeben werden wie Farben für den Farbbereich ausgewählt werden.

Festlegen der Legendenoptionen

Um die Start- und Endpunkte für die Legende festzulegen, aktivieren Sie das Kontrollkästchen **Legendenbereich festlegen**. Angenommen, Ihre Daten weisen einen Datenbereich von 0 bis 100.000 auf. Im Kacheldiagramm sollen die Werte zwischen 22.000 und 26.000 farbig hervorgehoben werden. In diesem Fall würde man 22.000 als Minimalwert (Startwert) und 26.000 als Maximalwert (Endwert) festlegen. Die Kacheln mit Werten kleiner 22.000 bzw. größer 26.000 werden grau gekennzeichnet. Die Kacheln mit Werten zwischen 22.000 und 26.000 erhalten den von Ihnen festgelegten Farbbereich.

Standardmäßig verwendet die Farbzuordnung in der Legende die Farben der Response-Daten. Sie können jedoch festlegen, ob die Farbzuordnung auf beiden Seiten symmetrisch zu einem festen Wert ist, indem Sie das Kontrollkästchen **Festlegen, dass Farbzuordnung der Legende symmetrisch zu einem festen Wert ist** aktivieren. Mit dieser Option können Sie den Midpoint in der Farbskala festlegen. Standardmäßig ist der Symmetrie- oder Mittelpunkt 0. Sie können einen neuen Symmetriepunkt oder Mittelpunkt angeben.

Kacheldiagramm: Festlegen erweiterter Optionen

Klicken Sie im Auswahlfenster auf **Erweitert**, um auf diese Optionen zuzugreifen.

Hinweis: Sie können nur dann in Ihrem Kacheldiagramm einen Link zu einer Webseite erstellen, wenn Sie in SAS Enterprise Guide Java als Format für die Grafikausgabe und HTML als Format für Ihre Ergebnisse ausgewählt haben. Um das Grafikformat und das Format für die Ergebnisse festzulegen, siehe Fenster **Ergebnisse** ⇨ **Diagramm** im Dialogfenster **Optionen**. Da Java benötigt wird, werden Links im SAS Add-In for Microsoft Office nicht unterstützt.

Sie können aus einem Kacheldiagramm Links auf Webseiten erstellen. Ein Link wird in einem Kontextmenü angezeigt, wenn Sie mit der rechten Maustaste in einer Ebene für die ausgewählte Linkvariable auf eine Kachel klicken. (Die Linkvariablen sind die Variablen, die Sie zur Rolle **Klassifizierungsvariablen** zugewiesen haben.) In welcher Ebene diese Links angezeigt werden, hängt davon ab, ob Sie eine Klassifizierungsausprägung angegeben haben.

- Wenn Sie keine Klassifizierungsausprägung angeben, sind Links aus allen Ebenen des Kacheldiagramms verfügbar. Wenn die Hierarchie für dieses Kacheldiagramm beispielsweise Land > Region > Verwaltungsbezirk ist, sind die Links verfügbar, wenn Sie mit der rechten Maustaste auf die Kacheln in den Ebenen Land, Region und Verwaltungsbezirk klicken.
- Wenn Sie eine Klassifizierungsausprägung angeben, sind nur Links aus den untersten Ebenen des Kacheldiagramms verfügbar. Die Hierarchie für das Kacheldiagramm besteht beispielsweise aus drei Ebenen: Land, Region und Verwaltungsbezirk. Wenn die Klassifizierungsausprägung 2 ist, sind die Links in den beiden unteren Ebenen Region und Verwaltungsbezirk verfügbar. Wenn die Klassifizierungsausprägung 1 ist, sind nur die Links der untersten Ebene Verwaltungsbezirk verfügbar.

In der Anwendungsroutine **Kacheldiagramm** werden die Linkvariablen im Feld **Verfügbare Variablen** aufgelistet. Die Anwendungsroutine weist jeder Linkvariablen automatisch eine Ersetzungsnummer zu. Diese Nummer wird im Feld **Verfügbare Variablen** neben der mit ihr verknüpften Variablen angezeigt. Sie können diese Ersetzungsnummer in die URL für den Web-Link oder das Etikett einbeziehen, um für jeden Wert einer Linkvariablen einen eindeutigen Link zu erstellen.

In Ihren Daten sind beispielsweise die Werte für Land fr (Frankreich), jp (Japan) und uk (United Kindom) enthalten. Wenn 1 die Ersetzungsnummer ist, die automatisch für die Variable Land zugewiesen wird, und Sie http://www.sales.{&1}.com als URL für den Web-Link angeben, sind die folgenden Links im Kacheldiagramm verfügbar:

- http://www.sales.fr.com

- http://www.sales.jp.com
- http://www.sales.uk.com

So fügen Sie Links zu einem Diagramm hinzu:

1 Geben Sie das Etikett für die Verknüpfung an. Um für jeden Wert der Linkvariablen ein eindeutiges Etikett zu erstellen, können Sie die Ersetzungsnummer der Linkvariablen für das Etikett angeben. Um die Ersetzungsnummer für eine Linkvariable im Etikett hinzuzufügen, klicken Sie auf **`Variable einfügen`**.

   Hier einige Beispiele für die Angabe des Link-Etiketts:

   - **`Staaten-Webseite für {&1}`**
   - **`Wikipedia Webseite für {&2}`**

2 Geben Sie die URL für die Webseite an. Sie können für jede Linkvariable einen eindeutigen Web-Link erstellen, indem Sie die Ersetzungsnummer für die Linkvariable in die URL einfügen. Die Syntax zur Angabe der Linkvariable lautet {&*Ersetzungsnummer*}. Dieser Ersetzungswert kann an einer beliebigen Stelle in der URL für den Web-Link eingefügt werden. Um die Ersetzungsnummer für eine Linkvariable in der URL hinzuzufügen, klicken Sie auf **Variable einfügen**.

   Hier einige Beispiele für die Angabe einer URL für den Web-Link:

   - http://www.state.{&1}.us
   - http://en.wikipedia.org/wiki/{&2}

3 Klicken Sie auf **Hinzufügen**, um diesen Weblink zur Liste mit Weblinks für das aktuelle Kacheldiagramm hinzuzufügen.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

91

Transponieren

Allgemeines über die Anwendungsroutine 'Transponieren' 717
Transponieren: Zuweisen von Variablen zu Analyserollen 718
Transponieren: Festlegen der Analyseoptionen 719
Transponieren: Festlegen von Ergebnisoptionen 720
Anzeigen von Eigenschaften 720

Allgemeines über die Anwendungsroutine 'Transponieren'

Die Anwendungsroutine **Transponieren** wandelt die ausgewählten Spalten einer Eingabedatei in die Zeilen einer Ausgabedatei um. Wenn Sie keine Gruppierungsvariablen verwenden, wird jede ausgewählte Spalte in eine einzelne Zeile umgewandelt. Wenn Sie jedoch Gruppierungsvariablen verwenden, werden die ausgewählten Spalten basierend auf den Werten der Gruppierungsvariablen in Unterspalten aufgeteilt, und jede Unterspalte wird in eine Zeile der Ausgabedatei umgewandelt.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	TRANSPOSE
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

Transponieren: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Variablen transponieren	Jede Spalte, der Sie diese Rolle zuweisen, wird in der Ausgabedatei zu einer oder mehreren Zeilen. Wenn Sie keine Gruppierungsvariablen auswählen, wird die gesamte Spalte in eine einzige Zeile umgewandelt. Wenn Sie eine oder mehrere Gruppierungsvariablen auswählen, werden mit den Gruppierungsvariablen alle Spalten in Unterspalten segmentiert, von denen jede in eine Zeile umgewandelt wird. In diesem Fall gehen aus einer Spalte soviele Zeilen hervor wie von den Gruppierungsvariablen Gruppen definiert werden. Sie müssen mindestens eine Spalte der Rolle **Variablen transponieren** zuweisen. Um eine Gruppierungsvariable auszuwählen, weisen Sie der Rolle **Analyse gruppieren nach** eine Spalte zu.
Variablen kopieren	Jede Spalte, die Sie dieser Rolle zuweisen, wird direkt aus der Eingabedatei in die Ausgabedatei kopiert, ohne transponiert zu werden. Da diese Spalten direkt in die Ausgabedatei kopiert werden, ist die Anzahl der Zeilen in der Ausgabedatei gleich der Anzahl der Zeilen in der Eingabedatei. Die Ausgabedatei wird mit fehlenden Werten aufgefüllt, wenn die Zeilenanzahl in der Eingabedatei nicht der Anzahl der Variablen entspricht, die transponiert werden.
Neue Spaltennamen	Wenn Sie dieser Rolle eine Eingabespalte zuweisen, werden mit den Werten in dieser Spalte die Suffixe der Datenspaltenüberschriften in der Ausgabedatei gebildet. Wenn Sie dieser Rolle keine Eingabespalte zuweisen, werden die Suffixe der Datenspaltenüberschriften in der Ausgabedatei mit

Optionsname	Beschreibung
	ganzen Zahlen gebildet. Um das Präfix für die Datenspaltenüberschriften anzugeben, klicken Sie im Auswahlfenster auf **Optionen** und aktivieren das Kontrollkästchen **Präfix verwenden**. Geben Sie dann im Feld **Präfix für Spaltenname** das Präfix ein. Im Allgemeinen sollte eine Spalte, welche für die Spaltennamen der Ausgabedatei verwendet wird, keine doppelten Werte enthalten, da dies zu doppelten Spaltennamen führen würde. Wenn Sie eine Spalte mit doppelten Werten verwenden möchten, wählen Sie die Spalte nach dem Zuweisen aus und markieren anschließend das Kontrollkästchen **Duplikate zulassen**. In diesem Fall wird nur die letzte Zeile, die einen gegebenen Wert enthält, bei der Transposition verwendet. Wenn Sie eine Spalte mit doppelten Werten verwenden und das Kontrollkästchen **Duplikate zulassen** nicht markieren, kann die Transposition nicht durchgeführt werden.
Analyse gruppieren nach	Mit jeder Variablen, die Sie dieser Rolle zuweisen, werden die zu transponierenden Spalten in Unterspalten segmentiert, die wiederum transponiert werden. Jede Unterspalte, von den Werten der Gruppierungsvariablen definiert, wird zu einer Zeile der Ausgabedatei.
Neue Spaltenetiketten	Diese Rolle wird verfügbar, wenn Sie der Rolle **Neue Spaltennamen** eine Eingabespalte zuweisen. Wenn Sie dieser Rolle eine Eingabespalte zuweisen, werden mit den Werten in dieser Spalte die Etiketten der Datenvariablen in der Ausgabedatei gebildet. Wenn Sie dieser Rolle keine Eingabespalte zuweisen, sind die Etiketten der Datenvariablen in der Ausgabedatei mit den Namen dieser Variablen identisch.

Transponieren: Festlegen der Analyseoptionen

Klicken Sie im Auswahlfenster auf **Optionen**, um auf diese Optionen zuzugreifen.

Jede Zeile der Ausgabedatei enthält den Namen und das Etikett der Variablen in der Einlesedatei. Sie können eine Überschrift für die Ausgabespalte angeben, die diese Variablennamen oder Variablenetiketten enthält. Der Name darf Sonderzeichen und Whitspaces enthalten und darf mit einer Zahl beginnen, darf jedoch nicht länger als 32 Zeichen sein. Die Standardüberschrift für die

Variablennamensspalte lautet **Quelle**. Die Standardüberschrift für die Variablenetikettspalte lautet **Etikett**.

Die Überschriften aller Datenspalten der Ausgabedatei beginnen mit dem gleichen Präfix. Das Präfix darf Sonderzeichen und Whitspaces enthalten und mit einer Zahl beginnen. Es darf jedoch nicht länger als 32 Zeichen sein. Das Standardpräfix ist **Spalte**.

Hinweis: Die Angabe eines Präfixes ist optional, wenn Sie der Rolle **Neue Spaltennamen** eine Variable zuweisen. Wird dieser Rolle keine Variable zugewiesen, ist ein Präfix erforderlich und das Kontrollkästchen **Präfix verwenden** ist nicht verfügbar.

Transponieren: Festlegen von Ergebnisoptionen

Klicken Sie im Auswahlfenster auf **Ergebnisse**, um auf diese Optionen zuzugreifen.

Sie können der Ausgabedatei einen anderen Namen zuweisen.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

92

u-Regelkarte

Allgemeines über die 'u-Regelkarte' 723
u-Regelkarte: Zuweisen von Variablen zu Analyserollen 724
Regelkarten: Festlegen von Regelgrenzen 726
Regelkarte: Auswählen der durchzuführenden Tests 727
Regelkarten: Festlegen von Optionen für die Plotdarstellung 728
Regelkarten: Speichern von Analyseergebnissen 729
Festlegen von Titeln und Fußnoten 730
Anzeigen von Eigenschaften 730

Allgemeines über die 'u-Regelkarte'

Die Shewhart-Regelkarte ist ein Grafik- und Analysetool, mit dem bestimmt werden kann, ob ein Prozess innerhalb der statistischen Kontrolle liegt. Die Anwendungsroutine **u-Regelkarte** erstellt u-Regelkarten für die Anzahl der Nichtkonformitäten (Defekte) pro Inspektionseinheit in den Untergruppenstichproben mit einer willkürlichen Einheitenzahl.

Sie können diese Anwendungsroutine zum Ermitteln der Anzahl der Fehler pro Untersuchungseinheit verwenden, die aus einem Fertigungsprozess resultieren. Angenommen, ein Textilunternehmen verwendet eine u-Regelkarte zur Überwachung der Fehler pro Quadratmeter Stoff. Der Stoff wird bei der Untersuchung auf Rollen gewickelt. Jeder Stoff ist einen Meter breit und 30 Meter lang. Die Fehlerzählungen für 20 Rollen werden in einer SAS-Datei gespeichert, die zum Erstellen der u-Regelkarte verwendet wird. Jeder Punkt auf der u-Regelkarte stellt die Anzahl der Nichtkonformitäten pro Untersuchungseinheit in einer bestimmten Untergruppe dar. Beispielsweise wird für die erste Untergruppe der Wert 12/30=0,4 dargestellt (da 12 Fehler auf der ersten Rolle erkannt werden und sich auf dieser 30 Quadratmeter Stoff befinden). Wenn keiner der Punkte die Regelgrenze überschreitet (standardmäßig ist dies 3 Sigma), weist die u-Regelkarte darauf hin, dass der Stoffherstellungsprozess innerhalb der statistischen Kontrolle liegt.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	SHEWHART
Erforderliche SAS-Produkte	Base SAS, SAS/QC, SAS/GRAPH
Empfohlene zusätzliche SAS-Produkte	SAS/STAT

u-Regelkarte: Zuweisen von Variablen zu Analyserollen

Klicken Sie im Auswahlfenster auf **Daten**, um auf diese Optionen zuzugreifen.

Standardmäßig wird die Datenquelle als Eingabedatenquelle für die Anwendungsroutine verwendet, die Sie vor dem Öffnen der Anwendungsroutine ausgewählt haben. Der Name der Eingabedatenquelle wird im Feld **Datenquelle** angezeigt. Sämtliche Filter, die in der Datenquelle festgelegt wurden, werden im Feld **Filter** angezeigt. Klicken Sie auf **Bearbeiten**, um die Eingabedatenquelle zu ändern oder zu filtern.

Nachdem Sie eine Datenquelle ausgewählt haben, können Sie Rollen eine oder mehrere Variablen zuweisen.

Optionsname	Beschreibung
Anzahl der Nichtkonformitäten	gibt die Variablen an, welche die Anzahl der Nichtkonformitäten enthalten. Sie müssen dieser Rolle mindestens eine Variable zuweisen
Untergruppen-Identifier	gibt die Variable an, welche die Untergruppen in den Daten identifiziert. Die Werte der Variablen geben an, wie die Beobachtungen in der Eingabetabelle in rationale Untergruppen gegliedert werden. Dies sind die üblichen Werte: ■ die Indizes, welche die Reihenfolge angeben, in der die Untergruppenstichproben gesammelt wurden ■ die Daten oder Zeiten, zu denen die Untergruppenstichproben gesammelt wurden ■ die Etiketten, welche die Untergruppenstichproben eindeutig identifizieren.

Optionsname	Beschreibung
	Ein Untergruppen-Identifier ist erforderlich. Sie können dieser Rolle nur eine Variable zuweisen. Um die Daten nach den identifizierten Untergruppen zu sortieren, aktivieren Sie das Kontrollkästchen **Nach Untergruppe sortieren**.
Stichprobengröße der Untergruppe	gibt die Stichprobengröße der Untergruppe als Werte einer Variablen an. Die Stichprobengröße der Untergruppe ist erforderlich, und Sie können dieser Rolle nur eine numerische Variable zuweisen. Aktivieren Sie das Kontrollkästchen **Numerischen Wert für Untergruppe verwenden (statt Variable)**, wenn die Stichprobengröße der Untergruppe fest ist und Sie einen konstanten Wert für die Größe angeben müssen. Geben Sie den gewünschten Wert im Feld **Stichprobengröße der Untergruppe** ein. Der Wert für die Stichprobengröße der Untergruppe muss zwischen dem maximalen Wert der Prozessvariablen und 99999 (einschließlich) liegen. **Hinweis:** Diese Rolle ist nicht erforderlich, wenn Sie das Kontrollkästchen **Numerischen Wert für Untergruppe verwenden (statt Variable)** im Fenster **Optionen** aktiviert haben.
Analyse gruppieren nach	erzeugt separate Analysen auf Zeilen in von der/den Gruppierungsvariable(n) definierten Gruppen. Standardmäßig werden die Daten aufsteigend nach den Werten dieser Gruppierungsvariablen sortiert. **Hinweis:** Um zu verhindern, dass die Eingabedatei sortiert wird, entfernen Sie das Häkchen im Kontrollkästchen **Nach Variablen sortieren**. **Hinweis:** Sie können Analysen nicht nach einer Variablen gruppieren, die bereits einer Analyserolle zugewiesen ist.
Blockvariablen	gibt bis zu zwei Variablen an, welche die Daten in Blöcke aufeinanderfolgender Untergruppen gruppieren. Diese Blöcke werden in einer Legende beschriftet. Jede Blockvariable liefert eine Etikettenausprägung in der Legende.

Regelkarten: Festlegen von Regelgrenzen

Klicken Sie im Auswahlfenster auf **Regelgrenzen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Sigma-Grenzen	gibt die Breite der Regelgrenzen als Mehrfaches des Standardfehlers (Sigma) der beschreibenden Untergruppenstatistik an, die im Diagramm dargestellt wird. Die Breite muss positiv sein. Der Standardwert für das Mehrfache ist 3.
Berechnungsmethode auswählen	gibt die Berechnungsmethode für die Regelgrenzen an. Folgende Optionen können ausgewählt werden: ■ **Regelgrenzen aus den aktiven Daten berechnen** ■ **Regelgrenzen aus der ausgewählten Datei berechnen** Klicken Sie auf **Durchsuchen**, um die Datei anzugeben, die die Regelgrenzen bzw. die Parameter zu deren Berechnung enthält. Wählen Sie die gewünschte Datei aus und klicken Sie auf **Öffnen**. **Hinweis:** Bei der ausgewählten Datei muss es sich um eine LIMITS= Datei handeln. Weitere Informationen über die in einer LIMITS= Datei erforderlichen Variablen finden Sie in der Prozedurendokumentation zur erstellten Regelkarte. Wenn Sie die beschreibenden Statistiken für alle Untergruppen darstellen möchten, aktivieren Sie das Kontrollkästchen **Beschreibende Statistiken für alle Untergruppen darstellen, unabhängig davon, ob die Stichprobengröße der Untergruppe der nominalen Stichprobengröße für Regelgrenzen entspricht**. Sie können diese Statistiken beispielsweise darstellen, wenn fast alle Untergruppen eine gemeinsame Stichprobengröße haben und wenn Sie die festen (anstelle der variablen) Regelgrenzen, die der nominalen Stichprobengröße *n* entsprechen, anzeigen möchten. Wenn Sie diese Option auswählen und sich Ihre Stichprobengrößen für die Untergruppen deutlich unterscheiden, ist die Interpretation der Regelgrenzen nur für die Untergruppen sinnvoll, deren Stichprobengröße *n* entspricht.

Optionsname	Beschreibung
	Um spezielle Marker für die Datenpunkte hinzuzufügen, bei denen die Stichprobengröße ungleich `n` ist, aktivieren Sie das Kontrollkästchen **Spezielle Marker für Punkte hinzufügen, die Stichprobengrößen entsprechen, welche ungleich nominalen Stichprobengrößen für feste Regelgrenzen sind**. ■ **Regelgrenzen angeben** Geben Sie in den Feldern **Oben**, **Mitte** und **Unten** die obere Regelgrenze, die Mittellinie bzw. die untere Regelgrenze an. Wenn die Stichprobengrößen Ihrer Untergruppe variieren, müssen Sie eine nominale Stichprobengröße angeben. **Hinweis:** Wenn Sie der Rolle **Analyse gruppieren nach** Variablen zuweisen, können Sie keine Werte für Regelgrenzen angeben.

Regelkarte: Auswählen der durchzuführenden Tests

Klicken Sie im Auswahlfenster auf **Tests**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Tests wenden nicht die Regelgrenzen an, die keine Drei-Sigma-Grenzen sind oder die mit der Untergruppenstichprobengröße variieren.

Optionsname	Beschreibung
Test auswählen	Fordern Sie über die Kontrollkästchen einen oder mehrere Tests auf besondere Ursachen an, die auch als Ausführungstests, Mustertests und Western-Electric-Regeln bekannt sind. Diese Tests entdecken bestimmte nichtzufällige Muster in den Punkten, die in der primären Regelkarte dargestellt werden. Das Vorkommen eines nichtzufälligen Musters wird als Signal bezeichnet. Die Tests können auf das Vorhandensein einer besonderen Variationsursache hinweisen. Wenn Sie die Tests auswählen, erscheint eine Beschreibung des jeweiligen Tests im Feld **Beschreibung**.

Optionsname	Beschreibung
Etikett	Geben Sie in den Textfeldern die Etiketten für Punkte an, bei denen die Tests auf besondere Ursachen positiv sind. Die Etiketten dürfen maximal 16 Zeichen lang sein. Jedes Etikett erscheint bei allen Punkten im Test, an denen ein Signal vorliegt.
Text zum Identifizieren der Testsignale	■ Wählen Sie über die Dropdown-Liste eine Farbe für die angegebenen Etiketten aus. ■ Markieren Sie das Kontrollkästchen **Zonenlinien anzeigen**, um Linien einzuschließen, welche die Zonen A, B und C in einem primären Diagramm kennzeichnen. Diese Zonen werden gebildet, indem die Intervalle zwischen den Regelgrenzen in sechs gleichmäßig verteilte Unterintervalle unterteilt werden. ■ Verwenden Sie das Kontrollkästchen **3 Sigma-Grenzen überschreiben**, um nach speziellen Ursachen zu suchen, wenn Sie eine andere Sigma-Regelgrenze als das Standardvielfache von 3 angeben. Um die Sigma-Grenze zu ändern, klicken Sie im Auswahlfenster auf **Regelgrenzen**. ■ Über das Kontrollkästchen **Test auf überlappende Punktmuster anwenden** wenden Sie Tests auf besondere Ursachen auf überlappende Punktmuster an.

Regelkarten: Festlegen von Optionen für die Plotdarstellung

Einige der Darstellungsoptionen sind für unterschiedliche Regelkarten-Anwendungsroutinen gleich. Weitere Informationen zu diesen Optionen erhalten Sie unter folgenden Themen:

- “Festlegen von Achsenoptionen” auf Seite 759
- “Hinzufügen von Referenzlinien” auf Seite 760
- “Festlegen weiterer Plotoptionen” auf Seite 760
- “Festlegen von Blockvariablenoptionen” auf Seite 762

Regelkarten: Speichern von Analyseergebnissen

Klicken Sie im Auswahlfenster auf **Tabellen**, um auf diese Optionen zuzugreifen.

Um die Analyseergebnisse zu speichern, können Sie Ausgabedateien mit diversen Statistiken speichern. Markieren Sie die Kontrollkästchen der Ausgabedaten, die Sie permanent speichern möchten.

Optionsname	Beschreibung
Datei mit Untergruppenstatistiken und Regelgrenzen	erstellt eine Ausgabedatei, welche die im Diagramm dargestellten Informationen enthält. Dies beinhaltet die Werte der Untergruppenvariablen und ihre entsprechenden beschreibenden Statistiken und Regelgrenzen. **Hinweis:** Diese Option erstellt eine Datei OUTTABLE=. Sie können diese Datei mit der Option TABLE= aus der SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Untergruppenstatistiken	erstellt eine Ausgabedatei mit den beschreibenden Statistiken der Untergruppe. **Hinweis:** Diese Option erstellt eine Datei mit OUTHISTORY=. Sie können diese Datei mit der Option HISTORY= aus der SHEWHART-Prozedur wieder in SAS einlesen. Weitere Informationen finden Sie in der Hilfe zur Prozedur SHEWHART.
Ausgabedatei mit Regelgrenzen	erstellt eine Ausgabedatei, welche die Regelgrenzen speichert. Sie können diese Datei verwenden, wenn Sie die Berechnungsmethode für die Regelgrenzen später in dieser Anwendungsroutine auswählen.

Hinweis: SAS Enterprise Guide durchsucht die Liste der Bibliotheken, die im Dialogfeld **Optionen** in **Anwendungsroutinen** > **Ausgabebibliothek** definiert ist und speichert die Ausgabedaten in der ersten, nicht schreibgeschützten Bibliothek aus dieser Liste.

Das SAS Add-In for Microsoft Office speichert die Ausgabedaten in der Bibliothek **Work**. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für Ihre Ausgabedaten anzugeben.

Festlegen von Titeln und Fußnoten

Klicken Sie im Auswahlfenster auf **Titel**, um auf diese Optionen zuzugreifen.

Der Standardtitel- und -fußnotentext für die Ergebnisse wird automatisch beibehalten.

So geben Sie Ihre eigenen Titel und Fußnoten an:

1 Wählen Sie Titel oder Fußnote im Feld **Abschnitt** aus.

2 Deaktivieren Sie das Kontrollkästchen **Standardtext verwenden**.

3 Bearbeiten Sie den Text für den Titel bzw. die Fußnote im Textfeld.

Hinweis: Sie können in Titeln und Fußnoten Makrovariablen verwenden.

Anzeigen von Eigenschaften

Klicken Sie im Auswahlfenster auf **Eigenschaften**, um auf diese Optionen zuzugreifen.

Sie können im Fenster **Eigenschaften** eine Zusammenfassung der Eigenschaften für eine Anwendungsroutine anzeigen. Bei einigen Anwendungsroutinen können Sie die Eigenschaften bearbeiten. Klicken Sie dazu auf **Bearbeiten**. Das Dialogfeld **Eigenschaften** wird geöffnet.

Im Fenster **Eigenschaften** können folgende Informationen angezeigt werden:

- das Etikett für die Anwendungsroutine
- der zur Ausführung dieser Anwendungsroutine verwendete Server
- Datum und Uhrzeit, zu dem/der die Anwendungsroutine erstellt wurde
- Datum und Uhrzeit der letzten Änderungen an der Anwendungsroutine
- Uhrzeit der letzten Ausführung der Anwendungsroutine
- alle Beschränkungen für Daten
- alle verwendeten Eingabeaufforderungen
- das Format aller Ergebnisse
- das Format für die Grafikausgabe in den Ergebnissen

Hinweis: Wenn Sie das Etikett für eine vorhandene Anwendungsroutine ändern, wird das neue Etikett gespeichert, wenn Sie im Dialogfeld **Eigenschaften** auf **OK** klicken. Das neue Etikett wird auch verwendet, wenn Sie die Anwendungsroutine abbrechen.

93

Aktualisieren von Bibliotheksmetadaten

Allgemeines über den Assistenten 'Bibliotheksmetadaten aktualisieren' 733

Bibliotheksmetadaten aktualisieren: Auswählen des SAS-Servers und der SAS-Bibliothek 734

Bibliotheksmetadaten aktualisieren: Ausführen eines Berichts für die Metadaten oder Aktualisieren der Metadaten 734

Allgemeines über den Assistenten 'Bibliotheksmetadaten aktualisieren'

Mit dem Assistenten **Bibliotheksmetadaten aktualisieren** können Sie einen Bericht über die Tabellendefinitionen im Metadaten-Repository erstellen oder diese Tabellendefinitionen aktualisieren. Sie können zwei Arten von Berichten erstellen: einen Bericht, in dem der Inhalt der Bibliothek basierend auf den Tabellendefinitionen im Metadaten-Repository aufgeführt ist, oder einen Bericht, der die Unterschiede zwischen den Tabellen in der physischen Bibliothek und den Tabellendefinitionen im Metadaten-Repository angibt. Abgesehen von der Berichtserstellung können Sie mit dem Assistenten **Bibliotheksmetadaten aktualisieren** auch Tabellendefinitionen im Metadaten-Repository aktualisieren, hinzufügen oder löschen.

Standardmäßig zeigt SAS Enterprise Guide nur die Tabellen an, die in den Metadaten registriert sind. Der Assistent **Bibliotheksmetadaten aktualisieren** verwendet die METALIB-Prozedur, um diese Tabellen im Metadaten-Repository programmatisch zu registrieren und die Metadaten für diese Tabellen zu synchronisieren.

Hinweis: Diese Anwendungsroutine ist nur im SAS Enterprise Guide verfügbar.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	METALIB
Erforderliche SAS-Produkte	Base SAS
Empfohlene zusätzliche SAS-Produkte	Ohne

Bibliotheksmetadaten aktualisieren: Auswählen des SAS-Servers und der SAS-Bibliothek

1. Wählen Sie den Server in der Liste der verfügbaren Server aus.
2. Wählen Sie eine Bibliothek aus.
3. Klicken Sie auf **Weiter**, um einen Bericht auszuführen oder um die Metadaten für die ausgewählte Bibliothek zu aktualisieren.

Bibliotheksmetadaten aktualisieren: Ausführen eines Berichts für die Metadaten oder Aktualisieren der Metadaten

1. Bevor Sie einen Bericht ausführen oder die Metadaten aktualisieren können, müssen Sie den SAS-Server und die SAS-Bibliothek auswählen.
2. Wählen Sie eine dieser Optionen aus:

Optionsname	Beschreibung
Bericht zu den Inhalten der Bibliothek, wie im Metadaten-Repository registriert	erstellt einen Bericht über den Inhalt der Bibliothek basierend auf den Tabellendefinitionen im Metadaten-Repository.
Bericht zu den Unterschieden zwischen physischen Tabellen und dem Metadaten-Repository	erstellt einen Bericht mit den folgenden Informationen: ■ die Tabellendefinitionen im Metadaten-Repository, die aktualisiert werden müssten, um mit den physischen Tabellen übereinzustimmen ■ die Tabellen in der physischen Bibliothek, die keine Tabellendefinitionen im Metadaten-Repository enthalten ■ die Tabellen, die nicht mehr in der physischen Bibliothek vorhanden sind, die aber noch Tabellendefinitionen im Metadaten-Repository enthalten
Tabellendefinitionen in Metadaten mit tatsächlichen Tabellen und Spalten aktualisieren und ergänzen	führt folgende Schritte aus: ■ Bei Tabellen in der physischen Bibliothek, die Metadaten enthalten, werden die Tabellendefinitionen im Metadaten-Repository aktualisiert, sodass sie mit den entsprechenden physischen Tabellen übereinstimmen. Diese Aktualisierungen enthalten Informationen über die Spalten, Indizes, eindeutige Schlüsse, Fremdschlüssel und Schlüsselassoziationen. ■ Für Tabellen in der physischen Bibliothek, die keine Metadaten enthalten, werden diese Tabellendefinitionen in das Metadaten-Repository eingefügt.
Nur existierende Tabellendefinitionen in Metadaten mit aktuellen Spalteninformationen aktualisieren	Aktualisiert die Tabellendefinitionen im Metadaten-Repository für diese Tabellen in der physischen Bibliothek, die Metadaten besitzen.

Optionsname	Beschreibung
Veraltete Einträge aus der Metadaten-Bibliothek löschen (Tabellen, die nicht mehr existieren)	Löscht die Tabellendefinitionen für Tabellen aus dem Metadaten-Repository, die in der physischen Bibliothek nicht mehr vorhanden sind.

3 Bevor Sie Metadaten aktualisieren oder löschen können, müssen Sie eventuell die Benutzer-ID und das Passwort für ein Konto mit Administratorrechten eingeben. Wenn Benutzername und Passwort nicht erforderlich sind, dann sind diese Optionen nicht verfügbar.

4 Klicken Sie auf **Fertig stellen**, um den Bericht auszuführen oder die Metadaten in der Bibliothek zu aktualisieren.

94

Zu CAS hochladen

Allgemeines über die Anwendungsroutine 'Zu CAS hochladen' **737**
Hochladen von Daten auf einen SAS CAS Analytic Server **738**

Allgemeines über die Anwendungsroutine 'Zu CAS hochladen'

Die Anwendungsroutine **Zu CAS hochladen** ermöglicht es Ihnen, SAS 9.4-Daten auf einen CAS-Server (CAS, Cloud Analytics Services) hochzuladen. Die Anwendungsroutine verwendet SAS/CONNECT, um Daten auf einen Remoteserver zu verschieben. Mit der Einführung von SAS 9.4M5 stehen zusätzliche und gleichfalls effizientere Methoden zum Verschieben von Daten auf einen CAS-Server zur Verfügung. So können Sie zum Beispiel die CAS LIBNAME-Engine oder die Prozedur CASUTIL verwenden. Weitere Information, siehe *SAS Cloud Analytic Services: User's Guide*.

Hinweis: Um diese Anwendungsroutine in SAS Enterprise Guide oder SAS Add-In for Microsoft Office auszuführen, müssen Sie Ihre Umgebung konfigurieren. Weitere Informationen, siehe *SAS Intelligence Platform: Desktop Application Administration Guide*.

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	Ohne
Erforderliche SAS-Produkte	Base SAS, SAS Viya

Name der Anforderung	Prozedur- und Produktnamen
Empfohlene zusätzliche SAS-Produkte	Ohne

Hochladen von Daten auf einen SAS CAS Analytic Server

1. Wählen Sie den CAS-Server aus.
2. Wählen Sie die Bibliothek aus, in der die Daten gespeichert werden sollen.
3. Geben Sie den Membernamen der Tabelle an. Standardmäßig ist der Membername identisch mit dem Namen der Datei. Bei verteilten Ladeprozessen ist es nicht möglich, den Membernamen zu ändern.
4. Geben Sie an, wo die Anmeldedaten für die ausgewählte Tabelle zu finden sind. Besteht bereits eine Verbindung zu einem SAS Metadata Server, können Sie die in den Metadaten gespeicherten Anmeldedaten verwenden. Wenn Sie lokale Daten auf den CAS-Server hochladen, müssen der Benutzername und das Passwort für den CAS-Server verwendet werden.

95

Dateien auf Server hochladen

Allgemeines über die Anwendungsroutine 'Dateien auf Server hochladen' 739
Dateien auf Server hochladen: Auswählen der hochzuladenden SAS-Datendateien . . 740
Dateien auf Server hochladen: Angeben der Optionen zum Hochladen 740

Allgemeines über die Anwendungsroutine 'Dateien auf Server hochladen'

Mit der Anwendungsroutine **Dateien auf Server hochladen** können Sie aus Ihrem lokalen Microsoft Windows-Dateisystem eine oder mehrere SAS-Dateien (*.sas7bdat) auswählen und dann auf einen SAS-Server kopieren.

Die Anwendungsroutine erstellt ein Log, das eine Zusammenfassung folgender Informationen enthält: die kopierten Datendateien, die Größe in Byte, Zeiten für jede Datei und für alle Dateien als Gruppe sowie Fehler, die während der Übertragung aufgetreten sind. Standardmäßig werden alle Dateien als Ausgabedaten in Ihr SAS Enterprise Guide-Projekt eingefügt, sodass sie im Projekt verwendet werden können.

Dateien auf Server hochladen: Auswählen der hochzuladenden SAS-Datendateien

1 Klicken Sie auf **Hinzufügen**, um die SAS-Datendateien auszuwählen, die hochgeladen werden sollen.

2 Wählen Sie den Namen der zu kopierenden Datei.

Hinweis: Bei den Dateien muss es sich um SAS7BDAT-Dateien handeln.

3 Klicken Sie auf **Öffnen**, um die SAS-Datendateien auszuwählen, die Sie kopieren möchten.

Hinweis: Alle kopierten Dateien erhalten auf dem Zielserver Namen in Kleinbuchstaben.

4 Klicken Sie auf **Weiter**, um die Optionen zum Hochladen festzulegen.

Dateien auf Server hochladen: Angeben der Optionen zum Hochladen

1 Bevor Sie die Optionen zum Hochladen angeben können, müssen Sie zuerst die hochzuladenden Datendateien auswählen.

2 Wählen Sie in der Dropdown-Liste **SAS-Server auswählen** den Server aus, auf den die ausgewählten Datendateien kopiert werden sollen.

3 Wählen Sie in der Dropdown-Liste **Zielbibliothek auswählen** die SAS-Bibliothek, die verwendet werden soll.

4 Wenn Sie bereits auf dem SAS-Server vorhandene Dateien mit demselben Namen wie die zu kopierenden Dateien überschreiben möchten, aktivieren Sie das Kontrollkästchen **Vorhandene Dateien mit demselben Namen überschreiben** . Standardmäßig werden vorhandene Dateien nicht überschrieben

5 Wenn die zu kopierenden Datendateien nicht in Ihrem aktuellen SAS Enterprise Guide-Projekt gespeichert werden sollen, deaktivieren Sie das Kontrollkästchen **Dateien zu Ihrem aktuellen Projekt hinzufügen**. Standardmäßig werden die kopierten Dateien dem aktuellen Projekt hinzugefügt.

6 Klicken Sie auf **Fertig stellen**, um die Dateien zu kopieren.

Hinweis: Wenn es sich bei der Zielbibliothek um eine Base SAS Engine-Bibliothek mit einem einzelnen Dateipfad handelt, erfolgt die Dateiübertragung direkt in diesen Dateipfad. Handelt es sich bei der Zielbibliothek nicht um eine Base SAS Engine-Bibliothek oder handelt es sich um eine verkettete Bibliothek, werden die Dateien zuerst in eine Staging Area in WORK übertragen. Anschließend werden die Dateien mittels der PROC SQL in die Zielbibliothek kopiert.

96

Auf LASR hochladen

Allgemeines über die Anwendungsroutine 'Auf LASR hochladen' **743**
Hochladen von Daten auf einen SAS LASR Analytic Server **744**

Allgemeines über die Anwendungsroutine 'Auf LASR hochladen'

Der SAS LASR Analytic Server ist eine In-Memory-Analytics-Plattform, der speziell dafür entwickelt wurde, die Zeit für analytische Berechnungen zu verringern. Der Server wird in Umgebungen eingesetzt, die aus nicht-verteilten Deployments ohne große Datenvolumen bestehen. Zur Analyse großer Dateien verwendet der verteilte Server eine verteilte Berechnungsumgebung sowie parallele Arbeitsprozesse. Mit der Anwendungsroutine **Auf LASR hochladen** haben Sie die Möglichkeit, Ihre Daten auf einen LASR Analytic Server hochzuladen und auf diese Daten über SAS Visual Analytics zuzugreifen.

Hinweis: Um diese Anwendungsroutine ausführen zu können, muss eine Verbindung zwischen SAS Enterprise Guide bzw. SAS Add-In for Microsoft Office und demselben Metadatenserver bestehen, der auch für SAS Visual Analytics verwendet wird.

Weitere Informationen darüber, wie Sie eine Verbindung zu einem SAS LASR Analytic Server einrichten, finden Sie unter *SAS Intelligence Platform: Desktop Application Administration Guide*

Um diese Anwendungsroutine ausführen zu können, müssen folgende Anforderungen erfüllt sein:

Name der Anforderung	Prozedur- und Produktnamen
Verwendete SAS-Prozeduren	Ohne
Erforderliche SAS-Produkte	Base SAS, SAS LASR Analytic Server
Empfohlene zusätzliche SAS-Produkte	SAS Visual Analytics

Hochladen von Daten auf einen SAS LASR Analytic Server

1. Wählen Sie die Bibliothek aus, in der die Daten gespeichert werden sollen.
2. Geben Sie den Membernamen der Tabelle an. Standardmäßig ist der Membername identisch mit dem Namen der Datei. Bei verteilten Ladeprozessen ist es nicht möglich, den Membernamen zu ändern.
3. Geben Sie den Verzeichnispfad für die hochgeladenen Daten an. Standardmäßig besitzt jede Bibliothek einen vordefinierten Speicherort.

Anhang 1

Allgemeine Darstellungsoptionen für Grafik-Anwendungsroutinen

Festlegen von Achsenoptionen . . . **745**
Siehe auch . . . 748
Hinzufügen von Referenzlinien . . . **748**
Siehe auch . . . 749
Festlegen von Legendenoptionen . . . **749**
Siehe auch . . . 750
Anpassen des Diagrammbereichs . . . **750**
Siehe auch . . . 751
Festlegen der Balkendarstellung . . . **751**
Siehe auch . . . 752
Festlegen von Balkenoptionen . . . **753**
Siehe auch . . . 753
Festlegen der Liniendiagramm-Optionen . . . **753**
Siehe auch . . . 754
Auswählen der Interpolationsmethode . . . **754**
Siehe auch . . . 755
Festlegen erweiterter Optionen . . . **756**
Siehe auch . . . 757

Festlegen von Achsenoptionen

Diese Optionen werden im Auswahlfenster unter **Darstellung > Achsen** angezeigt.

Hinweis: Die Achsenoptionen werden nicht angewendet, wenn Sie ein Oberflächenplot erstellen, bei dem Sie GIF als grafisches Ausgabeformat verwenden. Um dieses Achsenoptionen in einem Oberflächenplot anzuwenden, müssen Sie ActiveX als grafisches Ausgabeformat verwenden.

Optionsname	Beschreibung
Achsen und Markierungen deaktivieren	unterdrückt die Achsen und Achsenmarkierungen.
Achse	passt die horizontale bzw. vertikale Achse an. ■ Bei einigen Plots können Sie auch Optionen für die vertikale rechte Achse festlegen. ■ Bei dreidimensionalen Streuungs-, Nadel- und Oberflächendiagrammen können Sie diese Optionen für die Tiefenachse definieren. ■ Für gruppierte Diagramme oder Balkendiagramme mit mehreren Maßen können Sie diese Optionen für die Gruppenachse angeben. So passen Sie die Darstellung einer Achse an: ■ Geben Sie bei horizontalen, vertikalen und Tiefenachsen die Farbe, die Breite und den Stil der Linie für die Achse an. ■ Um die Reihenfolge der Werte für die horizontale bzw. vertikale Achse umzukehren, aktivieren Sie die Option **Achse umkehren**. Wenn Sie dieses Kontrollkästchen für die horizontale Achse aktivieren, ist die Reihenfolge der X-Werte umgekehrt. Wenn Sie dieses Kontrollkästchen für die vertikale Achse aktivieren, ist die Reihenfolge der Y-Werte umgekehrt. **Hinweis:** Diese Option ist nur für Flächendiagramme, Box-Plots, Blasendiagramme, Konturdiagramme, Liniendiagramme und zweidimensionale Streuungsdiagramme verfügbar. ■ Um das Etikett einer Achse anzupassen, klicken Sie auf den Reiter **Etikett**. Geben Sie den Text für Ihr benutzerdefiniertes Etikett in das Feld **Etikett** ein. Legen Sie mithilfe der Dropdown-Liste **Etikettrotation** fest, wie das Etikett in Beziehung zur Achse angezeigt werden soll. Sie können ebenso den Typ, die Größe und die Farbe der Schriftart festlegen. Sie können wählen, ob der Text unterstrichen, fett oder kursiv geschrieben werden soll. Ändern Sie mithilfe der Schaltflächen die Ausrichtung. **Hinweis:** Wenn Sie ein Balkendiagramm mit mehreren Maßen erstellen, ist das Standardetikett für die abhängige Variable die Statistik, die zur Berechnung der Länge der einzelnen Balken verwendet wurde.

Optionsname	Beschreibung
	■ Um die Werte einer Achse anzupassen, klicken Sie auf den Reiter **Werte**. Sie können den Typ, die Größe und die Farbe der Schriftart für die Werte festlegen. Sie können wählen, ob der Text unterstrichen, fett oder kursiv geschrieben werden soll. Ändern Sie mithilfe der Schaltflächen die Ausrichtung. Legen Sie mithilfe der Dropdown-Liste **Werterotation** fest, wie die Werte in Beziehung zur Achse angezeigt werden sollen.
Hauptmarkierungen	legt die Hauptmarkierungen für die horizontale und vertikale Achse fest. Bei einigen Plots können Sie auch Optionen für die vertikale rechte Achse festlegen. Bei 3D-Streuungs-, 3D-Nadel- und 3D-Oberflächendiagrammen können Sie auch Achsenmarkierungen für die Tiefenachse definieren. Bei Balkendiagrammen können Sie nur Achsenmarkierungen für die abhängige Variable festlegen. Wenn Sie z. B. ein vertikales Balkendiagramm erstellen, können Sie Achsenmarkierungen für die vertikale, nicht jedoch für die horizontale Achse festlegen. Wählen Sie **Achsenhauptmarkierungen für *Achsenname* anzeigen** und legene Sie fest, wie die Anzahl der zu verwendenden Markierungen ermittelt werden soll. ■ Um die Anzahl und Häufigkeit der Hauptmarkierungen durch die Software festlegen zu lassen, markieren Sie **Automatisch**. Wenn Sie ein Flächendiagramm, Blasendiagramm, Liniendiagramm oder ein 2D-Streuungsdiagramm erstellen, aktivieren Sie das Kontrollkästchen **Bei null beginnen**, wenn Sie möchten, dass die Werte der Markierungen bei 0 beginnen. ■ Um die Anzahl der zu verwendenden Achsenmarkierungen festzulegen, wählen Sie **Anzahl**. Geben Sie die gewünschte Anzahl in das Textfeld ein. ■ Um ein logarithmisches Achsenmarkierungsmuster anzugeben, wählen Sie **Log**. Wählen Sie Basis und Potenz aus und bestimmen Sie, ob die Markierungen nach Werten der Basis oder nach Werten der Potenz angeordnet werden sollen. ■ Um spezielle Positionen für die Achsenmarkierungen festzulegen, wählen Sie **Angeben**. Geben Sie einen Achsenmarkierungswert in das Textfeld ein und klicken Sie auf **Hinzufügen**. Um einen Wert zu ändern, wählen Sie diesen aus der Liste aus und klicken auf **Bearbeiten**. Wenn Sie die Änderungen vorgenommen haben, klicken Sie auf **Speichern**.
Nebenmarkierungen	geben Sie die Nebenmarkierungen für die horizontale bzw. vertikale Achse an.

Optionsname	Beschreibung
	■ Bei einigen Plots können Sie auch Optionen für die vertikale rechte Achse festlegen. ■ Bei 3D-Streuungs-, 3D-Nadel- und 3D-Oberflächendiagrammen können Sie auch Achsenmarkierungen für die Tiefenachse definieren. ■ Bei Balkendiagrammen können Sie nur Achsenmarkierungen für die abhängige Variable festlegen. Wenn Sie z. B. ein vertikales Balkendiagramm erstellen, können Sie Achsenmarkierungen für die vertikale, nicht jedoch für die horizontale Achse festlegen.

Siehe auch

- Kapitel 3, “Flächendiagramm” auf Seite 15
- Kapitel 6, “Balkendiagramm” auf Seite 39
- Kapitel 8, “Balken-Linien-Diagramm” auf Seite 51
- Kapitel 11, “Box-Plot” auf Seite 75
- Kapitel 12, “Blasendiagramm” auf Seite 83
- Kapitel 19, “Konturdiagramm” auf Seite 137
- Kapitel 42, “Liniendiagramm” auf Seite 309
- Kapitel 77, “Streuungsdiagramm” auf Seite 619
- Kapitel 87, “Oberflächenplot” auf Seite 683

Hinzufügen von Referenzlinien

Diese Optionen werden im Auswahlfenster unter der Überschrift **Darstellung > Referenz** angezeigt.

Hinweis: Diese Optionen sind nicht für 3D-Streuungsdiagramme verfügbar.

Optionsname	Beschreibung
Gitterlinien anzeigen	zeigt die Gitterlinien an.

Optionsname	Beschreibung
	Hinweis: Diese Option ist für Balkendiagramme nicht verfügbar.
Referenzlinien	fügt Referenzlinien zur horizontalen bzw. vertikalen Achse hinzu. Bei einigen Plots können Sie auch Referenzlinien für die vertikale rechte Achse hinzufügen. **Hinweis:** Bei Balkendiagrammen werden die Referenzlinien für die abhängige Variable gezeichnet. Bei vertikalen Balkendiagrammen erscheinen die Referenzlinien als horizontale Linien und bei horizontalen Balkendiagrammen als vertikale Linien. Für 3D Plots können keine Referenzlinien erstellt werden.

Siehe auch

- Kapitel 3, “Flächendiagramm” auf Seite 15
- Kapitel 6, “Balkendiagramm” auf Seite 39
- Kapitel 8, “Balken-Linien-Diagramm” auf Seite 51
- Kapitel 11, “Box-Plot” auf Seite 75
- Kapitel 12, “Blasendiagramm” auf Seite 83
- Kapitel 19, “Konturdiagramm” auf Seite 137
- Kapitel 42, “Liniendiagramm” auf Seite 309
- Kapitel 77, “Streuungsdiagramm” auf Seite 619

Festlegen von Legendenoptionen

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Legende**, um auf diese Optionen zuzugreifen.

Sie können Legenden für Mehrbereichsplots, Mehrlinienplots, Balken-Linien-Diagramme, Konturplots, Karten, gestapelten Balkendiagramme, Kreisdiagramme und Ringdiagramme erstellen. Darüber hinaus stehen Legenden für Balkendiagramme mit mehreren Maßen zur Verfügung, wenn Sie der Rolle **Stapel** eine Variable zugewiesen haben.

Für Balken-Liniendiagramme können zwei Legenden erstellt werden: eine für das Balkendiagramm und eines für den Linienplot.

Legenden sind auch verfügbar, wenn Sie ein Streuungsdiagramm erstellen und den Rollen **Vertikal** und **Vertikal (Rechts)** Variablen zugewiesen haben. Da in derselben Grafik mehrere Streuungsdiagramme vorhanden sind, können Sie mit Hilfe der Legende zwischen Datenpunkten für die **vertikale** Variable und Datenpunkten für die **vertikal rechte** Variable unterscheiden.

Siehe auch

- Kapitel 3, “Flächendiagramm” auf Seite 15
- Kapitel 6, “Balkendiagramm” auf Seite 39
- Kapitel 8, “Balken-Linien-Diagramm” auf Seite 51
- Kapitel 19, “Konturdiagramm” auf Seite 137
- Kapitel 30, “Ringdiagramm” auf Seite 207
- Kapitel 42, “Liniendiagramm” auf Seite 309
- Kapitel 61, “Kreisdiagramm” auf Seite 479

Anpassen des Diagrammbereichs

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Diagrammbereich**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Größe benutzerdef. Diagramm	legt die Bildgröße fest. Zum Vergleich wird die aktuelle Bildschirmgröße angezeigt.
Diagrammhintergrundfarbe	legt die Hintergrundfarbe fest. **Hinweis:** Wenn Sie ein Kartendiagramm erstellen, wählen Sie die Hintergrundfarbe für die Karte aus der Dropdown-Liste **Hintergrundfarbe für Karte** aus.
Rahmen um Plotfläche zeichnen	zeichnet einen Rahmen um den Diagrammbereich herum. Wählen Sie über die Dropdown-Liste **Hintergrund der Plotfläche** eine Farbe für die Plotfläche aus. **Hinweis:** Diese Option ist für Ringdiagramme, Kartendiagramme, Kreisdiagramme, dreidimensionale gruppierte Balkendiagramme, dreidimensionale Streuungsdiagramme oder Oberflächenplots nicht verfügbar.

Optionsname	Beschreibung
Anzeigen von Diagrammtipps	legt fest, ob Diagrammtipps angezeigt werden. Der Diagrammtipp gilt für das gesamte Diagramm. Er bezieht sich nicht auf die einzelnen Bestandteile, wie Balken oder Linien. **Hinweis:** Der Diagrammtipp ist bei Kartendiagrammen nur dann sichtbar, wenn Sie ActiveX oder Java als Format für die Grafikausgabe gewählt haben.

Siehe auch

- Kapitel 3, “Flächendiagramm” auf Seite 15
- Kapitel 6, “Balkendiagramm” auf Seite 39
- Kapitel 8, “Balken-Linien-Diagramm” auf Seite 51
- Kapitel 11, “Box-Plot” auf Seite 75
- Kapitel 12, “Blasendiagramm” auf Seite 83
- Kapitel 19, “Konturdiagramm” auf Seite 137
- Kapitel 30, “Ringdiagramm” auf Seite 207
- Kapitel 42, “Liniendiagramm” auf Seite 309
- Kapitel 61, “Kreisdiagramm” auf Seite 479
- Kapitel 69, “Radardiagramm” auf Seite 559
- Kapitel 77, “Streuungsdiagramm” auf Seite 619
- Kapitel 87, “Oberflächenplot” auf Seite 683

Festlegen der Balkendarstellung

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Balken**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Angeben der Balkenfarbe	gibt die Farbe der Balken an. Folgende Optionen können ausgewählt werden: ■ **Standardfarbschema** verwendet das Standardfarbschema für die Balken. Das Standardfarbschema wird durch den aktuellen Stil bestimmt. Sie können den Stil mit Hilfe des

Optionsname	Beschreibung
	Stil-Managers oder des Dialogfelds **Eigenschaften** für das Ergebnis festlegen. In SAS Enterprise Guide können Sie im Dialogfeld **Optionen** auch den Stil für jeden Ergebnistyp festlegen. Im SAS Add-In for Microsoft Office können Sie im Dialogfeld **SAS Add-In - Optionen** Stile auf die SAS Report- und HTML-Formate anwenden. ■ **Existierende Farben (in einer früheren Version angegeben)** ermöglicht das Verwenden von Farben, die in einer früheren Version von SAS Enterprise Guide oder SAS Add-In for Microsoft Office festgelegt wurden. Diese Option ist nur für SAS-Inhalte verfügbar, die Sie von einer früheren Version migriert haben. ■ **Benutzerdefinierte Farben** - Ermöglicht die Angabe benutzerdefinierter Farben für die Balken. Sie können benutzerdefinierte Farben z. B. für Balkendiagramme festlegen, die mehr als eine Farbe benötigen, wie etwa gruppierte, gestapelte und überlappende Balkendiagramme. Mit den Dropdown-Listen können Sie bis zu 12 Farben für das Balkendiagramm bestimmen. wenn das Diagramm mehr als 12 Farben erfordert, werden die zusätzlichen Farben aus dem aktuellen Standardfarbschema ausgewählt. Klicken Sie auf **Zurücksetzen**, um die benutzerdefinierten Farben wieder auf den Standard zurückzusetzen. **Hinweis:** Je nach Balkendiagrammtyp werden eventuell nicht alle benutzerdefinierten Farben verwendet.
Balkenanzahl festlegen	gibt die Anzahl der Balken im Diagramm an. Standardmäßig wird die Anzahl der Balken automatisch bestimmt. Bei numerischen Spalten werden die Werte in Bereiche geteilt, mit einem Balken pro Bereich. Wenn die numerische Spalte diskrete Werte darstellt, möchten Sie eventuell die Option **Ein Balken pro eindeutigem Datenwert** verwenden. **Hinweis:** Sie können die Anzahl der Balken für Balkendiagramme mit mehreren Messgrößen angeben. Aktivieren Sie das Kontrollkästchen **Balkenanzahl angeben**, um anzugeben, wie viele Balken im Diagramm angezeigt werden sollen.

Siehe auch

- Kapitel 6, “Balkendiagramm” auf Seite 39
- Kapitel 8, “Balken-Linien-Diagramm” auf Seite 51

Festlegen von Balkenoptionen

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Balkenoptionen**, um auf diese Optionen zuzugreifen.

Mit Hilfe der Dropdown-Liste **Sortierreihenfolge** können Sie die Sortierreihenfolge für die Balken bestimmen. Standardmäßig wird die Sortierreihenfolge automatisch festgelegt. Sie können die Balken in aufsteigender oder absteigender Reihenfolge der Länge oder Höhe anordnen.

Aktivieren Sie das Kontrollkästchen **Umrissfarbe**, wenn Sie jeden Balken mit einem Umriss versehen möchten. Sie können eine Farbe für den Umriss aus der Dropdown-Liste auswählen.

Mit Hilfe der Dropdown-Liste **Balkengröße** können Sie die Balkengröße bestimmen. Standardmäßig wird die Balkengröße automatisch festgelegt. Sie können jedoch die Balkenbreite oder den Abstand zwischen den Balken als Prozentwert der gesamten Diagrammbreite im Textfeld angeben.

Siehe auch

Kapitel 8, “Balken-Linien-Diagramm” auf Seite 51

Festlegen der Liniendiagramm-Optionen

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Plots**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Optionen sind nur dann verfügbar, wenn Sie der Rolle **Liniensumme von** mindestens eine Variable zuweisen. Wenn Sie der Rolle **Liniensumme von** keine Variable zugewiesen haben, können Sie in der Dropdown-Liste **Verwendete Statistik bei Liniendarstellung** im Fenster **Erweitert** die bei der Darstellung der Linie in einem Balken-Liniendiagramm verwendete Statistik auswählen.

So steuern Sie die Darstellung eines Liniendiagramms für eine bestimmte Variable:

1 Wählen Sie eine Variable aus der Liste aus. Bei den Variablen in der Liste handelt es sich um die Variablen, die Sie der Rolle **Liniensumme von** zugewiesen haben.

2 Geben Sie im Bereich **Linie** den Linienstil, die Linienbreite- und die Linienfarbe an.

3 Wählen Sie eine Umrissfarbe aus.

4 Geben Sie im Bereich **Datenpunktmarker** den Typ, das zu verwendende Symbol, die Höhe für jeden Marker und die Farbe der Datenpunktmarker an.

Hinweis: Manche Symbole werden nicht von allen Grafikausgabeformaten unterstützt. Wenn Sie ein nicht durch das Ausgabeformat unterstütztes Symbol auswählen, wird stattdessen ein anderes Symbol (ein unterstütztes Symbol) angezeigt.

5 Wählen Sie in der Dropdown-Liste **Verwendete Statistik bei Liniendarstellung** die zu verwendende Statistik aus, wenn die Linie auf einem Balken-Liniendiagramm dargestellt wird. Folgende Statistiken stehen zur Auswahl:

- **Durchschnitt** ist der durchschnittliche Wert der Variablen, welcher der Rolle **Liniensumme von** zugewiesen wurde. Dieser Mittelwert wird für jeden eindeutigen Wert berechnet, der in der Linie als Punkt dargestellt wird. Um beispielsweise die durchschnittlichen Umsätze jeder Stadt in der Datei anzuzeigen, weisen Sie der Rolle **Liniensumme von** die Variable **Sales** zu und wählen **Durchschnitt** als Statistik zur Darstellung der Linie.
- **Summe** ist die Summe der Variablen, die der Rolle **Liniensumme von** zugewiesen wurde. Diese Summe wird für jeden eindeutigen Wert berechnet, der in der Linie als ein Punkt dargestellt wird. Wenn Sie z. B. der Rolle **Liniensumme von** die Rolle Sales zuweisen und die Werte der Variablen Sales für die Stadt Denver 8734, 982, 1504, 3207, 4502, 624 und 918 sind, dann ist die Summenstatistik für Denver 20.471.

Siehe auch

Kapitel 8, “Balken-Linien-Diagramm” auf Seite 51

Auswählen der Interpolationsmethode

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Interpolationen**, um auf diese Optionen zuzugreifen.

So wählen Sie eine Interpolationsmethode aus:

1 Wählen Sie eine Variable aus. Wenn Sie ein Liniendiagramm oder ein zweidimensionales Streuungsdiagramm erstellen, entsprechen die Variablen in der Liste den Variablen, die Sie den Rollen **Vertikal** und **Vertikal (Rechts)** zugewiesen haben. Wenn Sie **Mehrere Liniendiagramme nach Gruppenspalte**

als Typ für das Liniendiagramm ausgewählt haben, enthält diese Liste die eindeutigen Werte der Variablen, die Sie der Rolle **Gruppe** zugewiesen haben.

2 Wählen Sie eine Interpolationsmethode aus der Dropdown-Liste aus. Je nach ausgewählter Interpolationsmethode kann die Angabe weiterer Einstellungen erforderlich sein. Diese Interpolationsmethoden stehen zur Verfügung:

- **LaGrange** zeichnet eine glatte Kurve durch die Datenpunkte. Diese Methode wird in erster Linie verwendet, wenn die Daten aus exakten, durch Tabulatoren getrennten Werten bestehen. Ein Polynom des angegebenen Grads (1, 3 oder 5) wird durch die nächsten 2, 4 oder 6 Punkte angepasst. Im Allgemeinen ist die erste Ableitung nicht stetig. Wenn die Werte der X-Variablen nicht streng zunehmend sind, wird die entsprechende parametrische Methode verwendet.
- **Linie** - Verbindet die Datenpunkte durch Geraden. Die Punkte werden in der Reihenfolge verbunden, in der sie in der Eingabedatei vorkommen.
- **Nadel** zieht eine vertikale Linie von jedem Datenpunkt zu einer horizontalen Linie am 0-Wert auf der vertikalen Achse oder am Minimalwert auf der vertikalen Achse, wenn dieser größer als 0 ist. Die horizontale Linie wird automatisch gezeichnet.
- **Regression** erstellt ein Regressionsanalyseplot. Diese Interpolationsmethode ist nur verfügbar, wenn Sie den Rollen Horizontal und Vertikal numerische Variablen zuweisen.
- **Streuung** unterdrückt die Interpolation. Es wird eine Darstellung der Datenpunkte erstellt.
- **Glättung** - passt eine glatte Linie mithilfe einer Splineroutine an die Daten an. Diese Methode dient zum Glätten verrauschter Daten. Die Punkte auf dem Plot liegen nicht unbedingt auf der Linie.
- **Spline** - stellt die Linie mithilfe einer Splineroutine dar.
- **STD** verbindet den mittleren Y-Wert mit ± 1, 2 oder 3 Standardabweichungen für jeden X-Wert mit einer durchgehenden Linie. Standardmäßig werden zwei Standardabweichungen verwendet.
- **Stufe** - stellt die Daten mithilfe einer Stufenfunktion dar.

3 (Optional) Aktivieren Sie das Kontrollkästchen **Für alle übernehmen**, um die gleiche Interpolationsmethode und die gleichen Eigenschaftseinstellungen für alle Variablen der Liste zu verwenden. Diese Option ist nur dann verfügbar, wenn sich mehr als eine Variable in der Liste befinden.

Siehe auch

- Kapitel 42, “Liniendiagramm” auf Seite 309
- Kapitel 77, “Streuungsdiagramm” auf Seite 619

Festlegen erweiterter Optionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Erweitert**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Statistik für Balkenberechnung	gibt an, wie die Länge des Balkens berechnet werden soll. Welche Statistiken verfügbar sind, hängt von der Grafik ab, die Sie erstellen. ■ Wenn Sie der Rolle **Summe von** in der Anwendungsroutine **Balkendiagramm** oder der Rolle **Balkensumme von** in der Anwendungsroutine **Balken-Liniendiagramm** keine Variable zugewiesen haben, sind die Statistiken Häufigkeit, kumulative Häufigkeit, Prozentwert und kumulativer Prozentwert verfügbar. ■ Wenn Sie der Rolle **Summe von** oder **Balkensumme von** eine Variable zugewiesen haben oder wenn Sie Gruppe mit mehreren Maßen als Typ für das Balkendiagramm zugewiesen haben, sind die Statistiken Summe und Durchschnitt verfügbar.
Verwendete Statistik bei Liniendarstellung	legt die Statistik fest, die verwendet wird, wenn eine Linie in einem Balken-Liniendiagramm gezeichnet wird. Die Statistiken Häufigkeit, Kumulative Häufigkeit, Prozentwert und Kumulativer Prozentwert sind verfügbar. Wenn Sie der Rolle **Liniensumme von** in der Anwendungsroutine **Balken-Linien-Diagramm** eine Variable zugewiesen haben, ist diese Option im Fenster **Erweitert** nicht verfügbar. Stattdessen steht sie im **Plots**-Fenster zur Verfügung, sodass Sie die Statistik für jede Linie steuern können, wenn der Rolle **Liniensumme von** mehr als eine Variable zugewiesen ist.
Fehlerbalken anzeigen	zeigt die Konfidenzintervalle auf den Balken an. Wählen Sie aus der Dropdown-Liste **Typ** eine der folgenden Optionen aus: ■ **Balken** stellt Fehlerbalken als Balken mit der halben Breite der Hauptbalken dar. ■ **Beides** stellt Fehlerbalken als zwei Markierungen dar, die durch eine Linie verbunden sind. ■ **Oben** stellt Fehlerbalken als Markierung für die obere Konfidenzgrenze dar, die mit der Oberseite des Balkens durch eine Linie verbunden wird.

Optionsname	Beschreibung
	Wählen Sie eine Farbe für die Fehlerbalken aus, und geben Sie im Feld **CLM** ein Konfidenzniveau an. **Hinweis:** Sie können Fehlerbalken nur anzeigen, wenn folgendes erfüllt ist: ■ Sie haben **Durchschnitt** oder **Prozentwert** als Statistik ausgewählt, die für die Berechnung der Balkenlänge verwendet werden soll ■ Sie erstellen kein dreidimensionales gruppiertes Balkendiagramm ■ Sie haben der Rolle **Stapel** keine Variable zugewiesen
Fehlende Werte akzeptieren	lässt fehlende Werte als gültige Werte für die Diagrammspalte zu.
Nullwerte auslassen	unterdrückt alle Werte der Diagrammspalte mit einer Balkenlänge von Null.
Prozentwerte und kumulative Prozentwerte für jede Gruppe berechnen (G100)	berechnet Prozentsätze und kumulative Prozentsätze separat für jede Gruppe in gruppierten Balkendiagrammen. Wenn Sie diese Option verwenden, stellt jeder einzelne Prozentwert den Anteil des Midpoints an der Gruppe dar. Die Summe der Werte ergibt für jede Gruppe 100 Prozent. **Hinweis:** Diese Option ist nur verfügbar, wenn Sie der Rolle **Balken gruppieren nach** für gruppierte Balkendiagramme eine Variable zugewiesen haben.
Statistiken neben Balken anzeigen	zeigt den Wert der Statistik an, die zum Berechnen des Balkens verwendet wird. **Hinweis:** Diese Option ist nur für horizontale Balkendiagramme verfügbar.
Einen Statistikwert zum Anzeigen für Balken angeben	zeigt den Wert einer Statistik an, die Sie aus der Dropdown-Liste auswählen. Der Statistikwert kann mit dem Wert, mit dem die Balkenlänge berechnet wird, übereinstimmen, oder aber ein anderer Wert sein. Wenn Sie ein zweidimensionales Balkendiagramm erstellen, können Sie den Wert der Statistik in den Balken anzeigen.

Siehe auch

- Kapitel 6, “Balkendiagramm” auf Seite 39
- Kapitel 8, “Balken-Linien-Diagramm” auf Seite 51

- Kapitel 19, “Konturdiagramm” auf Seite 137

Anhang 2

Allgemeine Darstellungsoptionen für Regelkarten-Anwendungsroutinen

Festlegen von Achsenoptionen 759
Hinzufügen von Referenzlinien 760
Einfügungen hinzufügen 760
Festlegen weiterer Plotoptionen 760
Festlegen von Blockvariablenoptionen 762

Festlegen von Achsenoptionen

Klicken Sie im Auswahlfenster unter **Darstellung** ⇨ **Achsen** auf **Achsen**, um auf die Optionen zuzugreifen.

Folgende Optionen können angegeben werden:

- die Farbe der Achsen
- die Breite der Achsen
- die Farbe der Achsenmarkierungen
- ein Etikett für die horizontalen und vertikalen Achsen

Hinzufügen von Referenzlinien

Klicken Sie im Auswahlfenster unter **Darstellung** ⇨ **Achsen** auf **Horizontal** bzw. **Vertikal**, um auf die Optionen zuzugreifen.

Um Referenzlinien zu verwenden, wählen Sie **Referenzlinien verwenden**. Sie können die Referenzlinie formatieren und die Position jeder Referenzlinie festlegen.

Einfügungen hinzufügen

Klicken Sie im Auswahlfenster unter der Überschrift **Darstellung** auf **Einfügung**, um auf diese Optionen zuzugreifen.

Sie können einen Plot erweitern, indem Sie direkt in der Grafik ein Feld oder eine Tabelle (eine so genannte Einfügung) der beschreibenden Statistiken hinzufügen. Wenn Sie ein Vergleichshistogramm erstellen, erscheint in jedem Komponentenhistogramm eine Einfügung.

Wählen Sie **Einfügung einschließen**, um eine beschreibende Statistik zum Plot hinzuzufügen. Wählen Sie dann die Statistiken, die Sie in die Einfügung einschließen möchten.

Hinweis: Die Statistiken für die Berechnung der Prozessfähigkeitsindizes sind nur verfügbar, wenn Sie Spezifikationsgrenzen festgelegt und dann der Rolle **Analysevariablen** zugewiesen haben. Sie können Spezifikationsgrenzen für Histogramm, CDF-Plots, Wahrscheinlichkeitsplots und Q-Q-Plots erstellen. Weitere Informationen zur Berechnung der Prozessfähigkeitsindizes finden Sie in der Hilfe zur Prozedur CAPABILITY.

Festlegen weiterer Plotoptionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Optionen**, um auf diese Optionen zuzugreifen.

Optionsname	Beschreibung
Innerhalb der Regelgrenzen, Regelgrenzen, Verbinden der Liniensegmente, Segmente außerhalb Grenzen, Bereiche außerhalb der Grenzen und Rahmen	legt die Farben dieser Diagrammelemente fest. ■ **Innerhalb der Regelgrenzen** gibt die Farbe innerhalb der unteren und oberen Regelgrenzen an. ■ **Regelgrenzen** gibt die Farbe der Regelgrenzen und der Mittellinie sowie der Etiketten für diese Linien an. Diese Option gibt außerdem die Farbe der Etiketten für diese Linien an. ■ **Verbinden der Liniensegmente** gibt die Farbe der Liniensegmente an, mit denen die Punkte auf der Regelkarte verbunden werden. ■ **Segmente außerhalb der Grenzen** gibt die Farbe der Darstellungssymbole und die Anteile der Verbindungsliniensegmente an, die außerhalb der Regelgrenzen liegen. Diese Option eignet sich zum Hervorheben der Punkte, die außerhalb der Regelgrenzen liegen. ■ **Bereiche außerhalb der Grenzen** gibt die Füllfarbe für die Flächen außerhalb der Regelgrenzen an, die zwischen den verbundenen Punkten und den Regelgrenzen liegen und durch Verbindungslinien begrenzt sind. Diese Option eignet sich zum Hervorheben der Punkte, die außerhalb der Regelgrenzen liegen. ■ **Rahmen** gibt die Farben zum Füllen des Rechtecks an, das durch die Achsen und den Rahmen begrenzt ist.
Symbol	legt das Symbol für die Datenpunkte fest.
Anzahl der angezeigten Seiten	gibt die Anzahl der Seiten an, die zum Anzeigen der Regelkarte verwendet werden.
Moving Range	gibt die Zahl der aufeinanderfolgenden Maße an, aus denen die Moving Ranges berechnet werden. Der angegebene Wert sollte zwischen 2 und 10 (inklusive) liegen. Der Standardwert beträgt 2. Ein Bereich von 2 bis 10 ist für die Anwendungsroutine üblich. Wenn Sie die Prozedur SHEWHART in SAS verwenden, können Sie einen höheren Wert angeben. Weitere Informationen finden Sie unter 'The SHEWHART Procedure' im *Base SAS Procedures Guide*. **Hinweis:** Diese Option ist nur für die Anwendungsroutine **Urwertkarte** verfügbar.
Legende für Untergruppenstichprobe unterdrücken	unterdrückt die Standardlegende für Untergruppenstichproben. Diese Option ist nützlich, wenn die Stichprobengrößen der Untergruppe konstant und mit der Stichprobengröße der Regelgrenze identisch sind, da die

Optionsname	Beschreibung
	Stichprobengröße der Regelgrenze automatisch in der oberen rechten Ecke der Karte angezeigt wird.
Numerischen Wert für Untergruppe verwenden (statt Variable)	gibt einen konstanten Wert für die Größe an. Verwenden Sie diese Option, wenn die Stichprobengröße der Untergruppe fix ist. Geben Sie den gewünschten Wert im Feld **Stichprobengröße der Untergruppe** ein. Der Wert für die Stichprobengröße der Untergruppe muss zwischen dem maximalen Wert der Prozessvariablen und 99999 (einschließlich) liegen. **Hinweis:** Wenn Sie der Rolle **Stichprobengröße der Untergruppe** eine Variable zuweisen, wird diese Option deaktiviert, und die dieser Rolle zugewiesene Variable wird für die Stichprobengröße der Untergruppe verwendet. **Hinweis:** Diese Option ist nur für die Anwendungsroutinen np-, p- und u-Regelkarte verfügbar.
Kein sekundäres Diagramm	unterdrückt die Anzeige eines sekundären Diagramms. Üblicherweise verwenden Sie diese Option mit einer Urwertkarte, um ein Diagramm für Einzelmessungen zu erstellen und das begleitende Diagramm für gleitende Spannweiten zu unterdrücken. **Hinweis:** Diese Option ist nur für die Anwendungsroutinen Urwertkarte, X-quer- und R-Regelkarte und die X-quer- und S-Regelkarte verfügbar.
Typ des Randverteilungsdiagramms	gibt den Typ des Randverteilungsdiagramms an, das in der Urwertkarte erscheint. Ein Randverteilungsdiagramm ist ein univariater Plot der Regelkartenstatistiken. Sie können einen Digidot-Plot, ein Histogramm oder einen Box-and-Whisker-Plot auswählen. Der Plot erscheint entweder links oder rechts von der Urwertkarte. Die Standardauswahl ist **Kein Randverteilungsdiagramm**. **Hinweis:** Diese Option ist nur für die Anwendungsroutine **Urwertkarte** verfügbar.

Festlegen von Blockvariablenoptionen

Klicken Sie im Auswahlfenster unter der Überschrift 'Darstellung' auf **Blockoptionen**, um auf diese Optionen zuzugreifen.

Hinweis: Diese Optionen sind nur verfügbar, wenn Sie der Rolle **Blockvariablen** eine Variable zugewiesen haben.

Optionsname	Beschreibung
Blocketikett	gibt die Etiketten für die Blockvariablen an.
Legendenposition	gibt die vertikale Position der Legende für die Werte der Blockvariablen an.
Lange Werte	geben Sie an, was mit langen Werten von Blockvariablen geschieht, wenn in der Blocklegende nicht genügend Platz ist, um diese anzuzeigen. Standardmäßig werden lange Werte nicht angezeigt. Wenn Sie die Option **Einheitlich in Höhe verringern** auswählen, wird die Höhe der Werte gleichmäßig verringert, sodass die Werte in die Legende passen. Wenn Sie die Option **Rechts abschneiden** auswählen, werden lange Werte auf der rechten Seite abgeschnitten, bis diese in die Legende passen.
Etikettenposition	gibt die Position eines Blockvariablenetiketts in der Blocklegende an. Die verfügbaren Optionen hängen von der Position der Legende ab. Folgende Optionen können ausgewählt werden: ■ **Oben** platziert das Etikett direkt über der Legende. Hierbei handelt es sich um die Standardeinstellung. ■ **Links** platziert das Etikett links neben der Legende ■ **Rechts** platziert das Etikett rechts neben der Legende. Diese Option ist nicht verfügbar, wenn Sie ausgewählt haben, die Legende über der Regelkarte zu positionieren. Sie sollten **Links** bzw. **Rechts** als Etikettenposition nur bei Etiketten verwenden, die kurz genug sind, dass sie in die Seitenränder des Diagramms passen. Zu lange Etiketten werden abgeschnitten.
Farbe des Blocketiketts	gibt die Farbe des Textes auf dem Blocketikett an. **Hinweis:** Diese Option ist nur verfügbar, wenn als Etikettenposition **Oben** ausgewählt wurde.
Wiederholte Werte in Legende anzeigen	zeigt die Blockvariablenwerte für alle Untergruppen an. Standardmäßig wird nur der erste Blockvariablenwert in einem Block angezeigt. Sich wiederholende Blockvariablenwerte werden nicht angezeigt.

Anhang 3

Erstellen eines Modells

Allgemeines zu Modellen ... **765**
Siehe auch ... 765
Angeben von Haupteffekten ... **766**
Festlegen von Kreuzeffekten (Interaktionen) ... **766**
Angeben von Verschachtelungseffekten ... **766**
Erstellen mehrerer Interaktionen für zwei oder mehr Variablen ... **767**
Erstellen polynomialer Interaktionen quantitativer Variablen ... **767**
Entfernen der Effekte der Konstante aus der Analyse ... **767**

Allgemeines zu Modellen

Klicken Sie im Auswahlfenster auf **Modell**, um auf diese Optionen zuzugreifen.

Standardmäßig sind keine Effekte angegeben. Deshalb passt diese Anwendungsroutine ein ausschließlich aus Konstanten bestehendes Modell an. Um einen Effekt anzugeben, müssen Sie der Rolle **Quantitative Variablen** oder **Klassifizierungsvariablen** mindestens eine Variable zuweisen. Sie können Kombinationen der Variablen auswählen, um Kreuz-, Verschachtelungs-, Faktor- oder Polynomeffekte zu erzeugen.

Siehe auch

- Kapitel 37, “High-Performance - Lineare Regression” auf Seite 257
- Kapitel 38, “High-Performance - Logistische Regression” auf Seite 269
- Kapitel 48, “Logistische Regression” auf Seite 357

Angeben von Haupteffekten

1. Wählen Sie den Variablennamen im Feld **Klassifizierungs- und quantitative Variablen**.
2. Klicken Sie auf **Haupt**, um die Variable zum Feld **Effekte** hinzuzufügen.

Festlegen von Kreuzeffekten (Interaktionen)

1. Wählen Sie mindestens einen Variablennamen aus. Um mehr als eine Variable auszuwählen, halten Sie die Strg-Taste gedrückt und wählen die zusätzlichen Variablennamen aus.
2. Klicken Sie auf **Kreuz**. Wenn Sie nur eine Variable auswählen, wird die Variable mit sich selbst gekreuzt.

Angeben von Verschachtelungseffekten

1. Wählen Sie einen oder mehrere Haupteffekte bzw. Kreuzeffekte aus, die Sie verschachteln möchten.
2. Wählen Sie mindestens eine Klassifizierungsvariable aus, innerhalb derer Sie die Haupt- bzw. Kreuzeffekte verschachteln möchten.
3. Klicken Sie auf **Verschachteln**. Der Verschachtelungseffekt wird im Feld **Effekte** in der Form *Variable 1*(*Variable 2*) angezeigt.

Erstellen mehrerer Interaktionen für zwei oder mehr Variablen

1. Geben Sie Interaktionen höheren Grades an, indem Sie die Zahl im Feld **Grade** hochsetzen.
2. Klicken Sie auf **Faktoriell**, um dem Feld **Effekte** die faktoriellen Effekte hinzuzufügen. Standardmäßig werden Kombinationen mit zwei Variablen verwendet.

Erstellen polynomialer Interaktionen quantitativer Variablen

1. Geben Sie Interaktionen höheren Grades an, indem Sie die Zahl im Feld **Grade** hochsetzen.
2. Klicken Sie auf **Polynomial**, um dem Feld **Effekte** die Polynomialeffekte hinzuzufügen.

Entfernen der Effekte der Konstante aus der Analyse

Deaktivieren Sie das Kontrollkästchen **Inkl. Konstante**. Wenn Sie das Häkchen im Kontrollkästchen entfernen, müssen Sie mindestens einen Effekt angegeben haben. Standardmäßig wird die Konstante eingeschlossen.

Anhang 4

Integrierte Anwendungsroutinen

Integrierte SAS-Anwendungsroutinen und dazugehörige SAS-Prozeduren und -Lizenzen 769

Integrierte SAS-Anwendungsroutinen und dazugehörige SAS-Prozeduren und -Lizenzen

In der folgenden Tabelle erhalten Sie eine Übersicht über die in SAS Enterprise Guide und SAS Add-In for Microsoft Office integrierten Anwendungsroutinen. Für jede Anwendungsroutine wird die jeweilige primäre SAS-Prozedur sowie die erforderliche SAS Software-Lizenz aufgeführt, die zum Ausführen der Anwendungsroutine benötigt wird.

Hinweis: Je nach verwendeter Version des SAS Add-In for Microsoft Offices oder des SAS Enterprise Guide stehen mglw. nicht alle der nachfolgend aufgeführten Anwendungsroutinen zur Verfügung. Die Verfügbarkeit einer Anwendungsroutine hängt davon ab, ob die erforderliche SAS Software lizenziert und in Ihrer Site installiert ist. Wenden Sie sich an Ihren Systemadministrator, wenn Sie weitere Informationen benötigen.

Sofern in Ihrer Seite SAS Studio lizenziert und installiert wurde, stehen darüber hinaus noch weitere Anwendungsroutinen (Tasks) zur Verfügung. Weitere Informationen finden Sie in SAS Studio-Tasks und zugehörige SAS-Prozeduren und -Lizenzen.

Integrierte Anwendungsroutine	Verwendete SAS-Prozeduren	Erforderliche SAS-Lizenzen
Tabelle anhängen	SQL	Base SAS
Flächendiagramm	GPLOT	Base SAS, SAS/GRAPH
ARIMA Modellerstellung und Prognose	ARIMA	Base SAS, SAS/ETS
Projektbibliothek zuweisen	LIBNAME	Base SAS
Balkendiagramm	GCHART	Base SAS, SAS/GRAPH
Assistent für Balkendiagramme auf Seite 47	GCHART	Base SAS, SAS/GRAPH
Balken-Linien-Diagramm	GBARLINE	Base SAS, SAS/GRAPH
Einfache Prognosen	FORECAST	Base SAS, SAS/ETS
Box-Regelkarte	SHEWHART	Base SAS, SAS/QC
Box-Plot	GPLOT	Base SAS, SAS/GRAPH
Blasendiagramm	GPLOT	Base SAS, SAS/GRAPH
c-Regelkarte auf Seite 97	SHEWHART	Base SAS, SAS/QC
Kanonische Korrelation	CANCORR	Base SAS, SAS/STAT
CDF-Plots	CAPABILITY	Base SAS, SAS/QC
Daten charakterisieren	CONTENTS, UNIVARIATE, FREQ	Base SAS
Clusteranalyse	CLUSTER, FASTCLUS, TREE	Base SAS, SAS/STAT
Daten vergleichen	COMPARE	Base SAS
Konturdiagramm	GCONTOUR, G3GRID	Base SAS, SAS/GRAPH

Integrierte Anwendungsroutine	Verwendete SAS-Prozeduren	Erforderliche SAS-Lizenzen
Korrelationen	CORR	Base SAS
Format erstellen	FORMAT	Base SAS
Format aus Datei erstellen auf Seite 163	FORMAT	Base SAS
Feature-Tabelle für Karten erstellen auf Seite 167	GPROJECT	Base SAS, SAS/GRAPH
Zeitreihendaten erstellen	TIMESERIES	Base SAS, SAS/ETS
Dateiattribute	DATASETS	Base SAS
Dateien und Formate löschen	SQL, CATALOG	Base SAS
Diskriminanzanalyse	DISCRIM	Base SAS, SAS/STAT
Verteilungsanalyse	UNIVARIATE	Base SAS
Ringdiagramm	GCHART	Base SAS, SAS/GRAPH
Dateien auf PC herunterladen	SQL	Base SAS
Faktoranalyse	FACTOR	Base SAS, SAS/STAT
Filtern und sortieren	SQL	Base SAS
Forecast Studio - Projekt erstellen	Ohne	SAS Forecast Server
Forecast Studio - Projekt öffnen	Ohne	SAS Forecast Server
Forecast Studio - Projekt überschreiben auf Seite 241	Ohne	SAS Forecast Server
Verallgemeinerte lineare Modelle	GENMOD	Base SAS, SAS/STAT
High-Performance - Lineare Regression	HPREG	Base SAS, SAS High Performance Analysis Server, SAS/STAT
High-Performance - Logistische Regression	HPLOGISTIC	Base SAS, SAS High Performance Analysis Server, SAS/STAT

Integrierte Anwendungsroutine	Verwendete SAS-Prozeduren	Erforderliche SAS-Lizenzen
Histogramme auf Seite 281	CAPABILITY	Base SAS, SAS/QC
Daten importieren (Microsoft Excel, Textdateien, Microsoft Access)*	DATA-Schritt, IMPORT	Base SAS, SAS/ACCESS to PC File Formats
JMP-Datei importieren	DATA-Schritt, IMPORT	Base SAS, SAS/ACCESS to PC File Formats
SPSS-Datei importieren	DATA-Schritt, IMPORT	Base SAS, SAS/ACCESS to PC File Formats
Stata-Datei importieren*	DATA-Schritt, IMPORT	Base SAS, SAS/ACCESS to PC File Formats
Urwertkarte	SHEWHART	Base SAS, SAS/QC
JMP Stored Process Packager*	DATA-Schritt	JMP
Sterbetafeln	LIFETEST	Base SAS, SAS/STAT
Liniendiagramm	GPLOT	Base SAS, SAS/GRAPH
Assistent für Liniendiagramme	GPLOT	Base SAS, SAS/GRAPH
Lineare Modelle	GLM	Base SAS, SAS/STAT
Lineare Regression	REG	Base SAS, SAS/STAT
Listenbericht	PRINT	Base SAS
Listenberichtassistent auf Seite 353	REPORT	Base SAS
Logistische Regression	LOGISTIC	Base SAS, SAS/STAT
Kartendiagramm	GMAP	Base SAS, SAS/GRAPH
X-quer- und R-Regelkarte	SHEWHART	Base SAS, SAS/QC
X-quer- und R-Regelkarte	SHEWHART	Base SAS, SAS/QC

Integrierte Anwendungsroutine	Verwendete SAS-Prozeduren	Erforderliche SAS-Lizenzen
Gemischte Modelle	MIXED	Base SAS, SAS/STAT
Modell-Scoring	Ohne	SAS Enterprise Miner
Nichtlineare Regression	NLIN	Base SAS, SAS/STAT
Nichtparametrische einfache ANOVA	NPAR1WAY	Base SAS, SAS/STAT
np-Regelkarte	SHEWHART	Base SAS, SAS/QC
Einfache ANOVA	ANOVA	Base SAS, SAS/STAT
Einfache Häufigkeiten	FREQ	Base SAS
SAS ODS Statistical Graphs	Ohne	ODS Graphics Designer
p-Regelkarte	SHEWHART	Base SAS, SAS/QC
Pareto-Diagramm	PARETO	Base SAS, SAS/QC
Kreisdiagramm	GCHART	Base SAS, SAS/GRAPH
Assistent für Kreisdiagramm	GCHART	Base SAS, SAS/GRAPH
P-P-Plots	CAPABILITY	Base SAS, SAS/QC
Zeitreihendaten vorbereiten	EXPAND	Base SAS, SAS/ETS
Hauptkomponenten	PRINCOMP	Base SAS, SAS/STAT
Wahrscheinlichkeitsplots	CAPABILITY	Base SAS, SAS/QC
Proportionale Hazards	PHREG	Base SAS, SAS/STAT
Q-Q-Plots	CAPABILITY	Base SAS, SAS/QC
Abfrage erstellen (Query Builder)*	SQL	Base SAS
Radardiagramm	GRADAR	Base SAS, SAS/GRAPH
Zufällige Stichprobe	SURVEYSELECT	Base SAS, SAS/STAT

Integrierte Anwendungsroutine	Verwendete SAS-Prozeduren	Erforderliche SAS-Lizenzen
Rangzuweisung für Daten	RANK	Base SAS
Rapid Predictive Modeler	SAS Enterprise Miner-Makros	SAS Enterprise Miner
RFM-Analyse (Recency, Frequency, and Monetary)	SAS-Makro	Base SAS
Regressionsanalyse von Paneldaten	TSCSREG	Base SAS, SAS/ETS
Regressionsanalyse mit autoregressiven Fehlern	AUTOREG	Base SAS, SAS/ETS
Streuungsdiagramm	GPLOT, G3D	Base SAS, SAS/GRAPH
Streuungsdiagramm-Matrix	SGSCATTER	Base SAS, SAS/GRAPH
ODS Statistical Graph anzeigen	SGDESIGN	Base SAS
Daten sortieren	SORT	Base SAS
Spalten teilen	TRANSPOSE	Base SAS
Spalten stapeln	TRANSPOSE	Base SAS
Daten standardisieren	STANDARD	Base SAS
Beschreibende Statistiken	MEANS	Base SAS
Assistent für beschreibende Statistiken	MEANS	Base SAS
Zusammenfassungstabellen auf Seite 669	TRANSPOSE	Base SAS
Assistent für Zusammenfassungstabellen	TRANSPOSE	Base SAS
Oberflächenplot	G3D	Base SAS, SAS/GRAPH
t-Test	TTEST	Base SAS, SAS/STAT

Integrierte Anwendungsroutine	Verwendete SAS-Prozeduren	Erforderliche SAS-Lizenzen
Tabellenanalyse	FREQ	Base SAS
Kacheldiagramm	GTILE	Base SAS, SAS/GRAPH
Transponieren	TRANSPOSE	Base SAS
u-Regelkarte	SHEWHART	Base SAS, SAS/QC
Bibliotheksmetadaten aktualisieren*	METALIB	Base SAS
Dateien auf Server hochladen*	SQL	Base SAS

Hinweis: Die mit einem Sternchen gekennzeichneten Anwendungsroutinen stehen nur in SAS Enterprise Guide zur Verfügung.