

Tâches SAS[®] dans SAS[®] Enterprise Guide[®] 8.4 et SAS[®] Add-In 8.4 for Microsoft Office

8.4*

* Ce document peut s'appliquer à des versions supplémentaires du logiciel. Ouvrez ce document dans [SAS Help Center](#) et cliquez sur la version dans la bannière pour voir toutes les versions disponibles.

La référence bibliographique correcte pour ce manuel est la suivante : SAS Institute Inc. 2024. *Tâches SAS® dans SAS® Enterprise Guide® 8.4 et SAS® Add-In 8.4 for Microsoft Office*. Cary, NC: SAS Institute Inc.

Tâches SAS® dans SAS® Enterprise Guide® 8.4 et 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.

Pour une version papier du manuel : 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.

Pour un téléchargement web ou livre électronique : 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

Table des matières

Chapitre 1 / Travailler avec les tâches	1
Tâches intégrées et tâches SAS Studio	1
Filtrer vos données de tâche	3
Ajouter le code utilisateur dans une tâche	5
Attribuer des propriétés à une colonne	5
Sélectionner un format de sortie pour les graphiques	6
Utiliser des modèles de tâches	6
Chapitre 2 / Tâche Ajouter une table	11
À propos de la tâche Ajouter une table	11
Ajouter une table : Créer la liste des tables d'entrée	12
Ajouter une table : enregistrer les résultats	12
Chapitre 3 / Aire	15
À propos de la tâche Aire	15
Aire : Sélectionner un type de graphe	16
Aire : Attribuer des variables aux fonctions d'analyse	16
Aire : Définir les options d'apparence	17
Indiquer des titres et des notes de bas de page	18
Afficher les propriétés	19
Chapitre 4 / Tâche Modélisation et prévision ARIMA	21
À propos de la tâche Modélisation et prévision ARIMA	21
Modélisation et prévision ARIMA : Attribuer des variables aux fonctions d'analyse	22
Modélisation et prévision ARIMA : Définir les options d'identification	23
Modélisation et prévision ARIMA : Définir les options d'estimation	25
Modélisation et prévision ARIMA : Définir les options de prévision	27
Indiquer des titres et des notes de bas de page	29
Afficher les propriétés	30
Chapitre 5 / Affecter une bibliothèque au projet	31
À propos de l'Assistant Affecter une bibliothèque de projets	31
Affecter une bibliothèque de projets : Indiquez un nom et un serveur	32
Affecter une bibliothèque de projets : Spécifier un moteur	32
Affecter une bibliothèque de projets : Spécifier les options de configuration	35
Affecter une bibliothèque de projets : Vérifier les informations récapitulatives	36
Chapitre 6 / Histogramme	37
À propos de la tâche Histogramme	37
Histogramme : Sélectionner un type de graphique	38
Histogramme : Attribuer des variables aux fonctions d'analyse	39
Histogramme : Définir les options d'apparence	40
Indiquer des titres et des notes de bas de page	41

Afficher les propriétés	42
Chapitre 7 / Assistant Histogramme	43
À propos de l'assistant Histogramme	43
Assistant Histogramme : Sélectionner les données	43
Assistant Histogramme : Attribuer des variables aux fonctions	44
Assistant Histogramme : Définir les options d'apparence	45
Assistant Histogramme : Indiquer les titres et les notes de bas de page	46
Chapitre 8 / Courbe-histogramme	47
À propos de la tâche Courbe-Histogramme	47
Courbe-histogramme : Attribuer des variables aux fonctions d'analyse	48
Courbe-histogramme : Définir les options d'apparence	49
Indiquer des titres et des notes de bas de page	50
Afficher les propriétés	51
Chapitre 9 / Prévion de base	53
À propos de la tâche Prévion de base	53
Prévion de base : Attribuer des variables aux fonctions d'analyse	54
Prévion de base : Définir les options de prévion	55
Prévion de base : Générer des graphes	57
Prévion de base : Définition des options des résultats	58
Indiquer des titres et des notes de bas de page	59
Afficher les propriétés	59
Chapitre 10 / Carte de la moyenne avec superposition des boîtes de Tukey	61
À propos de la carte de la moyenne avec superposition des boîtes de Tuckey	61
Cartes de contrôle : Attribuer des variables aux fonctions d'analyse	62
Cartes de contrôle : Indiquer des limites de contrôle	64
Carte de contrôle : Sélectionner les tests à effectuer	65
Cartes de contrôle : Définir les options d'apparence du graphe	66
Cartes de contrôle : Enregistrer les résultats d'analyse	68
Indiquer des titres et des notes de bas de page	69
Afficher les propriétés	69
Chapitre 11 / Boîte à moustaches	71
À propos de la tâche Boîte à moustaches	71
Boîte à moustaches : Sélectionner un type de graphe	72
Boîte à moustaches : Attribuer des variables aux fonctions d'analyse	72
Boîte à moustaches : Définir les options d'apparence	73
Indiquer des titres et des notes de bas de page	76
Afficher les propriétés	76
Chapitre 12 / Graphique à bulles	79
À propos de la tâche Bulle	79
Graphique à bulles : Attribuer des variables aux fonctions d'analyse	80
Graphique à bulles : Définir les options d'apparence	81
Indiquer des titres et des notes de bas de page	82
Afficher les propriétés	82
Chapitre 13 / Corrélation canonique	85
À propos de la tâche Corrélation canonique	85

Corrélation canonique : Attribuer des variables aux fonctions d'analyse	86
Corrélation canonique : Définir les options statistiques	87
Corrélation canonique : Générer des graphes	88
Corrélation canonique : Définition des options des résultats	89
Indiquer des titres et des notes de bas de page	90
Afficher les propriétés	90
Chapitre 14 / Carte c	93
À propos de la tâche Carte c	93
Carte c : Attribuer des variables aux fonctions d'analyse	94
Cartes de contrôle : Indiquer des limites de contrôle	95
Carte de contrôle : Sélectionner les tests à effectuer	96
Cartes de contrôle : Définir les options d'apparence du graphe	98
Cartes de contrôle : Enregistrer les résultats d'analyse	98
Indiquer des titres et des notes de bas de page	99
Afficher les propriétés	100
Chapitre 15 / Courbes CDF	101
À propos de la tâche Courbes CDF	101
Analyse de capacité : Attribuer des variables aux fonctions d'analyse	102
Courbes CDF : Sélectionner la distribution	103
Courbes CDF : Personnaliser la distribution	104
Analyse de capacité : Définir les options d'apparence du graphe	105
Analyse de capacité : Spécifier les types d'analyse	109
Indiquer des titres et des notes de bas de page	110
Afficher les propriétés	110
Chapitre 16 / Caractéristiques des données	113
À propos de l'Assistant Caractéristiques des données	113
Caractéristiques des données : Sélectionner les données	114
Caractéristiques des données : Sélectionner les options du rapport	114
Caractéristiques des données : Indiquer des limites pour les sorties	115
Chapitre 17 / Analyse de classification	117
À propos de la tâche Analyse de classification	117
Analyse de classification : Attribuer des variables aux fonctions d'analyse	118
Analyse de classification : Définir les options de classification	119
Analyse de classification : Générer des courbes et des diagrammes	120
Analyse de classification : Définition des options des résultats	122
Indiquer des titres et des notes de bas de page	124
Afficher les propriétés	124
Chapitre 18 / Comparer les données	127
À propos de l'Assistant Comparer les données	127
Comparer les données : Sélectionner les données	128
Comparer les données : Sélectionner les variables à comparer	128
Comparer les données : Spécifier les critères de comparaison	129
Comparer les données : Spécifier les options de sortie	129
Chapitre 19 / Courbe de niveau	133
À propos de la tâche Courbe de niveau	133
Courbe de niveau : Sélectionner un type de graphe	134

Courbe de niveau : Attribuer des variables aux fonctions d'analyse	135
Courbe de niveau : Définir les options d'apparence	136
Indiquer des titres et des notes de bas de page	137
Afficher les propriétés	138
Chapitre 20 / Copier des fichiers	139
À propos de la tâche Copier des fichiers	139
Copier des fichiers de l'ordinateur local vers un SAS Server	140
Chapitre 21 / Corrélations	143
À propos de la tâche Corrélations	143
Corrélation : Attribuer des variables aux fonctions d'analyse	144
Corrélation : Définir les options de corrélation	145
Corrélation : Définition des options des résultats	147
Corrélations : Créer la table de sortie	148
Indiquer des titres et des notes de bas de page	148
Afficher les propriétés	149
Chapitre 22 / Créer un format	151
À propos de la tâche Créer un format	151
Créer un format : Définir les options d'un format	152
Créer un format : Définir le nouveau format	153
Afficher les propriétés	156
Chapitre 23 / Créer un format à partir d'une table	159
À propos de la tâche Créer un format à partir d'une table	159
Créer un format à partir d'une table	161
Chapitre 24 / Créer une table de caractéristiques cartographiques	163
À propos de la tâche Créer une table de caractéristiques cartographiques	163
Créer une table de caractéristiques cartographiques : Attribuer des variables aux fonctions d'analyse	164
Créer une table de caractéristiques cartographiques : Définir les options	165
Créer une table de caractéristiques cartographiques : Spécifier les résultats	167
Indiquer des titres et des notes de bas de page	167
Afficher les propriétés	167
Chapitre 25 / Assistant Créer les données d'une série chronologique	169
À propos de l'Assistant Créer les données d'une série chronologique	169
Assistant Créer les données d'une série chronologique : Sélectionner les données	170
Assistant Créer les données d'une série chronologique : Attribuez des variables aux fonctions	170
Assistant Créer les données d'une série chronologique : Spécifiez des options pour les valeurs de cumul, manquantes et zéro.	171
Assistant Créer les données d'une série chronologique : Sélectionner une autre sortie	172
Assistant Créer les données d'une série chronologique : Indiquez le titre et les notes de bas de page	172
Chapitre 26 / Attributs de la table	175
À propos de l'Assistant Attributs de table	175

Attributs de la table : Sélectionner les données	176
Attributs de la table : Spécifier la sortie	176
Chapitre 27 / Supprimer les tables et les formats	179
À propos de la tâche Supprimer les tables et les formats	179
Supprimer des tables, des vues et des formats personnalisés	180
Afficher les propriétés	180
Chapitre 28 / Analyse discriminante	183
À propos de la tâche d'analyse discriminante	183
Analyse discriminante : Attribuer des variables aux fonctions d'analyse	184
Analyse discriminante : Définir les options de l'analyse	185
Analyse discriminante : Classifier les nouvelles données	187
Analyse discriminante : Définition des options des résultats	189
Indiquer des titres et des notes de bas de page	190
Afficher les propriétés	190
Chapitre 29 / Analyse de distribution	193
À propos de la tâche d'analyse de distribution	193
Analyse de distribution : Attribuer des variables aux fonctions d'analyse	194
Analyse de distribution : Sélectionner les distributions	195
Analyse de distribution : Personnaliser la distribution	196
Analyse de distribution : Générer des graphes	197
Analyse de distribution : Ajouter un encart	199
Analyse de distribution : Indiquer les options des résultats	199
Indiquer des titres et des notes de bas de page	200
Afficher les propriétés	200
Chapitre 30 / Anneau	203
À propos de la tâche Anneau	203
Anneau : Sélectionner un type de graphique	204
Anneau : Attribuer des variables aux fonctions d'analyse	204
Anneau : Définir les options d'apparence	205
Indiquer des titres et des notes de bas de page	208
Afficher les propriétés	208
Chapitre 31 / Télécharger des fichiers de données vers un PC	211
À propos de la tâche Télécharger des fichiers de données vers un PC	211
Télécharger des fichiers de données vers un PC : Sélectionner les tables SAS à télécharger	212
Télécharger des fichiers de données vers un PC : Définir les options de téléchargement	212
Chapitre 32 / Analyse factorielle	215
À propos de la tâche Analyse factorielle	215
Analyse factorielle : Attribuer des variables aux fonctions d'analyse	216
Analyse factorielle : Définir les options de la méthode de factorisation	217
Analyse factorielle : Définir les options des facteurs communs	219
Analyse factorielle : Définir les options de rotation et de graphe	221
Analyse factorielle : Définition des options des résultats	223
Indiquer des titres et des notes de bas de page	225

Chapitre 33 / Créer un projet Forecast Studio	227
À propos de l'Assistant Créer un projet Forecast Studio	227
Créer un projet Forecast Studio : Vérifier les données	228
Créer un projet Forecast Studio : Sélectionner les variables de prévision	228
Créer un projet Forecast Studio : Sélectionner les variables de classification	229
Créer un projet Forecast Studio : Enregistrer le projet	231
Chapitre 34 / Ouvrir un projet Forecast Studio	233
À propos de l'Assistant Ouvrir un projet Forecast Studio	233
Ouvrir un projet Forecast Studio : Sélectionner le projet de prévision	234
Ouvrir un projet Forecast Studio : Sélectionner la série de prévisions	234
Ouvrir un projet Forecast Studio : Sélectionner les types de résultat	235
Ouvrir un projet Forecast Studio : Indiquer les titres et les notes de bas de page	235
Chapitre 35 / Projet de valeur corrigée Forecast Studio	237
À propos de l'Assistant Projet de valeur corrigée Forecast Studio	237
Projet de valeur corrigée Forecast Studio : Sélectionner le projet de prévisions	238
Projet de valeur corrigée Forecast Studio : Sélectionner la série de prévisions et soumettre les valeurs corrigées	238
Chapitre 36 / Modèle linéaire généralisé	243
À propos de la tâche Modèles linéaires généralisés	243
Modèles linéaires généralisés : Attribuer des variables aux fonctions d'analyse	244
Modèles linéaires généralisés : Créer un modèle	245
Modèles linéaires généralisés : Définir les options du modèle	246
Modèles linéaires généralisés : Sélectionner les tests post hoc	248
Modèles linéaires généralisés : Générer des graphes	248
Modèles linéaires généralisés : Définir les options de prédiction	249
Indiquer des titres et des notes de bas de page	250
Afficher les propriétés	251
Chapitre 37 / Régression linéaire High-Performance	253
À propos de la tâche Régression linéaire High-Performance	253
Régression linéaire High-Performance : Attribuer des variables aux fonctions d'analyse	254
Régression linéaire High-Performance : Créer un modèle	256
Régression linéaire High-Performance : Définir les options du modèle	259
Régression linéaire High-Performance : Définir les options de la partition	260
Régression High-Performance : Définir les options de performances	261
Régression High-Performance : Définir les options de prédiction	261
Indiquer des titres et des notes de bas de page	262
Chapitre 38 / Régression logistique High-Performance	265
À propos de la tâche Régression logistique High-Performance	265
Régression logistique High-Performance : Attribuer des variables aux fonctions d'analyse	266
Régression logistique High-Performance : Créer un modèle	267
Régression logistique High-Performance : Définir les options du modèle	271
Régression High-Performance : Définir les options de performances	273
Régression High-Performance : Définir les options de prédiction	273
Indiquer des titres et des notes de bas de page	274

Chapitre 39 / Histogrammes	277
À propos de la tâche Histogramme	277
Analyse de capacité : Attribuer des variables aux fonctions d'analyse	278
Histogrammes : Sélectionner la distribution	279
Histogrammes : Personnaliser la distribution	281
Analyse de capacité : Définir les options d'apparence du graphe	282
Analyse de capacité : Spécifier les types d'analyse	286
Indiquer des titres et des notes de bas de page	287
Afficher les propriétés	287
Chapitre 40 / Carte aux mesures individuelles	289
À propos de la tâche Carte aux mesures individuelles	289
Carte aux mesures individuelles : Attribuer des variables aux fonctions d'analyse	290
Cartes de contrôle : Indiquer des limites de contrôle	291
Carte de contrôle : Sélectionner les tests à effectuer	293
Cartes de contrôle : Définir les options d'apparence du graphe	294
Cartes de contrôle : Enregistrer les résultats d'analyse	294
Indiquer des titres et des notes de bas de page	295
Afficher les propriétés	296
Chapitre 41 / Tables de survie	297
À propos de la tâche Tables de survie	297
Tables de survie : Attribuer des variables aux fonctions d'analyse	298
Tables de survie : Définir les options de la méthode	300
Tables de survie : Générer des graphes	301
Tables de survie : Définition des options des résultats	302
Indiquer des titres et des notes de bas de page	303
Afficher les propriétés	304
Chapitre 42 / Courbe	305
À propos de la tâche Courbe	305
Courbe : Sélectionner un type de graphe	306
Courbe : Attribuer des variables aux fonctions d'analyse	307
Courbe : Définir les options d'apparence	309
Indiquer des titres et des notes de bas de page	310
Afficher les propriétés	311
Chapitre 43 / Assistant Courbe	313
À propos de l'assistant Courbe	313
Assistant Courbe : Sélectionner les données	313
Assistant Courbe : Attribuer des variables aux fonctions	314
Assistant Courbe : Définir les options de la courbe	314
Assistant Courbe : Définir les options d'apparence	316
Assistant Courbe : Indiquer les titres et les notes de bas de page	316
Chapitre 44 / Modèles linéaires	317
À propos de la tâche Modèles linéaires	317
Modèles linéaires : Attribuer des variables aux fonctions d'analyse	318
Modèles linéaires : Créer un modèle	319
Modèles linéaires : Définir les options du modèle	319
Modèles linéaires : Définir les options avancées	320

Modèles linéaires : Sélectionner les tests post hoc	322
Modèles linéaires : Générer des graphes	326
Modèles linéaires : Définir les options de prédiction	327
Indiquer des titres et des notes de bas de page	328
Afficher les propriétés	328
Chapitre 45 / Régression linéaire	331
À propos de la tâche Régression linéaire	331
Régression linéaire : Attribuer des variables aux fonctions d'analyse	332
Régression linéaire : Définir les options du modèle	333
Régression linéaire : Définir les options statistiques	336
Régression linéaire : Générer des graphes	337
Régression linéaire : Définir les options de prédiction	339
Indiquer des titres et des notes de bas de page	340
Afficher les propriétés	341
Chapitre 46 / Liste des données	343
À propos de la tâche Liste des données	343
Liste des données : Attribuer des variables aux fonctions d'analyse	344
Liste des données : Définir les options de liste	345
Indiquer des titres et des notes de bas de page	347
Afficher les propriétés	347
Chapitre 47 / Assistant Rapport sous forme de liste	349
À propos de l'Assistant Rapport sous forme de liste	349
Assistant Rapport sous forme de liste : Sélectionner les données	350
Assistant Rapport sous forme de liste : Définir la liste	350
Assistant Rapport sous forme de liste : Afficher les totaux	351
Assistant Rapport sous forme de liste : Indiquer les titres et les notes de bas de page	352
Chapitre 48 / Régression logistique	353
À propos de la tâche Régression logistique	353
Régression logistique : Attribuer des variables aux fonctions d'analyse	354
Régression logistique : Créer un modèle	355
Régression logistique : Définir les options du modèle	359
Régression logistique : Générer des graphes	361
Régression logistique : Définir les options de prédiction	363
Indiquer des titres et des notes de bas de page	364
Afficher les propriétés	365
Chapitre 49 / Carte	367
À propos de la tâche Carte	367
Carte : Sélectionner le type de carte	368
Carte : Attribuer des variables aux fonctions d'analyse	369
Carte : Définir les options d'apparence	370
Indiquer des titres et des notes de bas de page	372
Carte : Afficher un aperçu des résultats	373
Afficher les propriétés	373
Chapitre 50 / Carte de la moyenne et de l'étendue	375
À propos de la Carte de la moyenne et de l'étendue	375

Carte de la moyenne et de l'étendue : Attribuer des variables aux fonctions d'analyse	376
Cartes de contrôle : Indiquer des limites de contrôle	377
Carte de contrôle : Sélectionner les tests à effectuer	379
Cartes de contrôle : Définir les options d'apparence du graphe	380
Cartes de contrôle : Enregistrer les résultats d'analyse	380
Indiquer des titres et des notes de bas de page	381
Afficher les propriétés	382
Chapitre 51 / Carte de la moyenne et de l'écart-type	383
À propos de la Carte de la moyenne et de l'écart-type	383
Carte de la moyenne et de l'écart-type : Attribuer des variables aux fonctions d'analyse	384
Cartes de contrôle : Indiquer des limites de contrôle	385
Carte de contrôle : Sélectionner les tests à effectuer	387
Cartes de contrôle : Définir les options d'apparence du graphe	388
Cartes de contrôle : Enregistrer les résultats d'analyse	388
Indiquer des titres et des notes de bas de page	389
Afficher les propriétés	390
Chapitre 52 / Modèles mixtes	391
À propos de la tâche Modèles mixtes	391
Modèles mixtes : Attribuer des variables aux fonctions d'analyse	392
Modèles mixtes : Créer un modèle d'effets fixes	393
Modèles mixtes : Définir les options du modèle d'effets fixes	393
Modèles mixtes : Définir des effets aléatoires	395
Modèles mixtes : Définir des effets répétitifs	397
Modèles mixtes : Spécifier les moyennes des moindres carrés	399
Modèles mixtes : Générer des graphes	400
Modèles mixtes : Définir les options de prédiction	402
Indiquer des titres et des notes de bas de page	402
Afficher les propriétés	403
Chapitre 53 / Scoring du modèle	405
À propos de l'Assistant Scoring du modèle	405
Scoring du modèle : Sélectionner les données	406
Scoring du modèle : Sélectionner un modèle de scoring	406
Scoring du modèle : Mettre en correspondance les variables de données avec les champs obligatoires du modèle.	407
Scoring du modèle : Sélectionner les colonnes des données de sortie.	407
Scoring du modèle : Enregistrer les données de sortie	408
Scoring du modèle : Confirmer vos sélections	408
Scoring du modèle : Interpréter les résultats	409
Chapitre 54 / Régression non linéaire	413
À propos de la tâche Régression non linéaire	413
Régression non linéaire : Attribuer des variables aux fonctions d'analyse	414
Régression non linéaire : Définir les options du modèle	415
Régression non linéaire : Définir les options de calcul	415
Régression non linéaire : Générer des graphes	417
Régression non linéaire : Définition des options des résultats	418
Indiquer des titres et des notes de bas de page	419

Afficher les propriétés	419
Chapitre 55 / Analyse de variance non-paramétrique à une dimension	421
À propos de la tâche Analyse de variance non paramétrique à une dimension ...	421
Analyse de variance non-paramétrique à une dimension : Attribuer des variables aux fonctions d'analyse	422
Analyse de variance non-paramétrique à une dimension : Sélectionner les analyses	423
Analyse de variance non-paramétrique à une dimension : Calculer les p-values exactes	424
Analyse de variance non-paramétrique à une dimension : Définition des options des résultats	424
Indiquer des titres et des notes de bas de page	425
Afficher les propriétés	426
Chapitre 56 / Carte np	427
À propos de la tâche Carte np	427
Carte np : Attribuer des variables aux fonctions d'analyse	428
Cartes de contrôle : Indiquer des limites de contrôle	429
Carte de contrôle : Sélectionner les tests à effectuer	431
Cartes de contrôle : Définir les options d'apparence du graphe	432
Cartes de contrôle : Enregistrer les résultats d'analyse	432
Indiquer des titres et des notes de bas de page	433
Afficher les propriétés	434
Chapitre 57 / Analyse de variance à une dimension	435
À propos de la tâche Analyse de variance à une dimension	435
Analyse de variance à une dimension : Attribuer des variables aux fonctions d'analyse	436
Analyse de variance à une dimension : sélection des tests	437
Analyse de variance à une dimension : Comparer les moyennes de groupe	438
Analyse de variance à une dimension : Sélectionner des statistiques pour les variables quantitatives	439
Analyse de variance à une dimension : Générer des graphes	440
Analyse de variance à une dimension : Définition des options des résultats	441
Indiquer des titres et des notes de bas de page	442
Afficher les propriétés	442
Chapitre 58 / Fréquences à un critère de classification	445
À propos de la tâche Fréquences à un critère de classification	445
Fréquences à un critère de classification : Attribuer des variables aux fonctions d'analyse	446
Fréquences à un critère de classification : Définir les options statistiques	447
Fréquences à un critère de classification : Générer des graphes	448
Fréquences à un critère de classification : Définition des options des résultats	448
Indiquer des titres et des notes de bas de page	450
Afficher les propriétés	450
Chapitre 59 / Carte p	453
À propos de la Carte p	453
Carte p : Attribuer des variables aux fonctions d'analyse	454
Cartes de contrôle : Indiquer des limites de contrôle	455

Carte de contrôle : Sélectionner les tests à effectuer	457
Cartes de contrôle : Définir les options d'apparence du graphe	458
Cartes de contrôle : Enregistrer les résultats d'analyse	458
Indiquer des titres et des notes de bas de page	459
Afficher les propriétés	460
Chapitre 60 / Graphique de Pareto	461
À propos de la tâche Graphique de Pareto	461
Graphique de Pareto : Attribuer des variables aux fonctions d'analyse	462
Graphique de Pareto : Limiter le nombre de catégories	463
Graphique de Pareto : Définir les options d'apparence du graphe	465
Indiquer des titres et des notes de bas de page	468
Afficher les propriétés	469
Chapitre 61 / Diagramme circulaire	471
À propos de la tâche Diagramme circulaire	471
Diagramme circulaire : Sélectionner un type de graphique	472
Diagramme circulaire : Attribuer des variables aux fonctions d'analyse	473
Diagramme circulaire : Définir les options d'apparence	474
Indiquer des titres et des notes de bas de page	476
Afficher les propriétés	477
Chapitre 62 / Assistant Diagramme circulaire	479
À propos de l'assistant Diagramme circulaire	479
Assistant Diagramme circulaire : Sélectionner les données	479
Assistant Diagramme circulaire : Attribuer des variables aux fonctions	480
Assistant Diagramme circulaire : Définir les options d'apparence	481
Assistant Diagramme circulaire : Indiquer les titres et les notes de bas de page	481
Chapitre 63 / Courbes P-P	483
À propos de la tâche Courbes P-P	483
Analyse de capacité : Attribuer des variables aux fonctions d'analyse	484
Courbes P-P : Sélectionner la distribution	486
Courbes P-P : Personnaliser la distribution	487
Analyse de capacité : Définir les options d'apparence du graphe	487
Analyse de capacité : Spécifier les types d'analyse	491
Indiquer des titres et des notes de bas de page	492
Afficher les propriétés	492
Chapitre 64 / Préparer les données d'une série chronologique	495
À propos de la tâche Préparer les données d'une série chronologique	495
Préparer les données d'une série chronologique : Attribuer des variables aux fonctions d'analyse	496
Préparer les données d'une série chronologique : Indiquer la fréquence de sortie	497
Préparer les données d'une série chronologique : Définir les caractéristiques d'une série chronologique	498
Préparer les données d'une série chronologique : Indiquer des transformations pour chaque variable d'une série chronologique	500
Préparer les données d'une série chronologique : Définition des options des résultats	501
Afficher les propriétés	502
Créer des lignes d'observations pour les lignes manquantes de la table	502

Convertir les données d'une série chronologique d'une fréquence d'échantillonnage à une autre	503
Modifier les caractéristiques des observations d'une table d'entrée	504
Convertir des séries apériodiques en estimations périodiques	505
Interpoler les valeurs manquantes	506
Appliquer des transformations génériques aux données	507
Chapitre 65 / Composantes principales	509
À propos de la tâche Composantes principales	509
Composantes principales : Attribuer des variables aux fonctions d'analyse	510
Composantes principales : Définir les options de l'analyse	511
Composantes principales : Générer des graphes	512
Composantes principales : Définition des options des résultats	513
Indiquer des titres et des notes de bas de page	514
Afficher les propriétés	515
Chapitre 66 / Courbes de probabilité	517
À propos de la tâche Courbes de probabilité	517
Analyse de capacité : Attribuer des variables aux fonctions d'analyse	518
Courbes de probabilité : Sélectionner la distribution	520
Courbes de probabilité : Personnaliser la distribution	521
Analyse de capacité : Définir les options d'apparence du graphe	521
Analyse de capacité : Spécifier les types d'analyse	525
Indiquer des titres et des notes de bas de page	526
Afficher les propriétés	526
Chapitre 67 / Hasards proportionnels	529
À propos de la tâche Hasards proportionnels	529
Hasards proportionnels : Attribuer des variables aux fonctions d'analyse	530
Hasards proportionnels : Définir les options du modèle	532
Hasards proportionnels : Définir les options de méthode	534
Hasards proportionnels : Générer des graphes	535
Hasards proportionnels : Définition des options des résultats	536
Indiquer des titres et des notes de bas de page	537
Afficher les propriétés	538
Chapitre 68 / Courbes Q-Q	539
À propos de la tâche Courbes Q-Q	539
Analyse de capacité : Attribuer des variables aux fonctions d'analyse	540
Courbes Q-Q : Sélectionner la distribution	542
Courbes Q-Q : Personnaliser la distribution	543
Analyse de capacité : Définir les options d'apparence du graphe	543
Analyse de capacité : Spécifier les types d'analyse	547
Indiquer des titres et des notes de bas de page	548
Afficher les propriétés	548
Chapitre 69 / Carte radar	551
À propos de la tâche Carte radar	551
Carte radar : Attribuer des variables aux fonctions d'analyse	552
Carte radar : Définir les options d'apparence	553
Indiquer des titres et des notes de bas de page	554
Afficher les propriétés	554

Chapitre 70 / Échantillon aléatoire	557
À propos de la tâche Échantillon aléatoire	557
Échantillon aléatoire : Créer un échantillon aléatoire	558
Chapitre 71 / Classer les données	561
À propos de la tâche Classer les données	561
Classer les données : Attribuer des variables aux fonctions d'analyse	562
Classer les données : Définir les options de l'analyse	562
Rang : Définition des options des résultats	564
Afficher les propriétés	565
Chapitre 72 / Analyse RFM (Récence, Fréquence et Montant)	567
À propos de la tâche Analyse RFM (Récence, Fréquence et Montant)	567
Analyse RFM (Récence, Fréquence et Montant) : Sélectionner le type des données à analyser	568
Analyse RFM (Récence, Fréquence et Montant) : Attribuer des variables aux fonctions d'analyse	569
Analyse RFM (Récence, Fréquence et Montant) : Regrouper les données dans des classes	570
Analyse RFM (Récence, Fréquence et Montant) : Enregistrer les résultats	571
Analyse RFM (Récence, Fréquence et Montant) : Sélectionner des graphes	572
Chapitre 73 / Analyse de la régression avec erreurs autorégressives	573
À propos de la tâche Analyse de la régression avec erreurs autorégressives	573
Analyse de la régression avec erreurs autorégressives : Attribuer des variables aux fonctions d'analyse	574
Analyse de la régression avec erreurs autorégressives : Définir les options du modèle	575
Analyse de la régression avec erreurs autorégressives : Définir les options statistiques	576
Analyse de la régression avec erreurs autorégressives : Générer des graphes	579
Analyse de la régression avec erreurs autorégressives : Définition des options des résultats	580
Indiquer des titres et des notes de bas de page	582
Afficher les propriétés	582
Chapitre 74 / Analyse de régression des données du panel	585
À propos de la tâche Analyse de régression des données du panel	585
Analyse de régression des données du panel : Attribuer des variables aux fonctions d'analyse	586
Analyse de régression des données du panel : Définir les options du modèle	587
Analyse de régression des données du panel : Définition des options des résultats	588
Indiquer des titres et des notes de bas de page	589
Afficher les propriétés	589
Chapitre 75 / Graphiques statistiques SAS ODS	591
À propos des Graphiques statistiques SAS ODS	591
Utiliser ODS Graphics Designer	592
Afficher le graphique statistique ODS	592

Chapitre 76 / SAS Rapid Predictive Modeler	595
À propos de SAS Rapid Predictive Modeler	595
Stratégies d'échantillonnage de SAS Rapid Predictive Modeler	596
SAS Rapid Predictive Modeler : Organiser les données	597
SAS Rapid Predictive Modeler : Attribuer des variables aux fonctions d'analyse	599
SAS Rapid Predictive Modeler : Sélectionner un modèle	600
SAS Rapid Predictive Modeler : Définir les options du rapport	605
SAS Rapid Predictive Modeler : Définir les options de sortie	606
SAS Rapid Predictive Modeler : Enregistrer un modèle	607
Chapitre 77 / Nuage de points	609
À propos de la tâche Nuage de points	609
Nuage de points : Sélectionner un type de graphe	610
Nuage de points : Attribuer des variables aux fonctions d'analyse	610
Nuage de points : Définir les options d'apparence	612
Indiquer des titres et des notes de bas de page	614
Afficher les propriétés	614
Chapitre 78 / Matrice nuage de points	617
À propos de la tâche Matrice du nuage de points	617
Matrice nuage de points : Créer une matrice de nuages de points	618
Chapitre 79 / Trier les données	621
À propos de la tâche Trier les données	621
Trier les données : Attribuer des variables aux fonctions d'analyse	622
Trier les données : Définir les options de tri des données	622
Trier les données : Définition des options des résultats	624
Afficher les propriétés	625
Chapitre 80 / Fractionner des colonnes	627
À propos de la tâche Fractionner des colonnes	627
Fractionner des colonnes : Attribuer des variables aux fonctions d'analyse	629
Fractionner des colonnes : Définition des options des résultats	630
Afficher les propriétés	630
Chapitre 81 / Empiler colonnes	633
À propos de la tâche Empiler des colonnes	633
Empiler des colonnes : Attribuer des variables aux fonctions d'analyse	635
Empiler des colonnes : Définir les options des noms de colonne	636
Empiler des colonnes : Définition des options des résultats	637
Afficher les propriétés	637
Chapitre 82 / Normaliser les données	639
À propos de la tâche Normaliser les données	639
Normaliser les données : Attribuer des variables aux fonctions d'analyse	640
Normaliser les données : Définir les options de normalisation	641
Normaliser les données : Définition des options des résultats	641
Afficher les propriétés	642
Chapitre 83 / Statistiques descriptives	643
À propos de la tâche Statistiques descriptives	643
Statistiques descriptives : Attribuer des variables aux fonctions d'analyse	644

Statistiques descriptives : Sélectionner des statistiques descriptives	645
Statistiques descriptives : Définir les options de tracé	647
Statistiques descriptives : Définition des options des résultats	648
Indiquer des titres et des notes de bas de page	650
Afficher les propriétés	651
Chapitre 84 / Assistant Statistiques descriptives	653
À propos de l'assistant Statistiques descriptives	653
Assistant Statistiques descriptives : Sélectionner les données	653
Assistant Statistiques descriptives : Attribuer des variables aux fonctions	654
Assistant Statistiques descriptives : Indiquer les statistiques du rapport et les résultats de la sortie	655
Assistant Statistiques descriptives : Indiquez le titre et les notes de bas de page	656
Chapitre 85 / Tables de synthèse	659
À propos de la tâche Tables de synthèse	659
Tables de synthèse : Attribuer des variables aux fonctions d'analyse	660
Tables de synthèse : Définir la table de synthèse	661
Tables de synthèse : Définition des options des résultats	664
Indiquer des titres et des notes de bas de page	664
Afficher les propriétés	665
Chapitre 86 / Assistant Tables de synthèse	667
À propos de l'assistant Tables de synthèse	667
Assistant Tables de synthèse : Sélectionner les données	667
Assistant Tables de synthèse : Attribuer des variables aux fonctions	668
Assistant Tables de synthèse : Sélectionner les variables de classification	669
Assistant Tables de synthèse : Spécifier les totaux	669
Assistant Tables de synthèse : Sélectionner une autre sortie	670
Assistant Tables de synthèse : Indiquez le titre et les notes de bas de page	670
Chapitre 87 / Surface	671
À propos de la tâche Surface	671
Surface : Sélectionner un type de graphe	672
Surface : Attribuer des variables aux fonctions d'analyse	673
Surface : Définir les options d'apparence	674
Indiquer des titres et des notes de bas de page	675
Afficher les propriétés	676
Chapitre 88 / Test t	677
À propos de la Tâche Test t	677
Test t : Choisir le type du test t	678
Test t : Attribuer des variables aux fonctions d'analyse	678
Test t : Définir les options de l'analyse	680
Test t : Générer des graphes	680
Indiquer des titres et des notes de bas de page	681
Afficher les propriétés	682
Chapitre 89 / Analyse d'une table	683
À propos de la tâche Analyse d'une table	683
Analyse d'une table : Attribuer des variables aux fonctions d'analyse	684
Analyse d'une table : Définir des tables	685

Analyse d'une table : Sélectionner les statistiques des cellules	685
Analyse d'une table : Spécifier les statistiques des tables	687
Analyse d'une table : Définition des options des résultats	692
Indiquer des titres et des notes de bas de page	693
Afficher les propriétés	694
Chapitre 90 / Diagramme en mosaïque	695
À propos de la tâche Diagramme en mosaïque	695
Diagramme en mosaïque : Attribuer des variables aux fonctions d'analyse	696
Diagramme en mosaïque : Définir les options d'apparence	697
Diagramme en mosaïque : Définir les options avancées	699
Indiquer des titres et des notes de bas de page	701
Chapitre 91 / Transposer	703
À propos de la tâche Transposer	703
Transposer : Attribuer des variables aux fonctions d'analyse	704
Transposer : Définir les options de l'analyse	705
Transposer : Définition des options des résultats	706
Afficher les propriétés	706
Chapitre 92 / Carte u	709
À propos de la Carte u	709
Carte u : Attribuer des variables aux fonctions d'analyse	710
Cartes de contrôle : Indiquer des limites de contrôle	711
Carte de contrôle : Sélectionner les tests à effectuer	713
Cartes de contrôle : Définir les options d'apparence du graphe	714
Cartes de contrôle : Enregistrer les résultats d'analyse	714
Indiquer des titres et des notes de bas de page	715
Afficher les propriétés	716
Chapitre 93 / Actualiser les métadonnées de la bibliothèque	717
À propos de l'Assistant Actualiser les métadonnées de la bibliothèque	717
Actualiser les métadonnées de la bibliothèque : Sélectionnez le SAS Server et la bibliothèque	718
Actualiser les métadonnées de la bibliothèque : Exécuter un rapport sur les métadonnées ou actualisez ces dernières.	718
Chapitre 94 / Téléchargement vers CAS	721
À propos de la tâche Téléchargement vers CAS	721
Télécharger des données vers un SAS CAS Analytic Server	722
Chapitre 95 / Télécharger des fichiers de données vers un serveur	723
À propos de la tâche Télécharger des fichiers de données vers un serveur	723
Télécharger des fichiers de données vers un serveur : Sélectionner les fichiers de données SAS à télécharger	724
Télécharger des fichiers de données vers un serveur : Spécifier les options de téléchargement	724
Chapitre 96 / Téléchargement vers LASR	727
À propos de la tâche Téléchargement vers LASR	727
Télécharger des données vers un SAS LASR Analytic Server	728

Annexe 1 / Options d'apparence courantes pour les tâches graphiques	729
Définir les options d'un axe	729
Ajouter des lignes de référence	732
Définir les options de légende	733
Personnaliser l'aire	734
Définir l'apparence des barres	735
Définir les options des barres	736
Définir les options des courbes	737
Sélectionner la méthode d'interpolation	738
Définir les options avancées	739
Annexe 2 / Options d'apparence courantes pour les tâches de la carte de contrôle	743
Définir les options d'un axe	743
Ajouter des lignes de référence	744
Ajouter un encart	744
Définir d'autres options de tracé	744
Définir les options de variable de bloc	746
Annexe 3 / Créer un modèle	749
À propos des modèles	749
Indiquer les effets principaux	750
Spécifier les effets croisés (interactions)	750
Spécifier des effets imbriqués	750
Créer plusieurs croisements de deux variables sélectionnées ou plus	751
Créer des croisements polynomiaux de variables quantitatives	751
Supprimer les effets de l'intercept dans l'analyse	751
Annexe 4 / Tâches intégrées	753
Tâches intégrées et Procédures et licences SAS associées	753

Travailler avec les tâches

<i>Tâches intégrées et tâches SAS Studio</i>	1
Voir aussi	3
<i>Filtrer vos données de tâche</i>	3
<i>Ajouter le code utilisateur dans une tâche</i>	5
<i>Attribuer des propriétés à une colonne</i>	5
<i>Sélectionner un format de sortie pour les graphiques</i>	6
<i>Utiliser des modèles de tâches</i>	6
À propos des modèles de tâches	6
Créer un modèle de tâche	7
Modifier un modèle de tâche	8
Exécuter un modèle de tâche	8
Importer un modèle de tâche	8
Exporter un modèle de tâche	9

Tâches intégrées et tâches SAS Studio

SAS Enterprise Guide et SAS Add-In for Microsoft Office sont livrés avec de nombreuses tâches SAS intégrées. Vous pouvez accéder à ces tâches dans le volet **Tâches**. Cette liste de tâches peut également inclure des tâches personnalisées et des tâches SAS Studio.

Pour accéder aux tâches de SAS Studio dans SAS Enterprise Guide ou SAS Add-In for Microsoft Office :

- 1 Assurez-vous que SAS Studio est bien installé sur votre site.
- 2 Assurez-vous que le serveur par défaut pour le profil actif est bien SAS Studio server.
- 3 Dans les paramètres de SAS Enterprise Guide ou SAS Add-In for Microsoft Office, activez l'option **Afficher les tâches SAS Studio de votre version de SAS Studio**.

- Si vous exécutez SAS Studio 3.4 ou version ultérieure, vous avez accès soit à SAS Studio Enterprise Edition soit à SAS Studio Single User Edition. SAS Studio Single User Edition doit être installé sur la même machine que SAS Enterprise Guide et SAS Add-In for Microsoft Office.
- Dans SAS Enterprise Guide, vous pouvez également accéder aux tâches de SAS Studio 5.2. Spécifiez l'URL et les informations d'identification pour votre version de SAS Studio.

Remarque : les tâches SAS Studio 5.2 ne sont pas disponibles dans SAS Add-in for Microsoft Office.

À présent, toutes les tâches de SAS Studio sont disponibles dans le volet **Tâches** et la fenêtre **Tâches** de SAS Enterprise Guide et SAS Add-In for Microsoft Office. Pour ne voir que les tâches de **SAS Studio** dans le volet **Tâches** et la Galerie de tâches, sélectionnez **SAS Studio** comme type de tâche.

L'interface pour les tâches de SAS Studio est légèrement différente des boîtes de dialogue utilisées pour les tâches intégrées dans SAS Enterprise Guide et SAS Add-In for Microsoft Office. Par exemple, voici l'interface de la tâche Simulation de lancement de pièces.

Simulation de lancer de pièces

OPTIONS **INFORMATIONS**

▼ **OBSERVATIONS**

Simule le lancer d'un nombre précis de pièces. Les résultats indiquent la fréquence et le pourcentage des fois où la pièce tombe sur Pile en fonction d'un nombre indiqué de lancers.

*Nombre de pièces : 10

*Nombre de lancers : 1 000 000

▼ **SORTIE**

*Nom de la table : WORK.COINS

☐ Afficher la table des graphiques

► Options graphiques

Aperçu du code Exécuter Enregistrer Annuler

La disponibilité de certaines tâches de SAS Studio dépend des licences sur votre site. Par exemple, de nombreuses tâches de SAS Studio permettent de travailler dans un environnement SAS Viya. Néanmoins, si vous ne disposez pas d'une licence et d'une installation de SAS Viya, ces tâches ne sont pas disponibles.

Voir aussi

- [Annexe 4, “Tâches intégrées” page 753](#)
- [“SAS Studio Tasks and Associated Procedures and Licenses” dans SAS Studio: Task Reference Guide](#)

Filtrer vos données de tâche

Vous pouvez spécifier les filtres à appliquer à la source de données en cours au moyen des listes déroulantes de la boîte de dialogue **Modifier les données et le filtre**. Après avoir créé un filtre, vous pouvez utiliser la macro-variable `_CLIENTTASKFILTER` dans les titres et notes de bas de page de la tâche pour afficher les informations de filtre dans le titre ou la note de bas de page des résultats.

Remarque : si vous utilisez SAS Add-In for Microsoft Office, les filtres que vous avez définis dans la fenêtre **Modifier la source de données** sont appliqués en premier. Ensuite, l'Add-in SAS utilise les filtres que vous créez au moyen de la boîte de dialogue **Modifier les données et le filtre**.

Pour créer un filtre pour des données :

- 1 Dans le volet Données de la tâche ou l'étape Vérifier les données de l'assistant, cliquez sur **Modifier**. La boîte de dialogue **Modifier les données et le filtre** apparaît.
- 2 (Facultatif) Dans SAS Enterprise Guide, sélectionnez la source de données en entrée. Lorsque vous ouvrez une tâche, celle-ci utilise la source de données active comme source de données en entrée par défaut. Le nom de la source de données active apparaît dans la liste déroulante Source de données.
- 3 Dans la première liste déroulante, sélectionnez la variable à filtrer. Cette liste contient toutes les variables de vos données.

Remarque : par défaut, cette liste déroulante présente les noms des variables. Pour qu'elle affiche leur libellé, activez la case à cocher **Afficher libellés à la place des noms de variables**.

- 4 Dans la deuxième liste déroulante, sélectionnez un opérateur.
- 5 Dans la troisième zone, indiquez la valeur de la variable en procédant de l'une des façons suivantes :
 - Saisissez la valeur de la variable dans le champ de saisie.

- Cliquez sur le bouton sélecteur de valeurs pour afficher la liste des valeurs disponibles.

Si vous avez sélectionné l'opérateur **Est compris dans** ou **N'est pas compris dans**, vous devez saisir des valeurs de début et de fin. Si vous avez sélectionné l'opérateur **Dans une liste** ou **Pas dans une liste**, vous devez indiquer la liste des valeurs pour la comparaison. Vous pouvez saisir ou coller des valeurs dans la zone. Pour sélectionner une valeur de la source de données, cliquez sur **Ajouter une valeur**. Chaque valeur doit figurer sur une nouvelle ligne.

Pour tous les autres opérateurs, les 25 premières valeurs distinctes de la source de données s'affichent dans la liste. Pour visualiser davantage de valeurs, cliquez sur **Autres valeurs**. Chaque fois que vous cliquez sur ce bouton, les 25 valeurs distinctes suivantes de la variable s'affichent en plus des 25 valeurs initiales.

Remarque : si la source de données SAS se trouve dans le système de fichiers de votre ordinateur ou sur un serveur accessible depuis votre ordinateur, la liste de valeurs est alimentée dans l'ordre où les valeurs uniques apparaissent au sein des données. Si, par contre, la source de données ne se trouve pas sur le système de fichiers de votre machine ni sur un serveur accessible depuis celle-ci, la liste des valeurs est alors alimentée par ordre croissant.

- 6 (Facultatif) Dans la dernière liste déroulante, sélectionnez **AND** ou **OR** pour appliquer d'autres filtres aux données. Répétez les étapes 3 à 5 pour créer ces filtres.

Remarque : si vous avez spécifié plusieurs filtres, ceux joints par une relation basée sur un ET logique sont évalués en premier. En effet, les filtres ne sont pas évalués dans l'ordre où vous les créez. Par exemple, supposons que vous ayez créé le filtre suivant : **Sex = 'M' or Name = 'Carol' and Date = '1980'**.

Dans cet exemple, SAS commence par éliminer par filtrage toutes les observations pour lesquelles Nom est égal à Carole et Date est égal à 1980. Ensuite, la relation basée sur un opérateur OR est évaluée.

- 7 Cliquez sur **OK** pour enregistrer vos modifications. Le filtre de la source de données en entrée est désormais visible dans le volet Données de la tâche.

Ajouter le code utilisateur dans une tâche

Lorsque vous visualisez le code d'une tâche, vous avez la possibilité d'insérer votre propre code dans celui que SAS génère pour la tâche. Vous pouvez insérer les options de procédure ou instructions du code SAS dans le code généré avec toutes les fonctionnalités de l'Éditeur de programmes. En revanche, vous ne pouvez pas modifier ni supprimer le code automatiquement généré dans la fenêtre **Aperçu du code de la tâche**; vous pouvez uniquement y insérer un nouveau code. Pour insérer votre propre code dans la fenêtre de tâche, cliquez sur **Aperçu du code**.

- 1 Dans la fenêtre **Aperçu du code de la tâche**, activez la case **Afficher les points d'insertion du code personnalisé**. Les lignes suivantes sont ajoutées dans le code de la tâche à chaque endroit où vous pouvez insérer du code personnalisé :

```
/* Début du code utilisateur personnalisé (Framework_BeforeTaskCode) */
<insérer le code personnalisé ici>
/* Fin du code utilisateur personnalisé (Framework_BeforeTaskCode) */
```

- 2 Cliquez à l'emplacement où vous voulez insérer du code et saisissez-le.
- 3 Pour enregistrer vos modifications, fermez la fenêtre **Aperçu du code de la tâche**.

Attribuer des propriétés à une colonne

Dans la liste de la fenêtre de tâche, cliquez sur **Données** pour exécuter cette tâche.

Faites un clic droit sur la variable et sélectionnez **Propriétés** pour attribuer de façon temporaire un libellé descriptif ainsi qu'un format. Les propriétés que vous attribuez ne sont appliquées que sur la tâche en cours.

Pour attribuer un libellé, saisissez un texte dans la zone **Libellé**. Le libellé ne peut contenir plus de 256 caractères. La liste déroulante contient la valeur du libellé associé de façon permanente à la variable.

La plupart des tâches permettent l'attribution de formats. Pour attribuer un format :

- 1 Cliquez sur **Modifier**. La fenêtre **Formats** apparaît.
- 2 Sélectionnez la catégorie du format dans la zone **Catégories**. Vous pouvez utiliser un format standard SAS ou bien un format personnalisé, que vous avez préalablement défini dans la Tâche Créer un format.

- 3 Sélectionnez le nom du format dans la zone **Formats**.
- 4 Indiquez les changements des attributs dans la zone **Attributs**, puis cliquez sur **OK**.

Remarque : dans la fenêtre **Propriétés**, la propriété **Tri** est définie à **Oui** si la table a été triée en fonction de cette variable.

Sélectionner un format de sortie pour les graphiques

Dans SAS Add-In for Microsoft Office, vous avez le choix entre les formats de sortie graphiques suivants : ActiveX, image ActiveX, GIF, JPEG, PNG et image Java. Dans SAS Enterprise Guide, vous pouvez également sélectionner les formats Java et SAS EMF.

Les formats de sortie ActiveX et Java sont interactifs. Vous pouvez faire un clic droit sur n'importe quelle sortie graphique générée avec ces formats de sortie et modifier certaines options. Les options que vous pouvez modifier ne sont pas les mêmes pour les formats de sortie ActiveX et Java. Les modifications d'option que vous effectuez n'ont d'incidence que sur la sortie, et non sur les sélections effectuées dans la fenêtre de la tâche lors de la création du graphique.

Les formats de sortie graphique Image ActiveX, GIF, JPEG, PNG, Image Java et SAS EMF ne sont pas interactifs. Vous ne pouvez pas changer l'apparence du graphique après l'avoir généré.

Utiliser des modèles de tâches

À propos des modèles de tâches

Lorsque vous exécutez plusieurs fois une tâche telle que la tâche Histogramme, vous souhaitez généralement utiliser les mêmes paramètres. Plutôt que d'ouvrir la tâche à chaque fois afin de définir ces paramètres pour différentes sources de données en entrée, vous pouvez enregistrer les paramètres dans un modèle de tâche. Vous pouvez ensuite exécuter ce modèle avec n'importe quelle source de données en entrée.

Remarque : tenez compte des limitations suivantes lorsque vous utilisez les modèles de tâches :

- Les valeurs ou paramètres qui dépendent de la source de données ne sont pas enregistrées dans le modèle de tâche. Par exemple, lorsque vous attribuez une variable à une fonction spécifique, cette affectation n'est pas enregistrée dans le modèle. Si une option est dépendante des valeurs de la source de données, le paramètre de l'option n'est pas enregistré dans le modèle. Par exemple, dans la tâche Courbe, les définitions de symboles qui dépendent des valeurs de données ne sont pas enregistrées dans le modèle de tâche.
- Vous ne pouvez pas créer de modèle de tâche à partir des tâches suivantes : Ajouter une table, Comparer les données, Supprimer les tables et les formats, Télécharger des fichiers de données vers un PC, Créer un projet Forecast Studio, Ouvrir un projet Forecast Studio, Projet de valeur corrigée Forecast Studio, Rapport sous forme de liste, Scoring du modèle, et Tables de synthèse.
- Dans SAS Enterprise Guide, si vous utilisez une invite dans une tâche et créez ensuite un modèle de la tâche, l'invite n'est pas enregistrée dans le modèle de tâche.

Remarque : dans SAS Enterprise Guide, vous pouvez aussi créer un modèle à partir du Générateur de requêtes.

Dans SAS Enterprise Guide, vous pouvez gérer les modèles en utilisant le Gestionnaire de modèles de tâches.

Créer un modèle de tâche

- 1 Ouvrez la tâche et spécifiez les options pour le modèle.
- 2 Cliquez sur la flèche vers le bas située à côté de **Exécuter** et sélectionnez **Créer un modèle** dans le menu contextuel. La boîte de dialogue **Créer un modèle de tâche** apparaît.
- 3 Saisissez le nom et la description du modèle.
- 4 Indiquez si vous souhaitez créer un nouveau modèle de tâche dans un groupe ou remplacer un modèle existant.
- 5 Cliquez sur **Créer**.

Modifier un modèle de tâche

Vous pouvez effectuer des modifications sur les modèles que vous avez créés. Si vous effectuez des modifications sur un modèle qui est basé sur une requête, vous devez d'abord vérifier que les références des données sont toujours valides.

- 1 Ouvrez le modèle de tâche que vous voulez modifier. La fenêtre Tâches s'ouvre avec les sélections faites dans le modèle.
- 2 Spécifiez les options de la tâche que vous voulez enregistrer et sélectionnez **Exécuter** ⇒ **Créer un modèle**. La boîte de dialogue **Créer un modèle de tâche** apparaît.
- 3 Sélectionnez **Remplacer un modèle de tâche existant** et sélectionnez le modèle existant que vous voulez remplacer. Cliquez sur **Créer**. Vos modifications sont enregistrées dans le modèle de tâche existant.

Exécuter un modèle de tâche

- 1 Ouvrir un modèle de tâche en sélectionnant **Tâches** ⇒ **Modèles de tâches** ⇒ **nom-du-groupe** ⇒ **nom-du-modèle**.
- 2 (Facultatif) Sélectionnez la source de données en entrée. Par défaut, la source de données active du projet est la source de données en entrée de la tâche. Vous pouvez néanmoins changer la source de données en entrée de la plupart des tâches à partir du volet **Données**.
- 3 Dans le volet **Données** de la fenêtre de tâche, attribuez des variables aux fonctions de la tâche à partir de la source de données en entrée.
- 4 (Facultatif) Modifiez les options supplémentaires de la tâche.
- 5 Cliquez sur **Exécuter**.

Importer un modèle de tâche

- 1 Dans le volet **Tâches**, cliquez sur **Importer un modèle**.
- 2 Pour sélectionner un ou plusieurs modèles à importer, cliquez sur **Ajouter**.
- 3 Dans la liste déroulante **Importer vers le groupe**, sélectionnez le groupe dans lequel vous voulez placer les modèles importés.

- 4 Cliquez sur **Importer**. Les modèles apparaissent à présent dans le volet **Tâches** et sont disponibles dans le menu **Tâches**.

Exporter un modèle de tâche

- 1 Dans le volet **Tâches**, cliquez sur **Exporter le modèle**.
- 2 Sélectionnez les modèles à exporter et ajoutez-les à la zone **Modèles de tâches sélectionnés**.
- 3 Pour sélectionner l'emplacement de ces modèles, cliquez sur **Exporter vers**.
- 4 Cliquez sur **Exporter**. Une copie des modèles est enregistrée à l'emplacement que vous avez sélectionné.

Tâche Ajouter une table

<i>À propos de la tâche Ajouter une table</i>	11
<i>Ajouter une table : Créer la liste des tables d'entrée</i>	12
<i>Ajouter une table : enregistrer les résultats</i>	12

À propos de la tâche Ajouter une table

La tâche Ajouter une table concatène au moins deux tables d'entrée pour ne créer qu'une seule table de sortie ou vue SQL. Vous pouvez concaténer jusqu'à 256 tables. La table résultante contient chaque ligne et chaque variable qui se trouvent dans l'une des tables en entrée. Une variable, qui n'est présente que dans certaines tables d'entrée, recevra une valeur manquante sur chaque ligne des résultats provenant d'une table d'entrée ne contenant pas la variable.

Par exemple, vous pouvez avoir quatre tables d'entrée indiquant les chiffres de vente d'un produit pour chaque trimestre de 2002. Pour créer une seule table de sortie présentant toutes les ventes de 2002, vous pouvez utiliser la tâche Ajouter une table.

Remarque : la Tâche Ajouter une table n'est disponible que dans SAS Enterprise Guide.

pour exécuter cette tâche, les prérequis sont :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	SQL
Produits SAS requis	Base SAS
Autres produits SAS recommandés	Néant

Ajouter une table : Créer la liste des tables d'entrée

Dans la liste, cliquez sur **Tables** pour sélectionner et ordonner les tables à concaténer.

La zone **Tables à ajouter** affiche la liste des tables qui vont être concaténées, dans l'ordre dans lequel elles seront concaténées. Utilisez cette zone pour créer la liste de tables voulue.

- 1 Cliquez sur **Ajouter une table**. La boîte de dialogue **Ouvrir les données** apparaît. Sélectionnez l'emplacement de la table.
- 2 Dans la boîte de dialogue qui s'ouvre, sélectionnez le nom de la table à ajouter. Pour sélectionner plus d'une table, maintenez la touche CTRL enfoncée tout en sélectionnant les tables à ajouter.
- 3 Cliquez sur **Ouvrir** ou sur **OK** pour ajouter la table dans la zone **Tables à ajouter**.

Remarque : vous ne pouvez pas supprimer la table que vous avez utilisée pour ouvrir la tâche Ajouter une table.

Ajouter une table : enregistrer les résultats

Dans la liste, cliquez sur **Résultats** pour créer une table SAS ou une vue SQL à partir de vos données.

Dans la zone **Format des résultats**, vous pouvez indiquer si vous voulez enregistrer les résultats ajoutés sous la forme d'une table SAS ou d'une vue SQL.

- Pour enregistrer les résultats sous forme de table SAS, cliquez sur **Table de données**. Il s'agit de la présentation par défaut.
- Pour enregistrer les résultats sous forme de vue SQL, cliquez sur **Vue des données**.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre

les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Aire

<i>À propos de la tâche Aire</i>	15
<i>Aire : Sélectionner un type de graphe</i>	16
<i>Aire : Attribuer des variables aux fonctions d'analyse</i>	16
<i>Aire : Définir les options d'apparence</i>	17
Personnaliser l'aire	17
Options d'apparence supplémentaires	18
<i>Indiquer des titres et des notes de bas de page</i>	18
<i>Afficher les propriétés</i>	19

À propos de la tâche Aire

La tâche Aire crée des aires, splines, escaliers ou superpositions qui présentent les relations mathématiques entre deux variables. Pour lesquelles l'aire située au-dessous de la courbe est importante ou dans lesquelles vous remplissez la zone située sous la courbe avec un motif ou une combinaison de couleurs pour créer un impact visuel.

Pour plus d'informations, voir "[Sélectionner un format de sortie pour les graphiques](#)" page 6.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GPLOT
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Aire : Sélectionner un type de graphe

Dans la liste, cliquez sur **Aire** pour accéder à ces options.

Nom du rôle	Description
Aire	crée une courbe à l'aide de deux variables. L'aire située au-dessous de la courbe est colorée.
Aire lissée	crée un graphique sous forme de courbe à l'aide de la routine de lissage. L'aire située sous la courbe lissée est colorée.
Aire en escalier	crée une courbe à l'aide de la fonction par paliers. L'aire située au-dessous du trait du diagramme en escaliers est colorée.
Aire à plusieurs colonnes verticales utilisant la superposition	crée une courbe distincte pour chaque variable Y par rapport à une même variable X. Toutes les courbes apparaissent sur le même graphique en raison de l'option de superposition. Une couleur différente est appliquée à l'aire située en dessous de chaque courbe.
Plusieurs aires par colonne de groupe	crée une courbe distincte pour chaque valeur unique de la variable attribuée à la fonction Groupe . Toutes les courbes apparaissent dans le même graphique. Une couleur différente est appliquée à l'aire située en dessous de chaque courbe.

Aire : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Horizontal	La colonne que vous attribuez à cette fonction est la variable d'axe horizontal ou X du graphique.
Vertical	La colonne que vous attribuez à cette fonction est la variable d'axe vertical ou Y du graphique.
Vertical (Droite)	La colonne que vous attribuez à cette fonction est la variable d'axe vertical ou Y du côté droit du graphique. Cette variable est représentée par rapport à la variable d'axe X de sorte qu'un autre graphe soit généré dans le diagramme.
Groupe	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de courbes tracées sur le graphique.
Grouper les graphiques par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de graphes créés. Un graphique distinct est créé pour chaque valeur unique. Activez la case à cocher Produire la même mise à l'échelle pour tous les graphiques si vous voulez que l'échelle de chaque axe soit la même pour tous les graphiques créés. Par défaut, la plage de valeurs de chaque axe est basée sur les valeurs minimales et maximales des données représentées, et peuvent donc varier d'un graphique à l'autre. Si vous activez cette option, la plage de chaque axe est la même pour tous les graphiques.

Aire : Définir les options d'apparence

Personnaliser l'aire

Dans la liste, sous l'en-tête Apparence, cliquez sur **Aire** pour accéder à ces options.

Sélectionnez un style de remplissage :

- **Vide**
- **Plein**
- **Trait**

Si vous sélectionnez **Trait**, sélectionnez un style de remplissage et la densité du trait.

Remarque : certains styles de remplissage ne sont pas pris en charge par tous les formats de sortie graphique. La sélection du style de remplissage **Vide** ou **Trait** ne produit aucun effet lorsque vous utilisez le format ActiveX, Java, Image ActiveX ou Image Java. Ces formats appliquent aux aires le style de remplissage Plein.

Options d'apparence supplémentaires

Certaines options d'apparence sont communes à plusieurs tâches graphiques. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir les options d'un axe" page 729](#)
- ["Ajouter des lignes de référence" page 732](#)
- ["Définir les options de légende" page 733](#)
- ["Personnaliser l'aire" page 734](#)

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Tâche Modélisation et prévision ARIMA

<i>À propos de la tâche Modélisation et prévision ARIMA</i>	21
<i>Modélisation et prévision ARIMA : Attribuer des variables aux fonctions d'analyse</i>	22
<i>Modélisation et prévision ARIMA : Définir les options d'identification</i>	23
Différencier les séries de réponse	23
Spécifier des tests de stationnarité	24
Générer des courbes et enregistrer les résultats	24
<i>Modélisation et prévision ARIMA : Définir les options d'estimation</i>	25
Activer les étapes d'estimation	25
Définir le modèle ARIMA	25
Sélectionner la méthode d'estimation et les options	26
Enregistrer les résultats	27
<i>Modélisation et prévision ARIMA : Définir les options de prévision</i>	27
Activer les étapes de prévision	27
Indiquer l'intervalle de prévision	28
Générer des courbes et enregistrer les résultats	28
<i>Indiquer des titres et des notes de bas de page</i>	29
<i>Afficher les propriétés</i>	30

À propos de la tâche Modélisation et prévision ARIMA

La méthode de modélisation et de prévision ARIMA analyse et établit des prévisions sur les données d'une série chronologique à une variable à intervalles réguliers, sur les données de fonction de transfert et sur les données d'intervention en utilisant le modèle de moyenne mobile intégrée autorégressive (ARIMA) ou le modèle de moyenne mobile autorégressive (ARMA). Un modèle ARIMA prédit une

valeur dans une série chronologique de réponse sous la forme d'une combinaison linéaire de ses propres valeurs, des erreurs passées (également appelées chocs ou innovations), ainsi que des valeurs actuelles et passées d'autres séries chronologiques.

Une analyse ARIMA se décompose en trois phases élémentaires :

- Phase 1 : Identification de la série chronologique
- Phase 2 : Estimation de la série chronologique
- Phase 3 : Prévision des valeurs futures de la série chronologique.

Vous pouvez utiliser la tâche de modélisation et de prévision ARIMA pour effectuer les analyses suivantes :

- Déterminer l'incidence de facteurs saisonniers. Par exemple, vous pouvez modéliser une série chronologique comprenant en entrée des variables météorologiques (température, pluviométrie, humidité, etc.).
- Prévoir la demande pour un produit en s'appuyant sur des données historiques.

Remarque : la tâche de modélisation et de prévision ARIMA est organisée de sorte que chacune des trois phases se trouve sur une page distincte. Lorsque vous cliquez sur **Exécuter**, seule la phase correspondant à la page sélectionnée ainsi que les étapes précédentes sont réalisées. Ainsi, si vous accédez à la page **Phase 3 : Prévisions** et faites des changements, puis accédez à la page **Phase 1 : Identification** et cliquez sur **Exécuter**, seule la phase Identification est exécutée.

Pour exécuter cette tâche, les prérequis sont :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	ARIMA
Produits SAS requis	SAS Base, SAS/ETS
Autres produits SAS recommandés	SAS/GRAPH

Modélisation et prévision ARIMA : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variable de la série chronologique	indique la variable à prédire. Vous devez attribuer une seule variable à cette fonction.
Variable ID temps	indique la variable utilisée pour attribuer des dates aux lignes. Vous devez attribuer une seule variable à cette fonction. La première variable ayant un format date, temps ou date temps SAS est automatiquement attribuée à cette fonction. Vous pouvez aussi créer une variable ID temps.
Grouper les prévisions par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper en fonction d'une variable que vous avez déjà sélectionnée comme variable dépendante ou variable indépendante.

Modélisation et prévision ARIMA : Définir les options d'identification

Différencier les séries de réponse

Dans le volet de sélection sous la phase 1 : En-tête Identification, cliquez sur **Différenciation** pour accéder à ces options.

Pour différencier les séries de réponse, cochez la case **Différencier les séries de réponse**. Lorsque vous activez cette case à cocher, les premières différences sont

effectuées par défaut. Pour changer le degré de différenciation, saisissez les nouvelles valeurs dans la zone **Retards de différenciation**. Par exemple, indiquez 1,1 pour réaliser la différenciation deux fois avec un retard de 1.

Spécifier des tests de stationnarité

Dans le volet de sélection sous la phase 1 : En-tête Identification, cliquez sur **Tests de stationnarité** pour accéder à ces options.

Pour indiquer un test de stationnarité, sélectionnez le bouton d'option du test voulu. Par défaut, aucun test n'est effectué. Si vous sélectionnez un test de stationnarité, vous pouvez indiquer les ordres autorégressifs pour le test. Par défaut, les ordres autorégressifs de 0, 1 et 2 sont indiqués. Dans la zone **Ordres autorégressifs**, saisissez les ordres autorégressifs souhaités en les séparant par des virgules ou un intervalle d'ordres séparés par des points de suspension. Par exemple, saisissez 1 . . 3 pour spécifier les ordres autorégressifs de 1, 2 et 3.

Si vous sélectionnez **Réaliser tests de Dickey-Fuller augmentés**, vous pouvez spécifier un décalage saisonnier afin d'effectuer un test saisonnier de Dickey-Fuller. Indiquez un entier dans la zone **Décalage saisonnier**. La valeur par défaut du décalage saisonnier est de 1, ce qui n'est pas un test saisonnier.

Générer des courbes et enregistrer les résultats

Dans le volet de sélection sous la phase 1 : En-tête Identification, cliquez sur **Courbes et résultats** pour accéder à ces options.

Nom d'option	Description
Courbe de valeurs réelles	affiche une courbe des séries chronologiques d'entrée dans les résultats. Vous pouvez également spécifier le nombre de retards à considérer dans le calcul des autocorrélations et des corrélations avec retard. Le nombre par défaut de retards est de 24, ou un quart du nombre des observations, en fonction de celui qui est le plus faible
Enregistrer les autocorrélations et les covariances croisées	enregistre une table de sortie contenant des autocorrélations et des covariances croisées. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un

Nom d'option	Description
	emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir .
Supprimer la sortie d'identification affichée	supprime la sortie d'identification des résultats.

Modélisation et prévision ARIMA : Définir les options d'estimation

Activer les étapes d'estimation

Dans la liste, sous l'en-tête Phase 2 : En-tête Estimation, cliquez sur **Activer les étapes d'estimation** pour accéder à ces options.

Pour activer les options pour les étapes d'estimation, activez la case à cocher **Exécuter les étapes d'estimation**.

Remarque : la tâche Modélisation et prévision ARIMA se déroule en trois phases. Les options pour l'étape 1 : Identification sont toujours activées. Vous devez cocher les cases appropriées pour activer les options de la phase 2 : Estimation et phase 3 : Prévisions. Vous devez activer les étapes de l'estimation avant d'activer celles de la prévision.

Définir le modèle ARIMA

Dans la liste, sous l'en-tête Phase 2 : Estimation, cliquez sur **Définition du modèle** pour accéder à ces options.

Vous pouvez ajouter des paramètres de modèle autorégressif ou de moyenne mobile à l'estimation du modèle.

Pour ajouter de nouveaux paramètres à l'estimation du modèle :

- 1 Dans la zone **Facteurs pour modèle AR** ou **Facteurs pour modèle MA**, saisissez un entier, une liste d'entiers séparés par des virgules ou un intervalle d'entiers pour représenter les retards des paramètres de moyenne mobile à estimer pour

chaque facteur multiplicatif de la liste. Par exemple, saisissez 1...3, 1 TO 3 ou 1, 2, 3 pour indiquer (1, 2, 3), ou saisissez 3 pour spécifier (3).

- 2 Cliquez sur **Ajouter** ou appuyez sur Enter pour ajouter le nouveau paramètre à la liste.
- 3 Répétez ce processus pour ajouter des conditions de multiplication. Par exemple, saisissez 1, cliquez sur **Ajouter**, saisissez 12, puis cliquez sur **Ajouter** pour spécifier (1)(12).

Une fois qu'un paramètre est créé, vous pouvez le modifier ou le supprimer de la liste dans la zone **Autorégressive** ou **Moyenne mobile**.

Sélectionner la méthode d'estimation et les options

Dans la liste, sous l'en-tête Phase 2 : Estimation, sélectionnez **Options du modèle** pour accéder à ces options.

Nom d'option	Description
Méthode d'estimation	Sélectionnez une méthode d'estimation dans la liste déroulante. Ces méthodes d'estimation sont disponibles : ■ La Méthode de vraisemblance maximale produit des estimations de vraisemblance maximales. La fonction de vraisemblance est optimisée via les moindres carrés non linéaires à l'aide de la méthode de Marquardt. ■ La Méthode inconditionnelle des moindres carrés produit des estimations inconditionnelles des moindres carrés. La méthode ULS est également appelée méthode exacte des moindres carrés (ELS). ■ La Méthode conditionnelle des moindres carrés produit des estimations conditionnelles des moindres carrés. Ces estimations supposent que les erreurs non observées passées sont égales à 0. Il s'agit de la méthode par défaut.
Nombre maximum d'itérations	précise un nombre maximum d'itérations pour l'étape d'estimation. Indiquez un entier dans la zone Nombre maximal d'itérations. La valeur par défaut est 50.
Suppression de la composante moyenne du modèle	Exclut la composante moyenne du modèle.

Enregistrer les résultats

Dans la liste, sous l'en-tête Phase 2 : Estimation, sélectionnez **Résultats** pour accéder à ces options.

Nom d'option	Description
Enregistrer les données de sortie	Vous pouvez enregistrer certains ou tous vos résultats dans une table de sortie. Activez les cases à cocher appropriées pour enregistrer vos résultats. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur Parcourir.
Supprimer la sortie de l'estimation affichée	Supprime la sortie de la phase d'estimation.

Modélisation et prévision ARIMA : Définir les options de prévision

Activer les étapes de prévision

Dans la liste, sous l'en-tête Phase 3 : En-tête Prévisions, cliquez sur **Activer les étapes de prévision** pour accéder à ces options.

Pour activer les options pour les étapes de prévision, activez la case à cocher **Exécuter les étapes de prévision**.

Remarque : la tâche Modélisation et prévision ARIMA se déroule en trois phases. Les options pour l'étape 1 : Identification sont toujours activées. Vous devez cocher

les cases appropriées pour activer les options de la phase 2 : Estimation et phase 3 : Prévisions. Vous devez activer les étapes de l'estimation avant d'activer celles de la prévision.

Indiquer l'intervalle de prévision

Dans la liste, sous l'en-tête Phase 3 : Prévisions, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Intervalle de temps entre les observations	Sélectionnez un intervalle de temps dans la liste déroulante pour représenter l'intervalle de temps entre les observations de la table d'entrée. L'intervalle que vous choisissez doit être approprié aux données d'entrée, sinon l'analyse échouera ou produira des résultats incorrects. Si vous sélectionnez Nombre d'unités comme intervalle de temps, vous pouvez également spécifier les unités de temps par intervalle. La valeur par défaut est 1. Pour tout intervalle de temps autre que Nombre d'unités, l'option Unités de temps par intervalle n'est pas disponible.
Nombre d'intervalles à prédire	Indiquez le nombre d'intervalles de temps pour lesquels vous voulez faire des prévisions. La valeur par défaut est 24. Par exemple, si vous indiquez un intervalle de temps entre les observations du mois et si vous indiquez 6 intervalles de temps pour la prévision, vous obtiendrez des prévisions pour 6 mois après la fin de vos données d'entrée. Sélectionnez un niveau de confiance dans la liste déroulante Niveau de confiance. La valeur par défaut est 95 pour cent.

Générer des courbes et enregistrer les résultats

Dans la liste, sous l'en-tête Phase 3 : Prévision, sélectionnez **Courbes et résultats** pour accéder à ces options.

Nom d'option	Description
Options des courbes de prévisions	Génère ces courbes : ■ Pour générer une courbe des valeurs réelles, des valeurs prédites d'une étape en avant, des prévisions à étapes

Nom d'option	Description
	multiples et des intervalles de confiance, sélectionnez la case à cocher Prévisions. ■ Pour générer des courbes de résidus, activez la case à cocher Résidus. Utilisez les options Limiter les interv. des données réelles affichées et Intervalles à afficher pour indiquer le nombre d'intervalles de données historiques à représenter avant le début des données de prévision.
Enregistrer les prévisions	Pour générer une table de sortie pour les prévisions, cochez la case Enregistrer les prévisions. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Supprimer la sortie de prévision affichée	Supprime la sortie de prévision des résultats.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Affecter une bibliothèque au projet

<i>À propos de l'Assistant Affecter une bibliothèque de projets</i>	31
<i>Affecter une bibliothèque de projets : Indiquez un nom et un serveur</i>	32
<i>Affecter une bibliothèque de projets : Spécifier un moteur</i>	32
<i>Affecter une bibliothèque de projets : Spécifier les options de configuration</i>	35
<i>Affecter une bibliothèque de projets : Vérifier les informations récapitulatives</i>	36

À propos de l'Assistant Affecter une bibliothèque de projets

L'Assistant **Affecter une bibliothèque de projets** permet de définir une bibliothèque pour le projet. L'assistant génère l'instruction LIBNAME dont vous avez besoin pour définir la bibliothèque. Cette bibliothèque est ensuite définie chaque fois que vous exécutez la tâche ou le flux de processus contenant la tâche dans le projet.

Par exemple, supposons que vous deviez analyser certaines données. Elles résident sur votre répertoire local et ne se trouvent pas dans une bibliothèque SAS définie par un administrateur. Il faut donc que vous définissiez une bibliothèque SAS pour pouvoir accéder aux données. Vous pouvez utiliser l'Assistant **Affecter une bibliothèque de projets** pour créer cette instruction LIBNAME.

.....
Remarque : cette tâche n'est disponible que dans SAS Enterprise Guide.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	Néant
Produits SAS requis	Base SAS
Autres produits SAS recommandés	Néant

Affecter une bibliothèque de projets : Indiquez un nom et un serveur

- 1 Saisissez le nom de la bibliothèque dans la zone **Nom**. Le nom peut comporter jusqu'à huit caractères et toutes les lettres doivent être des majuscules. Le nom doit être unique pour chaque serveur sur lequel la bibliothèque est créée. Il ne doit comporter aucun de ces caractères : \ / : * ? " < > |
- 2 Sélectionnez un serveur dans la liste des serveurs disponibles.
- 3 Cliquez sur **Suivant** pour sélectionner un moteur pour la bibliothèque.

Affecter une bibliothèque de projets : Spécifier un moteur

- 1 Avant de pouvoir sélectionner un moteur, vous devez indiquer un nom et un serveur pour la bibliothèque.
- 2 Sélectionnez le type de moteur. Vous avez le choix entre ces types de moteur :

Type de moteur	Description
Système de fichiers	Spécifiez ces options : ■ Pour que SAS puisse choisir le moteur en fonction du type des fichiers situés à l'emplacement que vous sélectionnez pour la bibliothèque, activez la case à cocher Laisser SAS choisir le moteur en fonction du contenu du répertoire indiqué.

Type de moteur	Description
	Remarque : n'activez pas cette case à cocher si des fichiers de différents types se trouvent à l'emplacement sélectionné pour la bibliothèque. ■ Sélectionnez le moteur de la bibliothèque. Si vous sélectionnez le moteur REMOTE-SAS/SHARE, vous devez spécifier la paire nom-valeur de slibref/SAMPSHR dans les options de configuration de la bibliothèque. Remarque : si vous activez la case à cocher Laisser SAS choisir le moteur en fonction du contenu du répertoire indiqué, vous ne pouvez pas sélectionner de moteur. ■ Indiquez le chemin physique de la bibliothèque sur le serveur dans la zone Chemin d'accès. Cliquez sur Parcourir pour sélectionner cet emplacement. Vous pouvez utiliser des caractères génériques dans le chemin. <userid> remplace l'identifiant Windows <sample> remplace l'exemple de répertoire d'Enterprise Client <serveruser> remplace l'identifiant pour la connexion au serveur <serverpassword> remplace le mot de passe pour la connexion au serveur <libraryuser> remplace l'identifiant pour la connexion à la bibliothèque <librarypassword> remplace le mot de passe pour la connexion à la bibliothèque Si des informations d'identification sont requises pour accéder au serveur sélectionné, vous êtes invité à saisir un identifiant et un mot de passe. Remarque : si vous sélectionnez le moteur BASE, V9, V8, V7, V604 ou V6, il est inutile de saisir un chemin d'accès. Pour les autres moteurs, le chemin est en revanche obligatoire.
Système de base de données	Spécifiez ces options : ■ Sélectionnez le moteur à utiliser pour accéder aux fichiers de la bibliothèque. ■ Sélectionnez le Database Server externe de la bibliothèque.

Type de moteur	Description
	Remarque : la liste déroulante affiche les Database Servers qui ont été définis dans le référentiel de métadonnées SAS actif. Si le serveur que vous voulez utiliser n'est pas défini dans les métadonnées, il n'apparaît pas dans la liste déroulante. Votre administrateur peut utiliser SAS Management Console pour définir d'autres serveurs. ■ Indiquez le schéma de base de données utilisé pour accéder aux données du serveur. Le nom du schéma que vous saisissez doit correspondre à celui d'un schéma déjà défini sur le Database Server. ■ Saisissez un identifiant et un mot de passe pour le Database Server.
WebDAV	Spécifiez ces options : ■ Sélectionnez le WebDAV Server à utiliser comme source de la bibliothèque. Remarque : la liste déroulante affiche les WebDAV Servers qui ont été définis dans le référentiel de métadonnées SAS actif. Si le serveur que vous voulez utiliser n'est pas défini dans les métadonnées, il n'apparaît pas dans la liste déroulante. Votre administrateur peut utiliser SAS Management Console pour définir d'autres serveurs. ■ Indiquez le chemin physique de la bibliothèque sur le serveur dans la zone Chemin d'accès. Vous pouvez utiliser ces caractères génériques lorsque vous spécifiez le chemin d'accès physique où réside la bibliothèque sur le serveur ou bien lorsque vous indiquez les options de la bibliothèque. <userid> remplace l'identifiant Windows <sample> remplace l'exemple de répertoire d'Enterprise Client <serveruser> remplace l'identifiant pour la connexion au serveur <serverpassword> remplace le mot de passe pour la connexion au serveur <libraryuser> remplace l'identifiant pour la connexion à la bibliothèque <librarypassword> remplace le mot de passe pour la connexion à la bibliothèque

Type de moteur	Description
	■ Saisissez un identifiant et un mot de passe pour le WebDAV Server.

- 3 Cliquez sur **Suivant** pour spécifier les options de configuration de la bibliothèque.

Affecter une bibliothèque de projets : Spécifier les options de configuration

- 1 Avant de spécifier les options de configuration, vous devez sélectionner un moteur pour la bibliothèque.
- 2 Définissez les options de bibliothèque comportant des paires nom-valeur. Vous ne pouvez spécifier que des options ayant des valeurs associées dans la table. ACCESS=READONLY constitue un exemple d'option avec une paire nom-valeur. Cette option déclare la bibliothèque en lecture seule de sorte que vous ne puissiez pas copier des données de sortie dans cette bibliothèque. Pour obtenir la liste des options disponibles, veuillez consulter la documentation de votre environnement d'exploitation.

Remarque : si vous sélectionnez le moteur **REMOTE-SAS/SHARE**, vous devez spécifier la paire nom-valeur de slibref/SAMPSHR dans les options de configuration de la bibliothèque.

Pour saisir un nom d'option ou une valeur :

- a Cliquez sur une cellule vide de la table. Lorsque vous spécifiez un nom et une valeur d'option, une ligne vide apparaît automatiquement dans la table.
- b Saisissez le nom de l'option dans la colonne **Nom**. Vous pouvez spécifier n'importe quelle option valide de l'instruction LIBNAME, mais il est conseillé toutefois d'éviter l'utilisation des options d'invite (telles que DBPROMPT=YES).
- c Saisissez la valeur correspondant à l'option dans la colonne Valeur. La valeur ne doit pas être un blanc.

Remarque : si la bibliothèque doit se connecter à la base de données Sybase, vous devez inclure READLOCK_TYPE=PAGE en tant qu'option.

- 3 Saisissez les options qui ne requièrent pas de valeurs dans **Options supplémentaires**.

- 4 Cliquez sur **Suivant** pour vérifier les informations du récapitulatif.

Affecter une bibliothèque de projets : Vérifier les informations récapitulatives

Avant de créer la nouvelle bibliothèque, vous pouvez vérifier les informations que vous avez indiquées.

- 1 Avant de vérifier ces informations, vous pouvez définir les options de configuration de la bibliothèque.
- 2 Vérifiez les informations dans le champ de saisie. Pour changer une valeur, cliquez sur **Précédent**.
- 3 Pour tester la nouvelle bibliothèque sur le serveur sélectionné, cliquez sur **Tester la bibliothèque**. Si SAS Enterprise Guide se connecte au serveur et attribue la bibliothèque, l'état du serveur passe à **OK**. Pour visualiser le journal du test, cliquez sur **Afficher le Journal**.
- 4 Cliquez sur **Terminer** pour créer la nouvelle bibliothèque.

Histogramme

<i>À propos de la tâche Histogramme</i>	37
<i>Histogramme : Sélectionner un type de graphique</i>	38
<i>Histogramme : Attribuer des variables aux fonctions d'analyse</i>	39
<i>Histogramme : Définir les options d'apparence</i>	40
Définir la présentation	40
Options d'apparence supplémentaires	41
<i>Indiquer des titres et des notes de bas de page</i>	41
<i>Afficher les propriétés</i>	42

À propos de la tâche Histogramme

La tâche Histogramme crée des histogrammes verticaux, horizontaux ou à trois dimensions qui comparent des valeurs numériques ou des statistiques entre différentes valeurs d'une variable de graphique. Les histogrammes représentent l'importance relative des données en affichant des barres de différentes hauteurs. Chaque barre représente une catégorie de données.

Vous pouvez utiliser un histogramme pour comparer le montant total des ventes de chaque succursale d'une enseigne. Dans ce type de graphique, chaque barre représente le total des ventes de chaque succursale.

Pour plus d'informations, voir ["Sélectionner un format de sortie pour les graphiques" page 6](#).

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GCHART
Produits SAS requis	SAS Base, SAS/STAT

Nom du prérequis	Noms des procédures et des produits
Autres produits SAS recommandés	SAS/GRAPH

Histogramme : Sélectionner un type de graphique

Dans la liste, cliquez sur **Histogramme** pour accéder à ces options.

Nom d'option	Description
Barres simples	Affiche la grandeur des données avec des barres représentant chacune une catégorie de données. La longueur (ou hauteur) des barres représente la valeur de la statistique pour l'abscisse correspondante.
Barres groupées	Produit un groupe de barres distinct pour chaque valeur unique de la variable attribuée à la fonction Grouper barres par .
Empilé(e)	Crée un segment distinct dans chaque barre pour chaque valeur unique de la variable attribuée à la fonction Empiler . Chaque segment à l'intérieur d'une barre est représenté par une couleur différente.
Barres groupées/empilées	Combine les graphiques à barres groupées et empilées. Elle produit un groupe de barres distinct pour chaque valeur unique de la variable attribuée à la fonction Grouper barres par et crée un segment distinct à l'intérieur de chaque barre pour chaque valeur unique de la variable attribuée à la fonction Empiler .
Barres colorées	Crée un histogramme simple dans lequel chaque barre est représentée par une couleur différente.
Barre colorée groupée	Produit un groupe de barres distinct pour chaque valeur unique de la variable attribuée à la fonction Grouper barres par . Chaque barre d'un groupe apparaît dans une couleur différente.
Barres colorées groupées	Produit un groupe de barres distinct pour chaque valeur unique de la variable attribuée à la fonction Grouper barres par . Les barres à l'intérieur d'un groupe sont de couleur identique.

Nom d'option	Description
Barre imbriquée	Produit un groupe de barres distinct pour chaque valeur unique de la variable attribuée à la fonction Grouper barres par . Les barres contenues dans chaque groupe sont jointes et chaque barre est représentée par une couleur différente.
Barres 3D groupées	Produit un groupe distinct de barres en 3D pour chaque valeur unique de la variable attribuée à la fonction Grouper barres par .
Barres 3D groupées/empilées	Combine les graphiques à barres groupées et empilées. Elle produit un groupe de barres distinct pour chaque valeur unique de la variable attribuée à la fonction Grouper barres par et crée un segment distinct à l'intérieur de chaque barre pour chaque valeur unique de la variable attribuée à la fonction Empiler .
Barres 3D colorées et groupées	Produit un groupe distinct de barres en 3D pour chaque valeur unique de la variable attribuée à la fonction Grouper barres par . Chaque barre d'un groupe apparaît dans une couleur différente.
Groupes 3D colorés et groupés	Produit un groupe distinct de barres en 3D pour chaque valeur unique de la variable attribuée à la fonction Grouper barres par . Les barres à l'intérieur d'un groupe sont de couleur identique.
Groupes de plusieurs mesures	Permet d'attribuer plusieurs variables d'analyse à la fonction Somme de. Le graphique obtenu contient un groupe de barres pour chaque valeur unique de la variable attribuée à la fonction Colonne à représenter . Les barres à l'intérieur d'un groupe sont de couleur identique.

Histogramme : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Remarque : Les rôles disponibles dépendent du type de graphique sélectionné.

Nom du rôle	Description
Colonne à représenter	Les valeurs de la colonne que vous attribuez à cette fonction définissent les catégories de données à représenter. Vous devez attribuer une variable à cette fonction.
Grouper les barres par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent les groupes dans lesquels les barres sont réparties. Vous devez attribuer une variable à cette fonction.
Empiler	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de segments de chaque barre.
Somme de	La colonne que vous attribuez à cette fonction détermine la longueur des barres. Pour sélectionner la statistique particulière utilisée pour déterminer la longueur de barre, sélectionnez Options avancées dans la liste. Si vous n'attribuez pas de colonne à cette fonction, la fréquence de chaque valeur de la Colonne à représenter détermine la longueur des barres ou bien la taille des sections.
Grouper les graphiques par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de graphes créés. Un graphique distinct est créé pour chaque valeur unique.

Histogramme : Définir les options d'apparence

Définir la présentation

Dans la liste, sous l'en-tête Apparence, cliquez sur **Présentation** pour accéder à ces options.

Par défaut, la plupart des histogrammes sont en 2D et la case à cocher **2D** est activée. Pour créer un graphique en 3D, désactivez la case à cocher **2D** et sélectionnez une forme de barre dans la liste déroulante **Forme**.

Remarque : la case à cocher **2D** n'est pas disponible si vous avez choisi le type de graphique histogramme en 3D.

Utilisez la liste déroulante **Ordre** pour indiquer l'ordre de tri des barres. Par défaut, il est déterminé automatiquement. Vous pouvez toutefois choisir d'organiser les barres par ordre croissant ou décroissant de longueur ou de hauteur.

Activez la case à cocher **Couleur du contour** si vous voulez appliquer un contour à chaque barre. Vous pouvez sélectionner une couleur de contour dans la liste déroulante.

Utilisez la liste déroulante **Taille de la barre** pour définir la taille des barres. Par défaut, elle est déterminée automatiquement. Vous pouvez toutefois saisir la largeur de barre ou l'espacement entre les barres sous forme de pourcentage de la largeur totale du graphique.

Options d'apparence supplémentaires

Certaines options d'apparence sont communes aux tâches graphiques. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir l'apparence des barres" page 735](#)
- ["Définir les options d'un axe" page 729](#)
- ["Ajouter des lignes de référence" page 732](#)
- ["Définir les options de légende" page 733](#)
- ["Personnaliser l'aire" page 734](#)
- ["Définir les options avancées" page 739](#)

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Assistant Histogramme

<i>À propos de l'assistant Histogramme</i>	43
<i>Assistant Histogramme : Sélectionner les données</i>	43
<i>Assistant Histogramme : Attribuer des variables aux fonctions</i>	44
<i>Assistant Histogramme : Définir les options d'apparence</i>	45
<i>Assistant Histogramme : Indiquer les titres et les notes de bas de page</i>	46

À propos de l'assistant Histogramme

L'Assistant Histogramme vous aide à créer des histogrammes verticaux ou horizontaux, à deux ou trois dimensions, avec des options pour les barres groupées ou empilées.

Assistant Histogramme : Sélectionner les données

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Cliquez sur **Suivant** pour attribuer des variables aux fonctions.

Assistant Histogramme : Attribuer des variables aux fonctions

- 1 Avant d'attribuer des variables aux fonctions, vous devez sélectionner les données.
- 2 Attribuez des variables aux fonctions suivantes :
 - **Barres** - les valeurs de la variable que vous attribuez à cette fonction déterminent le nombre de barres. La variable peut être alphanumérique ou numérique. Vous devez attribuer une variable à cette fonction. Vous pouvez trier les barres ou indiquer le type de variable :
 - **Trier** - vous pouvez indiquer l'ordre de tri des barres. Par défaut, la tâche trie les barres automatiquement par ordre croissant de la valeur de l'abscisse. Vous pouvez trier les barres par ordre croissant ou décroissant de la hauteur de barre.
 - **Type de variable** - Par défaut, le nombre de barres est déterminé automatiquement. Pour les colonnes numériques, les valeurs sont divisées par intervalles, avec une barre par intervalle. Si la colonne numérique représente des valeurs discrètes, vous pouvez sélectionner l'option **Discret**.
 - **Hauteur de barre** - la variable ou statistique que vous attribuez à cette fonction détermine la longueur des barres.
 - Si vous attribuez une variable à cette fonction, vous pouvez calculer la somme ou la moyenne de chaque barre. Par défaut, la somme est calculée.
 - Si vous n'attribuez pas de variable à cette fonction, vous pouvez calculer la fréquence, la fréquence cumulée, le pourcentage ou le pourcentage cumulé de chaque barre. Par défaut, la fréquence est calculée.

Remarque : si vous créez un histogramme horizontal, cette option est appelée **Longueur de barre**.

- 3 (Facultatif) Attribuez des variables à ces fonctions :
 - **Grouper par** - les valeurs de la variable attribuée à cette fonction déterminent les groupes dans lesquels les barres sont divisées. Sur un histogramme vertical, les groupes de barres apparaissent le long de l'axe X. Sur un histogramme horizontal, ils apparaissent le long de l'axe Y.

Remarque : cette option n'est disponible que si vous avez attribué une variable à la fonction **Profondeur**.

- **Profondeur** - les valeurs de la variable attribuée à cette fonction déterminent les groupes dans lesquels les barres sont divisées. Les groupes de barres apparaissent le long de l'axe Z.

Remarque : cette option n'est pas disponible si vous créez un histogramme horizontal ou attribuez une variable à la fonction **Grouper par**.

- **Empiler par** - les valeurs de la variable attribuée à cette fonction déterminent le nombre de segments de chacune des barres.
- **Représenter par** - des graphiques distincts sont générés pour chaque groupe. Les groupes sont déterminés par les valeurs de la variable que vous attribuez à cette fonction.

- 4 Cliquez sur **Suivant** pour définir les options d'apparence.

Assistant Histogramme : Définir les options d'apparence

- 1 Avant de définir l'apparence du graphique, vous devez attribuer des variables aux fonctions.
- 2 Activez la case à cocher **Graphique 3D** pour créer un histogramme à trois dimensions. Par défaut, les barres sont représentées en 2D.
- 3 Indiquez la façon dont les couleurs sont attribuées aux barres du graphique dans la liste déroulante **Colorier les barres par**. Les options **Ttes les barres identiques** et **Catégorie de barre** sont disponibles par défaut. Si vous attribuez une variable à la fonction **Grouper par** ou **Profondeur**, l'option **Variable de groupe** ou **Variable de profondeur** est également disponible.

Remarque : si une variable a été attribuée à la fonction **Empiler par**, cette option n'est pas disponible.

Si toutes les barres sont de la même couleur, sélectionnez cette dernière dans la liste déroulante **Couleur de la barre**. Si chaque catégorie de barre possède une couleur distincte, l'Assistant Histogramme sélectionne automatiquement un jeu de couleurs.

- 4 Si une légende est disponible, activez la case à cocher **Légende** pour l'inclure dans le graphique. Sélectionnez la position de la légende dans la liste déroulante.
- 5 Pour indiquer la statistique à afficher sous forme de libellé pour chaque barre, activez la case à cocher **Libellés des données** et sélectionnez la statistique dans la liste déroulante. Si aucune variable n'a été attribuée à la fonction **Hauteur de**

barre, les statistiques Fréquence, Fréquence cumulée, Pourcentage et Pourcentage cumulée sont disponibles. Si une variable a été attribuée à la fonction **Hauteur de barre**, les statistiques Somme et Moyenne sont également disponibles.

- 6 Pour modifier les libellés de chaque variable attribuée, cliquez sur **Libellé de l'axe**. La boîte de dialogue **Libellé de l'axe** apparaît.
- 7 Pour insérer des lignes de référence dans le graphique, activez la case à cocher **Utiliser les lignes de référence**.
- 8 Pour inclure des graduations sur le graphique, activez la case à cocher **Graduations**.
- 9 Cliquez sur **Suivant** pour indiquer des titres et des notes de bas de page.

Assistant Histogramme : Indiquer les titres et les notes de bas de page

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

- 1 Avant d'indiquer les titres et les notes de bas de page, vous devez spécifier les options d'apparence.
- 2 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie. Vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.
- 3 Cliquez sur **Terminer** pour exécuter la tâche.

Courbe-histogramme

<i>À propos de la tâche Courbe-Histogramme</i>	47
<i>Courbe-histogramme : Attribuer des variables aux fonctions d'analyse</i>	48
<i>Courbe-histogramme : Définir les options d'apparence</i>	49
Sélectionner la méthode d'interpolation	49
Options d'apparence supplémentaires	50
<i>Indiquer des titres et des notes de bas de page</i>	50
<i>Afficher les propriétés</i>	51

À propos de la tâche Courbe-Histogramme

La tâche Courbe-histogramme crée un graphique à barres verticales avec une courbe superposée. La courbe représente la valeur d'une statistique calculée pour l'une des variables de la table d'entrée.

Vous pouvez utiliser cette tâche pour :

- afficher et comparer les niveaux d'importance exacte et relative
- examiner la contribution de chaque partie par rapport à un ensemble
- déterminer les tendances et les motifs des données

Par exemple, vous pouvez utiliser cette tâche pour indiquer la relation entre les ventes et les coûts d'un produit pour chaque pays dans lequel ce produit est vendu. L'axe horizontal constitue l'abscisse, avec une barre pour chaque pays. Les axes verticaux sont les axes de réponse. Chaque barre représente le montant des ventes dans chaque pays. Sur l'axe vertical droit, une courbe représentant les coûts de production et de marketing du produit dans chaque pays est superposée aux barres. Ce graphique permet de comparer les coûts par rapport aux ventes pour chaque pays et sur l'ensemble des pays.

Pour plus d'informations, voir ["Sélectionner un format de sortie pour les graphiques"](#) page 6.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GBARLINE
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Courbe-histogramme : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Colonne à représenter	Les valeurs de la colonne que vous attribuez à cette fonction définissent les catégories de données à représenter. Vous devez attribuer une variable à cette fonction.
Somme de la barre de	La colonne que vous attribuez à cette fonction détermine la longueur des barres. Par défaut, si aucune colonne n'est attribuée à cette fonction, la longueur des barres est déterminée par le nombre total d'observations (statistique de fréquence). En revanche, si une colonne est attribuée à cette fonction, la longueur des barres est déterminée par la somme des valeurs de chaque catégorie. Pour choisir la statistique particulière utilisée pour déterminer la longueur des barres, sélectionnez Options avancées dans la liste. Pour plus d'informations, voir Définir les options avancées.

Nom du rôle	Description
Somme de la ligne de	Les colonnes que vous attribuez à cette fonction indiquent les variables à utiliser pour les courbes. Vous pouvez attribuer plusieurs variables à cette fonction. Pour choisir la statistique à utiliser pour le tracé de chaque courbe, sélectionnez Graphes dans la liste. Par défaut, la somme des valeurs de chaque catégorie est utilisée pour tracer chaque courbe. Remarque : si vous n'attribuez aucune colonne à cette fonction, le nombre total d'observations de chaque catégorie (statistique de fréquence) est utilisé pour tracer la courbe. Pour choisir la statistique à utiliser pour le tracé de la courbe, sélectionnez Options avancées dans le volet de sélection.
Empiler	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de segments de chaque barre.
Grouper les graphiques par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de graphes créés. Un graphique distinct est créé pour chaque valeur unique.

Courbe-histogramme : Définir les options d'apparence

Sélectionner la méthode d'interpolation

Dans la liste, sous l'en-tête Apparence, cliquez sur **Interpolations** pour accéder à ces options.

- 1 Sélectionnez une variable dans la liste. Ces variables sont celles que vous avez attribuées à la fonction **Somme de la courbe de**.
- 2 Sélectionnez une méthode d'interpolation dans la liste déroulante. En fonction de la méthode d'interpolation que vous choisissez, vous devrez peut-être spécifier des paramètres supplémentaires. Les méthodes d'interpolation suivantes sont disponibles :
 - **Courbe** relie les points de données par des lignes droites. Les points sont reliés dans l'ordre dans lequel ils apparaissent dans les données d'entrée.
 - **Nuage de points** supprime toute interpolation. Une courbe des points de données est créée.

- **Escalier** représente les données sous forme d'un escalier en utilisant la fonction par paliers.
- 3 (Facultatif) Activez la case à cocher **Appliquer à tout** pour utiliser les mêmes méthode d'interpolation et propriétés pour toutes les variables de la liste. Cette option n'est disponible que si la liste comporte plusieurs variables.

Options d'apparence supplémentaires

Certaines options d'apparence sont communes aux tâches graphiques. Pour plus d'informations sur ces options, voir ces rubriques :

- "Définir l'apparence des barres" page 735
- "Définir les options des barres" page 736
- "Définir les options des courbes" page 737
- "Définir les options d'un axe" page 729
- "Ajouter des lignes de référence" page 732
- "Définir les options de légende" page 733
- "Personnaliser l'aire" page 734
- "Définir les options avancées" page 739

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Prévision de base

<i>À propos de la tâche Prévision de base</i>	53
<i>Prévision de base : Attribuer des variables aux fonctions d'analyse</i>	54
<i>Prévision de base : Définir les options de prévision</i>	55
Sélectionner une méthode de prévision et des intervalles	55
Sélection des options de données en entrée	56
Définir les options de la méthode de prévision	56
<i>Prévision de base : Générer des graphes</i>	57
<i>Prévision de base : Définition des options des résultats</i>	58
<i>Indiquer des titres et des notes de bas de page</i>	59
<i>Afficher les propriétés</i>	59

À propos de la tâche Prévision de base

La tâche Prévision de base constitue un moyen rapide et automatique pour générer des prévisions pour plusieurs séries chronologiques, en une étape. Elle utilise des méthodes de prévision extrapolables dans lesquelles les prévisions d'une série portent uniquement sur les valeurs passées et de type heure et non sur les autres variables.

Vous pouvez utiliser cette tâche pour établir les prévisions de vente d'un produit. Par exemple, supposons que vous ayez enregistré les données mensuelles des ventes d'un produit dans une table et que vous souhaitez prévoir les ventes au cours des 10 prochains mois. Cette tâche peut calculer les dates pour les valeurs prévisionnelles et l'intervalle de confiance pour les ventes prévisionnelles. Vous pouvez enregistrer la sortie dans une table SAS. Vous pouvez également enregistrer dans une table de sortie SAS les paramètres des modèles de prévision utilisés, ainsi que les statistiques qui évaluent l'ajustement des modèles aux données. Enfin, vous pouvez représenter les prévisions, l'intervalle de confiance, les résidus et les valeurs réelles.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	FORECAST
Produits SAS requis	SAS Base, SAS/ETS
Autres produits SAS recommandés	SAS/GRAPH

Prédiction de base : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variable de prévision	indique les variables à utiliser pour la prévision. Vous devez attribuer une seule variable à cette fonction.
Variable ID temps	indique la variable utilisée pour attribuer des dates aux lignes. Vous devez attribuer une seule variable à cette fonction. La première variable ayant un format date, temps ou date temps SAS est automatiquement attribuée à cette fonction. Vous pouvez aussi créer une variable ID temps.
Grouper les prévisions par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper en fonction d'une variable déjà attribuée à une autre fonction.

Prévision de base : Définir les options de prévision

Sélectionner une méthode de prévision et des intervalles

Dans la liste, cliquez sur **Options de prévision** pour accéder à ces options.

Nom d'option	Description
Méthode de prévision	Pour indiquer une méthode de prévision, sélectionnez la méthode voulue dans la liste déroulante. Note ■ Autorégressive et pas à pas (pour traitements autorégressifs) ajuste un modèle de tendance temporelle aux séries et prend la différence entre chaque valeur et la tendance estimée. L'écart restant est ensuite ajusté à l'aide d'un modèle autorégressif. Il s'agit de la méthode par défaut. ■ Lissage exponentiel (pour les traitements de moyenne mobile) ajuste un modèle de tendance qui est tel que les données les plus récentes sont pondérées plus lourdement que les données du début de la série. La pondération d'une observation est une fonction géométrique (exponentielle) du nombre de périodes que l'observation étend dans le passé par rapport à la période en cours. ■ Méthode (multiplicative) de Winters (pour traitements saisonniers multiplicatifs) utilise des équations de mise à jour semblables au lissage exponentiel pour ajuster des paramètres pour le modèle. Dans ce modèle, vous avez deux paramètres de tendance et une fonction qui sélectionne le paramètre saisonnier correspondant à l'heure. ■ Méthode additive de Winters (pour traitements saisonniers additifs) est semblable à la méthode de Winters (multiplicative) excepté que les paramètres saisonniers sont ajoutés à la tendance et non multipliés par la tendance. Remarque : la sélection d'une méthode de prévision peut activer ou désactiver d'autres commandes dans cette fenêtre, en fonction de la méthode choisie.

Nom d'option	Description
Nombre d'intervalles à prédire	Indiquez le nombre d'intervalles de temps pour lesquels vous souhaitez produire des prévisions dans la zone Nombre d'intervalles à prévoir. La valeur par défaut est 12. Par exemple, si vous indiquez un intervalle de temps entre les observations du mois et indiquez 6 intervalles de temps pour la prévision, vous obtiendrez des prévisions pour 6 mois après la fin de vos données d'entrée.

Sélection des options de données en entrée

Dans la liste, cliquez sur **Options de prévision** pour accéder à ces options.

Nom d'option	Description
Intervalle de temps entre les observations	spécifie l'intervalle de temps entre les lignes de la table de données en entrée. L'intervalle de temps par défaut est Nombre d'unités. Si vous sélectionnez Nombre d'unités, vous pouvez également sélectionner les unités de temps par intervalle. Pour tous les autres intervalles de temps, l'option Unités de temps par intervalle n'est pas disponible. L'intervalle que vous choisissez doit être approprié aux données d'entrée, sinon l'analyse échouera ou produira des résultats incorrects.
Longueur du cycle saisonnier	spécifie une longueur de cycle saisonnier et les intervalles par cycle saisonnier pour une méthode de prévision saisonnière (méthode multiplicative de Winters ou méthode additive de Winters). Les longueurs de cycle saisonnier dépendent de l'intervalle de temps que vous avez sélectionné dans la zone au-dessus. La longueur par défaut d'un cycle saisonnier correspond au Nombre d'intervalles.

Définir les options de la méthode de prévision

Dans la liste, cliquez sur **Options de prévision** pour accéder à ces options.

Nom du graphe	Description
Degré du modèle de la tendance temporelle	Pour indiquer le degré du modèle de la tendance temporelle, sélectionnez une valeur dans la liste déroulante. Les modèles Constant , Linéaire et Quadratique sont disponibles. Par défaut, toutes les méthodes de prévision utilisent un modèle de tendance temporelle linéaire, sauf la méthode de lissage exponentiel qui utilise un modèle quadratique.
Niveau de confiance	Sélectionnez un niveau de confiance dans la liste déroulante. La valeur par défaut est 95 pour cent.
Pondération de lissage	Vous pouvez indiquer des constantes de lissage pour les méthodes Lissage exponentiel et la méthode multiplicative ou additive de Winters. Pour la méthode de lissage exponentiel, vous pouvez indiquer une constante de lissage uniquement pour la constante. Pour la méthode multiplicative ou additive de Winters, vous pouvez indiquer une constante de lissage pour le composant de la constante, le composant linéaire et quadratique ainsi que pour le composant saisonnier. Les composants linéaire et quadratique ne sont disponibles que si vous spécifiez une pondération de lissage de terme constant. De même, vous ne pouvez spécifier une pondération de lissage saisonnier qu'après en avoir spécifié un pour les composantes linéaire et quadratique. Les pondérations de lissage par défaut dépendent du degré de modèle de tendance temporelle spécifié et des autres pondérations de lissage spécifiées. Pour changer une constante de lissage, cochez la case appropriée et saisissez une nouvelle valeur dans le champ.

Prévision de base : Générer des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Utilisez les cases à cocher de la zone **Types** pour générer des graphes sur les valeurs réelles, les données de la prévision et les résidus.

Pour les tracés de données de prévision, vous pouvez choisir d'inclure ces données.

- **Valeurs réelles et préd. à un pas** inclut les valeurs des données réelles et les prédictions à un pas sur la courbe. Les prédictions à un pas font référence à une prédiction effectuée au cours de la période précédente ($t-1$) pour une période future.
- **Intervalle de confiance des prévisions** inclut l'intervalle de confiance pour chaque valeur prévisionnelle du graphe.

Pour tous les types de graphe, utilisez la case à cocher **Limitier l'affichage des données réelles** et la zone **Intervalles à afficher** pour sélectionner le nombre d'intervalles des données historiques à tracer avant les données de la prévision.

Remarque : la case à cocher **Limitier l'affichage des données réelles** n'est pas disponible si vous attribuez une variable à la fonction **Grouper les prévisions par**.

Prévion de base : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

La table contenant les valeurs de la prévision est toujours enregistrée. Pour générer une table sur les résultats estimés des paramètres, enregistrée de façon permanente, cochez la case **Résultats estimés des paramètres**. Sélectionnez les statistiques supplémentaires à inclure.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Vous pouvez indiquer si vous voulez inclure les prévisions et les résultats estimés des paramètres dans les résultats. Vous devez activer la case à cocher **Résultats estimés des paramètres** dans la zone **Résultats estimés des paramètres** pour pouvoir sélectionner **Résultats estimés des paramètres** dans la zone **Afficher la sortie**.

Remarque : si vous ne sélectionnez pas de graphes et ne demandez pas l'affichage des prévisions ni des estimations, aucune sortie ne sera générée.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Carte de la moyenne avec superposition des boîtes de Tukey

<i>À propos de la carte de la moyenne avec superposition des boîtes de Tukey</i>	61
<i>Cartes de contrôle : Attribuer des variables aux fonctions d'analyse</i>	62
<i>Cartes de contrôle : Indiquer des limites de contrôle</i>	64
<i>Carte de contrôle : Sélectionner les tests à effectuer</i>	65
<i>Cartes de contrôle : Définir les options d'apparence du graphe</i>	66
Définir les options de la Carte de la moyenne avec superposition des boîtes de Tukey	66
Options d'apparence supplémentaires	67
<i>Cartes de contrôle : Enregistrer les résultats d'analyse</i>	68
<i>Indiquer des titres et des notes de bas de page</i>	69
<i>Afficher les propriétés</i>	69

À propos de la carte de la moyenne avec superposition des boîtes de Tukey

La carte de contrôle de Shewhart est un outil graphique et analytique qui permet de déterminer si un processus est sous contrôle statistique. La tâche Carte de la moyenne avec superposition des boîtes de Tukey crée une carte de la moyenne pour les moyennes de sous-groupe. Les boîtes à moustaches des mesures pour chaque sous-groupe sont superposées sur la carte de la moyenne. La carte de la moyenne avec superposition des boîtes de Tukey fournit une analyse détaillée des distributions dans les sous-groupes. Elle est recommandée pour des échantillons de sous-groupes de taille importante (10 ou plus).

Vous pouvez utiliser une carte de la moyenne avec superposition des boîtes de Tuckey pour rechercher la distribution des sorties et déterminer si un processus est sous contrôle statistique. Par exemple, supposons qu'une compagnie pétrolière utilise une turbine qui vaporise l'eau pompée dans le sol afin de réduire la viscosité du pétrole et en faciliter l'extraction. Ce processus est répété 20 fois par jour, et la puissance utilisée (en kW) pour réchauffer l'eau à la température souhaitée est enregistrée. Vous pouvez utiliser une carte de la moyenne avec superposition des boîtes de Tuckey pour examiner la distribution de la puissance par jour et déterminer si le niveau moyen du processus de chauffage est sous contrôle statistique.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	SHEWHART
Produits SAS requis	SAS Base, SAS/QC, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Cartes de contrôle : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Mesure	indique les variables de traitement à analyser. Ces variables contiennent les mesures individuelles. Vous devez attribuer au moins une variable à cette fonction.
Identificateur du sous-groupe	indique la variable qui identifie les sous-groupes dans les données. Les valeurs de la variable indiquent de quelle manière les observations de la table d'entrée sont

Nom du rôle	Description
	organisées en sous-groupes rationnels. En général, les valeurs sont les suivantes : ■ les indices qui donnent l'ordre dans lequel les sous-échantillons sont collectés ■ les dates ou heures auxquelles les sous-échantillons ont été collectés ■ les libellés qui identifient de façon unique les échantillons de sous-groupe Un identificateur de sous-groupe est requis et vous ne pouvez attribuer qu'une seule variable à cette fonction. Pour trier les données selon les sous-groupes identifiés, activez la case à cocher Trier par sous-groupe.
Taille de l'échantillon du sous-groupe	spécifie la taille de l'échantillon du sous-groupe comme valeur de la variable. La taille de l'échantillon du sous-groupe est nécessaire et vous n'attribuez qu'une seule variable numérique à cette fonction. Activez la case à cocher Utiliser une valeur numérique pour le sous-groupe (au lieu d'une var.) lorsque la taille de l'échantillon de sous-groupe est déterminée et que vous devez définir une valeur constante pour la taille. Saisissez la valeur à utiliser dans la zone Taille de l'échantillon du sous-groupe. La taille de l'échantillon du sous-groupe doit être une valeur comprise entre la valeur maximale de la variable de traitement et 99999. Remarque : cette fonction n'est pas nécessaire si vous avez activé la case à cocher Utiliser une valeur numérique pour le sous-groupe (au lieu d'une var.) dans le volet Options.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.
Variables de bloc	indique deux variables maximum, qui regroupent les données par blocs de sous-groupes consécutifs. Ces blocs sont libellés dans une légende, et chaque variable de bloc fournit un niveau pour les libellés de la légende.

Cartes de contrôle : Indiquer des limites de contrôle

Dans la liste, cliquez sur **Limites de contrôle** pour accéder à ces options.

Nom d'option	Description
Limites Sigma	spécifie la largeur des limites de contrôle commun multiple de l'erreur type (sigma) des statistiques récapitulatives du sous-groupe tracé dans le graphique. La largeur doit être positive. Le multiple par défaut est 3.
Sélectionner une méthode de calcul	spécifie la méthode de calcul pour les limites de contrôle. Vous avez le choix entre ces options : ■ Calculer les limites de contrôle à partir des données actives ■ Calculer les limites de contrôle à partir de la table sélectionnée Pour désigner la table contenant les limites de contrôle ou les paramètres à partir desquels les limites de contrôle peuvent être calculées, cliquez sur Parcourir. Sélectionnez la table à utiliser et cliquez sur Ouvrir. Remarque : la table que vous sélectionnez doit être une table LIMITS=. Pour plus d'informations sur les variables requises dans une table LIMITS=, consultez la documentation des procédures du graphique que vous êtes en train de créer. Pour représenter les statistiques descriptives de tous les sous-groupes, activez la case à cocher Trace les statistiques descriptives de tous les sous-groupes, que la taille de l'échantillon du sous-groupe soit égale à la taille nominale des limites de contrôle ou non. Vous pouvez représenter ces statistiques lorsque la taille de l'échantillon est la même pour pratiquement tous les sous-groupes et lorsque vous voulez afficher les limites de contrôle fixes (non soumises à variation) correspondant à la taille d'échantillon nominale n. Si vous sélectionnez cette option alors que les tailles des échantillons de vos sous-groupes sont très différentes, l'interprétation des limites de contrôle n'a de sens que pour les sous-groupes dont la taille de l'échantillon est égale à n.

Nom d'option	Description
	Si vous souhaitez ajouter des symboles spéciaux pour les points de données où la taille de l'échantillon est différente de n, activez la case à cocher Ajouter des symboles spéciaux pour les points correspondant aux tailles d'échantillon différentes de la taille nominale pour les limites de contrôle fixes. ■ Spécifier les limites de contrôle Dans les champs Supérieure, Centrale, et Inférieure, indiquez respectivement la limite de contrôle supérieure, la ligne centrale et la limite de contrôle inférieure. Si l'effectif du sous-groupe varie, vous devez indiquer un effectif nominal. Remarque : si vous attribuez des variables à la fonction Grouper l'analyse par, vous ne pouvez pas saisir des valeurs pour la limite de contrôle.

Carte de contrôle : Sélectionner les tests à effectuer

Dans la liste, cliquez sur **Tests** pour accéder à ces options.

Remarque : ces tests ne s'appliquent pas aux limites de contrôle qui ne sont pas trois limites sigma ou qui varient en fonction de la taille de l'échantillon du sous-groupe.

Nom d'option	Description
Sélectionner un test	Utilisez les cases à cocher pour demander un ou plusieurs tests pour des sujets spéciaux, également connus sous les noms de test de dérive : Run test, Pattern et Western Electric tests. Ces tests détectent les dérives particulières non aléatoires dans les points représentés sur le graphique de contrôle principal. L'occurrence du modèle non aléatoire est appelé signal. Les tests peuvent indiquer la présence d'une cause spéciale de variation. Lorsque vous sélectionnez l'un des tests, une description de ce test apparaît dans la zone Description.

Nom d'option	Description
Libellé	Utilisez les champs de saisie afin de saisir des libellés pour les points où les tests pour causes spéciales sont positifs. Les libellés ne peuvent contenir plus de 16 caractères. Chaque libellé apparaît à tous les points du test où un signal est rencontré.
Texte identifiant les signaux de test	■ Utilisez la liste déroulante afin de sélectionner une couleur pour les libellés que vous avez indiqués. ■ Utilisez la case à cocher Afficher lignes de zone pour inclure des lignes qui représentent les zones A, B et C sur un graphique principal. Ces zones se construisent en divisant l'intervalle entre les limites de contrôle en six sous-intervalles espacés de façon égale. ■ Utilisez la case Surpasser la limite 3 sigma pour tester des sujets spéciaux si vous spécifiez une limite de contrôle sigma différente du multiple par défaut de 3. Pour changer la limite sigma, cliquez sur Limites de contrôle dans le volet de sélection. ■ Cochez la case Appliquer le test aux dérives des points se chevauchant afin d'appliquer des tests pour les causes spéciales aux dérives de points se chevauchant.

Cartes de contrôle : Définir les options d'apparence du graphe

Définir les options de la Carte de la moyenne avec superposition des boîtes de Tuckey

Dans la liste, sous l'en-tête Apparence, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Dans les limites de contrôle, Limites de	spécifie la couleur de ces éléments graphiques.

Nom d'option	Description
contrôle et Cadre	■ Dans les limites de contrôle spécifie la couleur à l'intérieur des limites de contrôle inférieure et supérieure. ■ Limites de contrôle spécifie la couleur des limites de contrôle et de la ligne centrale. Spécifie également la couleur des libellés de ces lignes. ■ Cadre spécifie les couleurs de remplissage du rectangle délimité par les axes et le cadre. Remarque : les options Dans les limites de contrôle et Limites de contrôle ne sont pas disponibles pour une carte de la moyenne avec superposition des boîtes de Tuckey si vous supprimez les limites de contrôle. Pour activer les limites de contrôle, désactivez la case à cocher Supprimer l'affichage des limites de contrôle.
Couleurs des boîtes	spécifie la couleur du contour et la couleur de remplissage des boîtes.
Supprime l'affichage des limites de contrôle	supprime la légende par défaut de la taille de l'échantillon du sous-groupe. Cette option est utile si les tailles des échantillons des sous-groupes sont constantes et égales à la taille de l'échantillon de limite de contrôle, car cette dernière apparaît automatiquement dans l'angle supérieur droit du graphique.
Faire varier la largeur de la boîte selon la taille du sous-groupe	a pour effet de faire varier les boîtes à moustaches proportionnellement à la taille du sous-groupe.
Boîtes à entailles	crée des boîtes à moustaches avec entailles. Les extrémités des entailles sont situées sur la médiane, plus et moins la valeur $1,58(IQR/\sqrt{n})$, où IQR est l'écart interquartile et $\sqrt{n}$ est la racine carrée de la taille de l'échantillon du sous-groupe. Les médianes (lignes centrales) de boîtes à moustaches sont sensiblement différentes au niveau 0,05 environ, si les entailles correspondantes ne se superposent pas.

Options d'apparence supplémentaires

Certaines options d'apparence sont communes à plusieurs tâches de la carte de contrôle. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir les options d'un axe" page 743](#)
- ["Ajouter des lignes de référence" page 744](#)

Cartes de contrôle : Enregistrer les résultats d'analyse

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Pour enregistrer les résultats de l'analyse, vous pouvez créer des tables de sortie contenant plusieurs statistiques. Activez les cases à cocher pour les données de sortie que vous voulez enregistrer de façon permanente.

Nom du rôle	Description
Statistiques de sous-groupe et table des limites de contrôle	crée une table de sortie contenant les informations représentées sur le graphe. Cela inclut les valeurs de la variable de sous-groupe ainsi que les statistiques descriptives correspondantes et les limites de contrôle. Remarque : cette option crée une table OUTTABLE=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option TABLE= décrite par la procédure SHEWHART. Pour plus d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des statistiques de sous-groupe	crée une table de sortie contenant les statistiques descriptives du sous-groupe. Remarque : cette option crée une table OUTHISTORY=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option HISTORY= décrite par la procédure SHEWHART. Pour plus d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des limites de contrôle	crée une table de sortie qui enregistre les limites de contrôle. Vous pouvez utiliser cette table lorsque vous sélectionnez la méthode de calcul pour les limites de contrôle lors d'une utilisation ultérieure de cette tâche.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),

- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Boîte à moustaches

<i>À propos de la tâche Boîte à moustaches</i>	71
<i>Boîte à moustaches : Sélectionner un type de graphe</i>	72
<i>Boîte à moustaches : Attribuer des variables aux fonctions d'analyse</i>	72
<i>Boîte à moustaches : Définir les options d'apparence</i>	73
Définir les options de la boîte à moustaches	73
Définir des options de courbes max-min	75
Options d'apparence supplémentaires	75
<i>Indiquer des titres et des notes de bas de page</i>	76
<i>Afficher les propriétés</i>	76

À propos de la tâche Boîte à moustaches

La tâche Boîte à moustaches crée des boîtes à moustaches, des courbes max-min ou max-min-clôture qui affichent des statistiques descriptives pour certaines variables numériques selon différentes valeurs d'une variable de graphique. Par exemple, une boîte à moustaches peut afficher l'intervalle, l'écart interquartile et la moyenne d'une variable. Une courbe max-min-clôture peut présenter la cote la plus élevée, la cote la plus basse et la cote de clôture d'une action à différentes dates.

Pour plus d'informations, voir ["Sélectionner un format de sortie pour les graphiques" page 6](#).

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GPLOT
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Boîte à moustaches : Sélectionner un type de graphe

Dans la liste, cliquez sur **Boîte à moustaches** pour accéder à ces options.

Nom d'option	Description
Boîte à moustaches	indique que les extrémités inférieures et supérieures de la boîte sont situées aux 25ème et 75ème centiles de l'échantillon. La ligne horizontale centrale est représentée au 50ème centile (médiane). Les lignes verticales, ou moustaches, sont tracées depuis la boîte jusqu'au point le plus extrême compris dans les intervalles interquartiles 1,5. Les valeurs situées en dehors de cet intervalle sont marquées d'un symbole graphique.
Courbe max-min	spécifie qu'une ligne verticale continue relie les valeurs Y minimale et maximale pour chaque valeur X. Une graduation indique la valeur Y moyenne pour chaque valeur X. Ce type de graphe est particulièrement adapté à la représentation des données boursières.
Courbe max-min-clôture	spécifie qu'une ligne verticale continue relie les valeurs Y minimale et maximale pour chaque valeur X. Une graduation indique une valeur de clôture. Ce type de graphe est particulièrement adapté à la représentation des données boursières.

Boîte à moustaches : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la

source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Horizontal	La colonne que vous attribuez à cette fonction est la variable d'axe horizontal ou X du graphique.
Vertical	La colonne que vous attribuez à cette fonction est la variable d'axe vertical ou Y du graphique.
Vertical (Droite)	La colonne que vous attribuez à cette fonction est la variable d'axe vertical ou Y du côté droit du graphique. Cette variable est représentée par rapport à la variable d'axe X de sorte qu'un autre graphe soit généré dans le diagramme.
Grouper les graphiques par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de graphes créés. Un graphique distinct est créé pour chaque valeur unique. Activez la case à cocher Produire la même mise à l'échelle pour tous les graphiques si vous voulez que l'échelle de chaque axe soit la même pour tous les graphiques créés. Par défaut, la plage de valeurs de chaque axe est basée sur les valeurs minimales et maximales des données représentées, et peuvent donc varier d'un graphique à l'autre. Si vous activez cette option, la plage de chaque axe est la même pour tous les graphiques.

Boîte à moustaches : Définir les options d'apparence

Définir les options de la boîte à moustaches

Dans la liste, sous l'en-tête Apparence, cliquez sur **Nuage de points** pour accéder à ces options.

Remarque : ces options ne sont disponibles que si vous sélectionnez une boîte à moustaches.

Nom du rôle	Description
Tracer les lignes médianes et les hauts et bas des moustaches	Vous pouvez choisir de tracer une ligne entre les points médians et de tracer les hauts et les bas des moustaches.
Spécifier les couleurs de remplissage et des valeurs hors norme	Pour sélectionner votre propre combinaison de couleurs de remplissage et de contour, activez la case à cocher Indiquer les couleurs , puis sélectionnez vos couleurs dans les listes déroulantes Rempliss. et valeurs hors norme et Contour . La seconde liste déroulante concerne les contours ainsi que les lignes des graphes max-min et des boîtes à moustaches.
Inclure des valeurs extérieures à la portée de l'axe	Pour que les valeurs extérieures à la portée de l'axe soient incluses dans les calculs d'interpolation, activez la case à cocher Inclure des valeurs hors de la portée de l'axe . Lorsque cette option est sélectionnée, les valeurs comprises en dehors de la portée de l'axe sont incluses dans les calculs d'interpolation, mais exclues du graphique. Par défaut, cette case à cocher est désactivée et les valeurs hors de la portée de l'axe sont omises des calculs d'interpolation.
Indiquer la longueur des moustaches	Vous pouvez sélectionner un centile de longueur pour les moustaches. Ces options sont disponibles : ■ + - 1,5 fois l'intervalle interquartile permet de tracer les lignes verticales, ou moustaches, depuis la boîte jusqu'au point le plus extrême compris dans les intervalles interquartiles 1,5. ■ Extrêmes supérieur/inférieur trace les moustaches à partir des valeurs supérieure et inférieure extrêmes. Remarque : si vous sélectionnez cette option, toutes les valeurs qui se situent en dehors de l'intervalle du centile de longueur pour les moustaches (également appelés valeurs hors norme) ne sont pas marquées de symboles graphiques. Si vous sélectionnez une autre option pour le centile de longueur, les valeurs hors norme sont marquées de symboles graphiques. ■ 1er centile inférieur, 99ème centile supérieur trace les moustaches à partir des valeurs extrêmes du 1er centile jusqu'au 99ème centile. ■ 5ème centile inférieur, 95ème centile supérieur trace les moustaches à partir des valeurs extrêmes du 5ème centile jusqu'au 95ème centile. ■ 10ème centile inférieur, 90ème centile supérieur trace les moustaches à partir des valeurs extrêmes du 10ème centile jusqu'au 90ème centile.

Nom du rôle	Description
	■ 25ème centile inférieur, 75ème centile supérieur ne produit aucune moustache car le graphique s'étend du 25ème au 75ème centile.
Personnaliser les valeurs hors norme	Vous pouvez spécifier le type de police à utiliser, le symbole ainsi que la hauteur.

Définir des options de courbes max-min

Dans la liste, sous l'en-tête Apparence, cliquez sur **max-min** pour accéder à ces options.

Remarque : ces options ne sont disponibles que si vous sélectionnez une courbe max-min ou max-min-clôture.

Nom du rôle	Description
Tracer les lignes médianes et les hauts et bas des moustaches	Vous pouvez choisir de tracer une ligne entre les points médians et de tracer les hauts et les bas des moustaches.
Connexion des valeurs maximale et minimale avec les barres vides	Vous pouvez choisir de tracer des barres à la place des lignes entre les valeurs Y maximale et minimale.
Tracer une graduation à la valeur de clôture	Vous pouvez choisir de tracer une graduation à la valeur de clôture plutôt qu'à la valeur de la moyenne. Pour utiliser cette option, vous devez disposer exactement de trois valeurs Y (maximale, minimale et clôture) pour chaque valeur X. S'il existe plus ou moins de trois valeurs Y pour chaque valeur X, la graduation est placée au niveau de la moyenne des valeurs Y.

Options d'apparence supplémentaires

Certaines options d'apparence sont communes aux tâches graphiques. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir les options d'un axe" page 729](#)

- [“Ajouter des lignes de référence” page 732](#)
- [“Personnaliser l'aire” page 734](#)

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,

- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Graphique à bulles

<i>À propos de la tâche Bulle</i>	79
<i>Graphique à bulles : Attribuer des variables aux fonctions d'analyse</i>	80
<i>Graphique à bulles : Définir les options d'apparence</i>	81
Définir les options des bulles	81
Options d'apparence supplémentaires	82
<i>Indiquer des titres et des notes de bas de page</i>	82
<i>Afficher les propriétés</i>	82

À propos de la tâche Bulle

La tâche Bulle crée des bulles simples en traçant des cercles aux proportions variantes pour les points de données qui sont représentés sur les axes verticaux et horizontaux. Deux des variables déterminent l'emplacement des points de données. La troisième variable contrôle la taille des cercles.

Vous pouvez utiliser un graphique à bulles pour comparer des informations sur les salaires. Par exemple, supposons que vous vouliez créer un graphique dont chaque bulle représente une catégorie d'ingénieurs dans les domaines de l'aérospatiale, de la chimie, du génie civil, de l'électricité et de la mécanique. L'emplacement de chaque bulle est déterminé par le salaire moyen de la catégorie représentée. La taille de la bulle représente le nombre d'ingénieurs de la catégorie par rapport au nombre total d'ingénieurs.

Pour plus d'informations, voir ["Sélectionner un format de sortie pour les graphiques" page 6](#).

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GPLOT

Nom du prérequis	Noms des procédures et des produits
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Graphique à bulles : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Horizontal	La colonne que vous attribuez à cette fonction est la variable d'axe horizontal ou X du graphique.
Vertical	La colonne que vous attribuez à cette fonction est la variable d'axe vertical ou Y du graphique.
Vertical (Droite)	La colonne que vous attribuez à cette fonction est la variable d'axe vertical ou Y du côté droit du graphique. Cette variable est représentée par rapport à la variable d'axe X de sorte qu'un autre graphe soit généré dans le diagramme.
Taille de la bulle	La colonne que vous attribuez à cette fonction détermine la taille relative des bulles du graphique représentant les variables Horizontal et Vertical. Remarque : si vous avez attribué une variable à la fonction Vertical (Droite) sans en avoir attribué une à la fonction Bulle2, cette fonction détermine la taille relative des bulles du graphique représentant les variables Horizontal et Vertical (Droite).

Nom du rôle	Description
Grouper les graphiques par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de graphes créés. Un graphique distinct est créé pour chaque valeur unique. Activez la case à cocher Produire la même mise à l'échelle pour tous les graphiques si vous voulez que l'échelle de chaque axe soit la même pour tous les graphiques créés. Par défaut, la plage de valeurs de chaque axe est basée sur les valeurs minimales et maximales des données représentées, et peuvent donc varier d'un graphique à l'autre. Si vous activez cette option, la plage de chaque axe est la même pour tous les graphiques.
Taille de la bulle2	La colonne que vous affectez à cette fonction détermine la taille relative des bulles sur le graphique des variables Horizontal et Vertical (droites). Remarque : cette option n'est disponible que si vous attribuez une variable à la fonction Vertical (Droite).

Graphique à bulles : Définir les options d'apparence

Définir les options des bulles

Dans la liste, sous l'en-tête Apparence, cliquez sur **Options de bulles** pour accéder à ces options.

Sélectionnez une couleur de bulle dans la liste déroulante.

Vous pouvez attribuer un libellé aux bulles. Le libellé est la valeur de la colonne attribuée à la fonction **Taille de la bulle** pour la combinaison des valeurs de variables horizontales et verticales. Si cette variable a un format, la valeur formatée est utilisée. Par défaut, les bulles n'ont pas de libellé.

Utilisez le curseur de défilement pour agrandir ou rétrécir toutes les bulles. Vous pouvez éventuellement visualiser la sortie avant d'ajuster cette valeur.

Remarque : si une variable a été attribuée à la fonction **Vertical (Droite)**, vous pouvez également définir les options Bulle2 pour les bulles combinant les variables Horizontal et Vertical (Droite).

Options d'apparence supplémentaires

Certaines options sont communes aux tâches graphiques. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir les options d'un axe" page 729](#)
- ["Ajouter des lignes de référence" page 732](#)
- ["Personnaliser l'aire" page 734](#)

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.
.....

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Corrélation canonique

<i>À propos de la tâche Corrélation canonique</i>	85
<i>Corrélation canonique : Attribuer des variables aux fonctions d'analyse</i>	86
<i>Corrélation canonique : Définir les options statistiques</i>	87
<i>Corrélation canonique : Générer des graphes</i>	88
<i>Corrélation canonique : Définition des options des résultats</i>	89
<i>Indiquer des titres et des notes de bas de page</i>	90
<i>Afficher les propriétés</i>	90

À propos de la tâche Corrélation canonique

La corrélation canonique examine la relation entre la combinaison linéaire d'un ensemble de variables X et celle d'un ensemble de variables Y. Une série d'hypothèses teste si chaque corrélation canonique et toutes les corrélations canoniques minimales sont nulles dans la population. La corrélation simple et multiple sont des cas spéciaux de corrélation canonique dans lesquels l'un des deux ensembles, ou bien les deux ensembles, contient(nent) une variable unique.

La tâche Corrélation canonique peut permettre de déterminer le degré de correspondance entre un ensemble de caractéristiques de travaux et un ensemble de mesures de satisfaction du personnel.

Pour exécuter cette tâche, les prérequis sont :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	CANCORR
Produits SAS requis	SAS Base, SAS/STAT

Nom du prérequis	Noms des procédures et des produits
Autres produits SAS recommandés	SAS/GRAPH

Corrélation canonique : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables de l'ensemble 1	effectue des analyses à l'aide des variables que vous avez affectées à cette fonction. Vous devez attribuer au moins une variable à cette fonction. Utilisez la zone Variables canoniques de l'ensemble 1 afin d'indiquer un préfixe et un libellé pour les variables de l'ensemble 1.
Variables de l'ensemble 2	effectue des analyses à l'aide des variables que vous avez affectées à cette fonction. Vous devez attribuer au moins une variable à cette fonction. Utilisez la zone Variables canoniques de l'ensemble 2 afin d'indiquer un préfixe et un libellé pour les variables de l'ensemble 2.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper des analyses en fonction d'une variable que vous avez déjà sélectionnée pour Variables de l'ensemble 1 ou Variables de l'ensemble 2.

Nom du rôle	Description
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Variables partielles	spécifie les variables pour baser l'analyse canonique sur des corrélations partielles. Les variables de cette fonction viennent en partie de Variables de l'ensemble 1 et de Variables de l'ensemble 2.
Poids relatif	spécifie la valeur de la variable dans chaque ligne pour calculer les coefficients de corrélation pondérés produit-moment. Une ligne est utilisée seulement si la valeur de la variable de cette fonction est supérieure à zéro. Vous ne pouvez attribuer qu'une seule variable à cette fonction.

Corrélation canonique : Définir les options statistiques

Dans la liste, cliquez sur **Statistiques** pour accéder à ces options.

Nom d'option	Description
Analyses de régression à effectuer	définit l'analyse de régression à effectuer. Vous avez le choix entre ces options : ■ Aucune analyse de régression ne réalise pas d'analyses de régression. ■ Ensemble 1 conditionne ensemble 2 et vice versa demande deux analyses de régressions multiples afin que l'option Variables de l'ensemble 1 corresponde aux variables dépendantes pour l'une des deux analyses et aux variables de régression pour l'autre. ■ Ensemble 1 conditionne ensemble 2 demande plusieurs analyses de régression avec l'option Variables de l'ensemble 2 comme variables dépendantes et l'option Variables de l'ensemble 1 comme variables de régression. ■ Ensemble 2 conditionne ensemble 1 demande plusieurs analyses de régression avec l'option Variables de l'ensemble 1 comme variables dépendantes et l'option Variables de l'ensemble 2 comme variables de régression.

Nom d'option	Description
Statistiques de régression	calcule une ou plusieurs de ces statistiques : ■ Coefficients de régression produit des coefficients de régression globaux à partir des analyses de régression. ■ Coefficients de régression standardisés produit des coefficients de régression standardisés. ■ Coefficients des erreurs types produit les erreurs types des coefficients de régression. ■ Statistique t et Prob > t produit des statistiques t pour les coefficients de régression et les niveaux de probabilité des statistiques t. ■ Corrélations multiples au carré produit plusieurs corrélations au carré et des tests F pour les analyses de régression.
Statistiques de corrélation	calcule une ou plusieurs de ces statistiques : ■ Corrélations des coefficients de régression produit des corrélations entre les estimations du coefficient de régression. ■ Corrélations partielles produit des corrélations partielles entre les variables dépendantes et les variables de régression, en supprimant de chaque variable de régression et dépendante les effets de toutes les autres variables de régression. ■ Corrélations partielles au carré produit des corrélations partielles au carré entre les variables de régression et les variables dépendantes, en supprimant de chaque variable de régression et dépendante les effets de toutes les autres variables de régression. ■ Corrélations semi-partielles produit des corrélations semi-partielles entre les variables de régression et les variables dépendantes, en supprimant de chaque variable de régression les effets de toutes les autres variables de régression. ■ Corrélations semi-partielles au carré produit des corrélations semi-partielles au carré entre les variables de régression et les variables dépendantes, en supprimant de chaque variable de régression les effets de toutes les autres variables de régression.

Corrélation canonique : Générer des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Remarque : les options graphiques ne sont pas disponibles si le nombre de variables canoniques pour lesquelles les coefficients et statistiques de redondance sont affichés est égal à zéro. Pour plus d'informations, voir l'option Nombre de variables canoniques dans Résultats.

pour afficher des graphes, activez la case à cocher **Afficher les graphes des variables canoniques**. Utilisez les cases **Variable de début** et **Variable de fin** pour saisir les variables à représenter. Voici quelques exemples :

- Pour représenter uniquement la première variable canonique (V1 versus W1), saisissez 1 dans les deux champs de saisie.
- Pour représenter la première, la seconde et la troisième variable canonique (V1 versus W1, V2 versus W2 et V3 versus W3), saisissez 1 dans le premier champ de saisie et 3 dans le second.

Corrélation canonique : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Enregistrer les résultats dans une table de sortie	enregistre les résultats dans une table de sortie. Vous avez le choix entre ces options : ■ Enregistrer les corrélations crée une table contenant diverses statistiques, notamment les corrélations et coefficients canoniques, ainsi que les statistiques de régressions multiples que vous demandez. ■ Enregistrer les scores crée une table contenant toutes les données d'origine et les scores des variables canoniques. Le nombre de nouvelles variables est égal à deux fois la valeur spécifiée dans la zone Nombre de variables canoniques. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur Parcourir.
Afficher les résultats	affiche la sortie. Vous pouvez spécifier :

Nom d'option	Description
	■ Nombre de variables canoniques indique le nombre de variables canoniques qui incluront une sortie complète dans les résultats. ■ Inclure l'analyse de redondance canonique effectue une analyse de redondance canonique incluant la variance globale (non normalisée) et normalisée ainsi que la proportion cumulée de la variance de chaque ensemble de variables, expliquée par leurs propres variables canoniques et par les variables canoniques inversées. ■ Inclure les scores affiche la table de sortie contenant toutes les données d'origine et les scores des variables canoniques. Si vous voulez créer une table permanente pour l'utiliser ultérieurement, cochez la case Enregistrer les scores.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Carte c

<i>À propos de la tâche Carte c</i>	93
<i>Carte c : Attribuer des variables aux fonctions d'analyse</i>	94
<i>Cartes de contrôle : Indiquer des limites de contrôle</i>	95
<i>Carte de contrôle : Sélectionner les tests à effectuer</i>	96
<i>Cartes de contrôle : Définir les options d'apparence du graphe</i>	98
<i>Cartes de contrôle : Enregistrer les résultats d'analyse</i>	98
<i>Indiquer des titres et des notes de bas de page</i>	99
<i>Afficher les propriétés</i>	100

À propos de la tâche Carte c

La carte de contrôle de Shewhart est un outil graphique et analytique qui permet de déterminer si un processus est sous contrôle statistique. La tâche Carte c crée des cartes c pour le nombre de non-conformités (défauts) dans les échantillons du sous-groupe.

Vous pouvez utiliser une carte c pour suivre le nombre de défauts décelés dans un nouveau produit. Par exemple, supposons qu'un constructeur automobile souhaite suivre le nombre de défauts de peinture d'un nouveau camion. Vingt camions du même modèle sont examinés, et le nombre de défauts de peinture par camion est enregistré. Chaque point de la carte c représente le nombre de défauts d'un camion donné. Les limites de contrôle supérieures et inférieures sont indiquées, de sorte que vous pouvez constater si des points sont au-dessus ou en dessous de ces limites. Si aucun de ces points n'est en dehors des limites de contrôle spécifiées, le processus de peinture est sous contrôle statistique.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	SHEWHART
Produits SAS requis	SAS Base, SAS/QC, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Carte c : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Nombre de non-conformités	indique les variables contenant le nombre de non-conformités. Vous devez attribuer au moins une variable à cette fonction
Identificateur du sous-groupe	indique la variable qui identifie les sous-groupes dans les données. Les valeurs de la variable indiquent de quelle manière les observations de la table d'entrée sont organisées en sous-groupes rationnels. En général, les valeurs sont les suivantes : ■ les indices qui donnent l'ordre dans lequel les sous-échantillons sont collectés ■ les dates ou heures auxquelles les sous-échantillons ont été collectés ■ les libellés qui identifient de façon unique les échantillons de sous-groupe Un identificateur de sous-groupe est requis et vous ne pouvez attribuer qu'une seule variable à cette fonction. Pour trier les données selon les sous-groupes identifiés, activez la case à cocher Trier par sous-groupe.

Nom du rôle	Description
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.
Variables de bloc	indique deux variables maximum, qui groupent les données par blocs de sous-groupes consécutifs. Ces blocs sont libellés dans une légende, et chaque variable de bloc fournit un niveau pour les libellés de la légende.

Cartes de contrôle : Indiquer des limites de contrôle

Dans la liste, cliquez sur **Limites de contrôle** pour accéder à ces options.

Nom d'option	Description
Limites Sigma	spécifie la largeur des limites de contrôle commun multiple de l'erreur type (sigma) des statistiques récapitulatives du sous-groupe tracé dans le graphique. La largeur doit être positive. Le multiple par défaut est 3.
Sélectionner une méthode de calcul	spécifie la méthode de calcul pour les limites de contrôle. Vous avez le choix entre ces options : ■ Calculer les limites de contrôle à partir des données actives ■ Calculer les limites de contrôle à partir de la table sélectionnée Pour désigner la table contenant les limites de contrôle ou les paramètres à partir desquels les limites de contrôle peuvent être calculées, cliquez sur Parcourir. Sélectionnez la table à utiliser et cliquez sur Ouvrir. Remarque : la table que vous sélectionnez doit être une table LIMITS=. Pour plus d'informations sur les

Nom d'option	Description
	variables requises dans une table LIMITS=, consultez la documentation des procédures du graphique que vous êtes en train de créer. Pour représenter les statistiques descriptives de tous les sous-groupes, activez la case à cocher Trace les statistiques descriptives de tous les sous-groupes, que la taille de l'échantillon du sous-groupe soit égale à la taille nominale des limites de contrôle ou non. Vous pouvez représenter ces statistiques lorsque la taille de l'échantillon est la même pour pratiquement tous les sous-groupes et lorsque vous voulez afficher les limites de contrôle fixes (non soumises à variation) correspondant à la taille d'échantillon nominale n. Si vous sélectionnez cette option alors que les tailles des échantillons de vos sous-groupes sont très différentes, l'interprétation des limites de contrôle n'a de sens que pour les sous-groupes dont la taille de l'échantillon est égale à n. Si vous souhaitez ajouter des symboles spéciaux pour les points de données où la taille de l'échantillon est différente de n, activez la case à cocher Ajouter des symboles spéciaux pour les points correspondant aux tailles d'échantillon différentes de la taille nominale pour les limites de contrôle fixes. ■ Spécifier les limites de contrôle Dans les champs Supérieure, Centrale, et Inférieure, indiquez respectivement la limite de contrôle supérieure, la ligne centrale et la limite de contrôle inférieure. Si l'effectif du sous-groupe varie, vous devez indiquer un effectif nominal. Remarque : si vous attribuez des variables à la fonction Grouper l'analyse par, vous ne pouvez pas saisir des valeurs pour la limite de contrôle.

Carte de contrôle : Sélectionner les tests à effectuer

Dans la liste, cliquez sur **Tests** pour accéder à ces options.

Remarque : ces tests ne s'appliquent pas aux limites de contrôle qui ne sont pas trois limites sigma ou qui varient en fonction de la taille de l'échantillon du sous-groupe.

Nom d'option	Description
Sélectionner un test	Utilisez les cases à cocher pour demander un ou plusieurs tests pour des sujets spéciaux, également connus sous les noms de test de dérive : Run test, Pattern et Western Electric tests. Ces tests détectent les dérives particulières non aléatoires dans les points représentés sur le graphique de contrôle principal. L'occurrence du modèle non aléatoire est appelé signal. Les tests peuvent indiquer la présence d'une cause spéciale de variation. Lorsque vous sélectionnez l'un des tests, une description de ce test apparaît dans la zone Description.
Libellé	Utilisez les champs de saisie afin de saisir des libellés pour les points où les tests pour causes spéciales sont positifs. Les libellés ne peuvent contenir plus de 16 caractères. Chaque libellé apparaît à tous les points du test où un signal est rencontré.
Texte identifiant les signaux de test	■ Utilisez la liste déroulante afin de sélectionner une couleur pour les libellés que vous avez indiqués. ■ Utilisez la case à cocher Afficher lignes de zone pour inclure des lignes qui représentent les zones A, B et C sur un graphique principal. Ces zones se construisent en divisant l'intervalle entre les limites de contrôle en six sous-intervalles espacés de façon égale. ■ Utilisez la case Surpasser la limite 3 sigma pour tester des sujets spéciaux si vous spécifiez une limite de contrôle sigma différente du multiple par défaut de 3. Pour changer la limite sigma, cliquez sur Limites de contrôle dans le volet de sélection. ■ Cochez la case Appliquer le test aux dérives des points se chevauchant afin d'appliquer des tests pour les causes spéciales aux dérives de points se chevauchant.

Cartes de contrôle : Définir les options d'apparence du graphe

Certaines options d'apparence sont communes à plusieurs tâches de la carte de contrôle. Pour plus d'informations sur ces options, voir ces rubriques :

- [“Définir les options d'un axe” page 743](#)
- [“Ajouter des lignes de référence” page 744](#)
- [“Définir d'autres options de tracé” page 744](#)
- [“Définir les options de variable de bloc” page 746](#)

Cartes de contrôle : Enregistrer les résultats d'analyse

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Pour enregistrer les résultats de l'analyse, vous pouvez créer des tables de sortie contenant plusieurs statistiques. Activez les cases à cocher pour les données de sortie que vous voulez enregistrer de façon permanente.

Nom du rôle	Description
Statistiques de sous-groupe et table des limites de contrôle	crée une table de sortie contenant les informations représentées sur le graphe. Cela inclut les valeurs de la variable de sous-groupe ainsi que les statistiques descriptives correspondantes et les limites de contrôle. Remarque : cette option crée une table OUTTABLE=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option TABLE= décrite par la procédure SHEWHART. Pour plus d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des statistiques de sous-groupe	crée une table de sortie contenant les statistiques descriptives du sous-groupe. Remarque : cette option crée une table OUTHISTORY=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option HISTORY= décrite par la procédure SHEWHART.

Nom du rôle	Description
	Pour plus d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des limites de contrôle	crée une table de sortie qui enregistre les limites de contrôle. Vous pouvez utiliser cette table lorsque vous sélectionnez la méthode de calcul pour les limites de contrôle lors d'une utilisation ultérieure de cette tâche.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Courbes CDF

<i>À propos de la tâche Courbes CDF</i>	101
<i>Analyse de capacité : Attribuer des variables aux fonctions d'analyse</i>	102
<i>Courbes CDF : Sélectionner la distribution</i>	103
<i>Courbes CDF : Personnaliser la distribution</i>	104
<i>Analyse de capacité : Définir les options d'apparence du graphe</i>	105
Définir les options d'un axe	105
Ajouter des lignes de référence	105
Ajouter un encart	106
Définir d'autres options de tracé	106
Définir les options des limites de spécification	108
<i>Analyse de capacité : Spécifier les types d'analyse</i>	109
<i>Indiquer des titres et des notes de bas de page</i>	110
<i>Afficher les propriétés</i>	110

À propos de la tâche Courbes CDF

Une analyse de capacité de processus compare la distribution des sorties d'un processus sous contrôle statistique à ses limites de spécification afin de déterminer la cohérence avec laquelle ces spécifications sont respectées. Une courbe CDF est utilisée dans l'analyse de capacité de processus pour observer la fonction de distribution cumulative (Cumulative Distribution Function ou CDF, en anglais) d'une variable. La fonction CDF est également appelée fonction de distribution cumulative empirique (Empirical Cumulative Distribution Function ou ECDF, en anglais). Il s'agit d'une fonction croissante par paliers.

Vous pouvez utiliser une courbe CDF pour déterminer le pourcentage d'observations situées en dessous de la limite de spécification inférieure ou au-dessus de la limite de spécification supérieure. Par exemple, supposons qu'une entreprise qui fabrique des câbles à fibres optiques souhaite connaître leur limite de rupture. Vous pouvez créer une table contenant 50 limites de rupture, mesurées en livres par pouce carré. La tâche Courbes CDF utilise les données pour créer une

courbe CDF qui affiche le pourcentage des observations de l'échantillon situées en dessous de la limite de spécification inférieure.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	CAPABILITY
Produits SAS requis	SAS Base, SAS/QC, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Analyse de capacité : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables d'analyse	indique les variables numériques faisant l'objet de l'analyse. Vous devez attribuer au moins une variable à cette fonction. Pour toutes les courbes (sauf la courbe P-P, de probabilités théoriques par rapport aux probabilités calculées) et pour chaque variable d'analyse, vous pouvez indiquer une limite supérieure, une valeur cible et une limite inférieure dans la zone Limites de spécification pour <i>nom_variable</i>. Lorsque vous indiquez ces limites, la sortie inclut automatiquement une table contenant les tests de la normalité, une table contenant les limites de spécification et une table contenant les indices de capacité du processus.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe.

Nom du rôle	Description
	Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Poids relatif	spécifie la variable à utiliser comme variable de poids relatif. Lorsque vous attribuez une variable à cette fonction, la valeur de la variable de chaque observation est utilisée pour calculer les statistiques pondérées. Vous ne pouvez attribuer qu'une seule variable à cette fonction. Remarque : lorsque vous créez un histogramme, une courbe de probabilité, une courbe P-P ou une courbe CDF, il est déconseillé d'attribuer une variable à la fonction Poids relatif. Si vous attribuez malgré tout une variable à cette fonction, le résultat doit être interprété avec précaution.
Variabes de classification (disponible pour la tâche d'histogramme uniquement)	indique une ou deux variables à utiliser en tant que variables de classification. Lorsque vous attribuez une variable à cette fonction, la tâche crée un histogramme comparatif vous permettant de comparer la distribution d'une variable d'analyse sur des niveaux de classification. Remarque : pour créer des courbes de probabilité comparatives ou des courbes Q-Q comparatifs, utilisez la tâche Analyse de distribution.

Courbes CDF : Sélectionner la distribution

Dans la liste, sous l'en-tête Distributions, cliquez sur **Récapitulatif** pour accéder à ces options.

Vous pouvez ajuster une distribution théorique sur une courbe CDF. Après avoir sélectionné la distribution, sélectionnez son nom dans la liste pour en spécifier les valeurs de paramètres et les propriétés de courbe.

Remarque : si vous sélectionnez plusieurs distributions, une courbe CDF est créée pour chacune d'elles.

Nom d'option	Description
Distributions	Utilisez les cases à cocher pour sélectionner une ou plusieurs de ces courbes de densité : ■ Normal ajuste une densité normale avec une moyenne (μ) et un écart-type (σ). ■ Lognormal ajuste une densité lognormale avec un paramètre d'échelle (ζ), un paramètre de seuil (θ) et un paramètre de forme (σ). ■ Exponentiel ajuste une densité exponentielle avec un paramètre d'échelle (σ) et un paramètre de seuil (θ). ■ Weibull ajuste une densité Weibull à trois paramètres avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et un paramètre de forme (c). ■ Bêta ajuste une densité bêta avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et des paramètres de forme (α et β). ■ Gamma ajuste une densité gamma avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et un paramètre de forme (α).
Style des graphiques	Sélectionnez le style des graphiques à utiliser dans la sortie. Vous avez le choix entre les graphiques SAS traditionnels et les graphiques statistiques ODS. Pour plus d'informations sur les graphiques ODS, voir SAS Output Delivery System: User's Guide.

Courbes CDF : Personnaliser la distribution

Dans la liste, sous l'en-tête Distributions, cliquez sur le nom de la distribution que vous voulez personnaliser pour pouvoir accéder à ces options.

Remarque : ces options de distribution ne sont disponibles que si vous avez sélectionné la distribution au préalable sur la page **Récapitulatif**, ou bien si vous activez la case à cocher **Nom de la distribution** sur la page de cette distribution.

Vous pouvez spécifier les paramètres de distribution pour chaque variable d'analyse ou pour toutes les variables d'analyse. Par défaut, la tâche estime les valeurs de paramètre de toutes les variables d'analyse.

Pour définir les paramètres de chaque variable d'analyse :

- 1 Désactivez la case à cocher **Appliquer la distribution à toutes les variables**.
- 2 Activez la case à cocher de la variable dans la zone **Variables d'analyse**.
- 3 Spécifiez les valeurs des paramètres.

Remarque : pour revenir aux valeurs des paramètres estimés, cliquez sur **Utiliser les estimations**.

Analyse de capabilité : Définir les options d'apparence du graphe

Définir les options d'un axe

Dans la liste, sous l'en-tête **Apparence** ⇒ **Axes**, cliquez sur **Axes** pour accéder à ces options.

Vous pouvez définir ces options :

- la couleur des axes
- la couleur des graduations
- un libellé pour les axes horizontaux et verticaux

Ajouter des lignes de référence

Dans la liste, sous l'en-tête **Apparence** ⇒ **Axes**, cliquez sur **Horizontal** ou **Vertical** pour accéder à ces options.

Pour utiliser les lignes de référence, sélectionnez **Utiliser les lignes de référence**. Vous pouvez formater la ligne de référence et spécifier l'emplacement de chaque ligne de référence.

Ajouter un encart

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Encart** pour accéder à ces options.

Vous pouvez améliorer un graphe en ajoutant une case ou une table (appelée encart) contenant les statistiques descriptives directement sur le graphe. Si vous créez un histogramme comparatif, un encart apparaît dans chaque histogramme.

Sélectionnez **Inclure l'encart** pour ajouter des statistiques descriptives au graphe. Sélectionnez ensuite les statistiques que vous voulez inclure dans l'encart.

Remarque : les statistiques permettant de calculer les indices de capabilité ne sont disponibles que si vous avez créé des limites de spécification pour la variable attribuée à la fonction **Variables d'analyse**. Vous pouvez créer des limites de spécification pour les histogrammes, les courbes CDF, les courbes de probabilité et les diagrammes Q-Q. Pour plus d'informations sur le calcul d'indices de capabilité, consultez l'aide relative à la procédure CAPABILITY.

Définir d'autres options de tracé

Dans la liste, sous l'en-tête Apparence, cliquez sur **Options** pour accéder à ces options.

Les options d'apparence disponibles dépendent du type de graphe créé par la tâche.

Nom du rôle	Description
Arrière-plan	spécifie la couleur d'arrière-plan du graphe.
Contour de la barre et remplissage de la barre	spécifie les couleurs pour les barres d'un histogramme. Remarque : ces options ne sont disponibles que pour les histogrammes.
Inverser les deux axes en plaçant les centiles sur l'axe opposé	inverse les axes du graphe. Si vous sélectionnez cette option, les options et les lignes de référence que vous avez sélectionnées pour les axes risquent d'être incorrectes en raison du changement d'orientation du graphe. Remarque : cette option n'est disponible que pour les courbes de probabilité et les courbes Q-Q.
Supprimer légendes par défaut	empêche l'affichage de la légende par défaut avec le graphe.

Nom du rôle	Description
Utiliser des milieux	spécifie les milieux dans l'histogramme. Les options disponibles dans la liste déroulante varient selon que vous créez un histogramme ou un histogramme comparatif. Si vous attribuez une variable à la fonction Variables de classification, vous créez un histogramme comparatif. Pour les histogrammes comparatifs, ces options sont disponibles : ■ Spécifier les milieux permet d'indiquer les milieux de début et de fin. Vous pouvez également spécifier l'intervalle d'escalier à utiliser pour calculer les autres milieux. Le milieu correspond à la valeur située au milieu de la plage de valeurs que représente chaque barre. ■ Clé détermine les milieux des données de la cellule clé. Par défaut, les niveaux de la variable de classification s'affichent de haut en bas (de gauche à droite) dans l'ordre croissant des valeurs internes (non formatées) de la première variable de classification. Si vous n'avez spécifié qu'une seule de ces variables, la cellule clé correspond au niveau qui apparaît le premier dans cet ordre. Si vous en avez spécifié deux, la cellule clé correspond à une combinaison des niveaux des variables 1 et 2 qui apparaissent en premier dans cet ordre. Ainsi, le choix de la cellule clé détermine l'axe horizontal uniforme utilisé pour les autres cellules. La liste des milieux de la cellule clé est donc étendue dans l'une ou l'autre direction si besoin jusqu'à ce qu'elle couvre les données dans les cellules restantes. ■ Uniforme détermine le nombre de milieux en fonction de la taille de l'échantillon complet. Si vous n'attribuez pas de variable à la fonction Variables de classification, vous créez un histogramme. Les options disponibles sont alors les suivantes : ■ Spécifier les milieux permet d'indiquer les milieux de début et de fin. Vous pouvez également spécifier l'intervalle d'escalier à utiliser pour calculer les autres milieux. Le milieu correspond à la valeur située au milieu de la plage de valeurs que représente chaque barre. ■ Indiquer le nombre de classes permet d'indiquer le nombre de classes (également appelées intervalles d'histogrammes) des données. Par défaut, le nombre de classes affichées dépend de l'étendue des données. Cependant, vous pouvez indiquer un maximum. Remarque : vous ne pouvez indiquer des milieux que si vous utilisez la tâche Histogramme.
Indiquer taille du graphique	spécifie la taille de l'image. À titre d'indication, la taille totale de l'écran est affichée dans les zones Largeur et Hauteur.

Nom du rôle	Description
personnalisé (en pixels)	
Remarque : des options de comparaison sont disponibles pour les histogrammes comparatifs. Vous pouvez créer un histogramme comparatif en attribuant une variable à la fonction Variables de classification dans la tâche Histogramme.	
Nom du rôle	Description
Lignes et Colonnes	Si vous attribuez au moins une variable de classification à la fonction d'analyse, vous pouvez utiliser les cases Lignes par page et Colonnes par page pour modifier la disposition des colonnes et des lignes dans l'histogramme comparatif. Si vous attribuez une variable de classification, la disposition par défaut est de deux lignes et d'une colonne par page. Si vous attribuez deux variables de classification, la disposition par défaut est de deux lignes et de deux colonnes par page. Vous pouvez également utiliser les listes déroulantes Côté du cadre et Haut du cadre afin d'indiquer les couleurs utilisées pour remplir les zones de libellés de ligne et libellés de colonne.
Nombre de classes	Par défaut, le nombre de classes affichées dépend de l'étendue des données. Cependant, vous pouvez indiquer un maximum. Vous pouvez également spécifier un écart-type (par exemple, le nombre de classes affichées), sur la base d'une valeur de l'écart-type comprise au-dessus et en dessous de la moyenne des données.

Définir les options des limites de spécification

Remarque : ces options sont disponibles pour les histogrammes, les courbes de probabilité, les graphiques Q-Q et les courbes CDF.

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Limites de spécification** pour indiquer la limite de spécification inférieure, la cible, et la limite de spécification supérieure.

Remarque : les options **Limites de spécification** ne sont disponibles que si vous indiquez une ou plusieurs limites pour au moins une variable d'analyse. Pour

indiquer une limite de spécification pour une variable d'analyse, cliquez sur **Données** dans la liste, puis sélectionnez la variable d'analyse.

Analyse de capacité : Spécifier les types d'analyse

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Par défaut, le rapport Analyse de capacité se compose des tables suivantes : une table d'intervalles de confiance de base, une table de mesures de base, une table de tests de position et une table de moments.

Pour ajouter une table au rapport, sélectionnez la table dans la zone **Tables**.

Pour spécifier les options d'analyse, sélectionnez la table dans la zone **Tables**. Si les options sont disponibles, une case s'affiche sur la droite avec les options à spécifier.

Pour supprimer toutes les tables de statistiques descriptives du rapport, activez la case à cocher **Supprimer stat. descriptives et indices de capacité**. La sélection de cette option ne supprime pas ces tables :

- tables créées par la distribution sélectionnée. Pour supprimer ces tables, activez la case à cocher **Supprimer les tables de distribution** dans le panneau correspondant à cette distribution.
- Tables créées par l'instruction INTERVALS. Pour supprimer ces tables, vous pouvez spécifier l'option NOPRINT. Pour plus d'informations sur l'instruction INTERVALS, consultez l'aide relative à la procédure CAPABILITY.

Pour enregistrer vos résultats dans une table de sortie, activez la case à cocher **Enregistrer les statistiques de sortie dans une table**. La table inclura les statistiques à une variable et les indices de capacité.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Caractéristiques des données

<i>À propos de l'Assistant Caractéristiques des données</i>	113
<i>Caractéristiques des données : Sélectionner les données</i>	114
<i>Caractéristiques des données : Sélectionner les options du rapport</i>	114
<i>Caractéristiques des données : Indiquer des limites pour les sorties</i>	115

À propos de l'Assistant Caractéristiques des données

L'Assistant Caractéristiques des données aide à créer un rapport de synthèse, des graphiques ainsi que des tables SAS de fréquence et statistiques univariées, décrivant les principales caractéristiques des données.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	CONTENTS, FREQ, UNIVARIATE
Produits SAS requis	Base SAS
Autres produits SAS recommandés	SAS/GRAPH

Caractéristiques des données : Sélectionner les données

- 1 Cliquez sur **Ajouter**. La boîte de dialogue **Ouvrir** apparaît.
- 2 Sélectionnez la table à utiliser. Pour sélectionner plusieurs tables, appuyez sur la touche CTRL. Cliquez sur **Ouvrir**. Les tables sélectionnées apparaissent dans la zone **Données**.
- 3 Cliquez sur **Suivant** pour sélectionner les options du rapport.

Caractéristiques des données : Sélectionner les options du rapport

- 1 Avant de sélectionner l'une des options du rapport, vous devez sélectionner les données.
- 2 Sélectionnez les types de sortie. Par défaut, un rapport de synthèse, des graphes et des tables SAS de données de fréquence et de données univariées sont créés.

Si vous sélectionnez **Tables SAS** comme type de sortie, SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **OPTIONS** et enregistre les données de sortie dans la première bibliothèque modifiable de cette liste. Pour indiquer un autre emplacement, cliquez sur **Parcourir**. Les tables obtenues sont également ajoutées au projet.
- 3 Cliquez sur **Suivant** afin de spécifier des limites pour les sorties.

Caractéristiques des données : Indiquer des limites pour les sorties

- 1 Avant de sélectionner les limites, vous devez sélectionner l'une des options du rapport.
- 2 Activez la case à cocher **Limiter les valeurs de catégorie** pour indiquer le nombre maximal de valeurs de catégorie par variable à insérer dans le rapport. Par défaut, 30 valeurs sont ajoutées dans le rapport.
- 3 Cliquez sur **Terminer** pour générer la sortie.

Analyse de classification

<i>À propos de la tâche Analyse de classification</i>	117
<i>Analyse de classification : Attribuer des variables aux fonctions d'analyse</i>	118
<i>Analyse de classification : Définir les options de classification</i>	119
<i>Analyse de classification : Générer des courbes et des diagrammes</i>	120
<i>Analyse de classification : Définition des options des résultats</i>	122
<i>Indiquer des titres et des notes de bas de page</i>	124
<i>Afficher les propriétés</i>	124

À propos de la tâche Analyse de classification

La tâche Analyse de classification crée des classifications hiérarchiques pour les observations d'une table SAS contenant soit des coordonnées soit des distances. Si la table contient des coordonnées, la tâche calcule les distances euclidiennes avant d'appliquer les méthodes de classification. Cette tâche permet de représenter les résultats d'une classification hiérarchique sous forme de dendrogramme.

La tâche Analyse de classification peut aussi créer des classifications non hiérarchiques des coordonnées en utilisant la méthode des nuées dynamiques. Vous ne pouvez pas utiliser la méthode des nuées dynamiques avec des distances. Cette tâche permet d'analyser les données de population.

Supposons que vous souhaitiez déterminer si les taux de natalité, de mortalité et de mortalité infantile nationaux permettent d'établir des types ou des catégories de pays. Vous pouvez effectuer une analyse de classification pour déterminer si les observations permettent d'établir des groupes suggérés par les données.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	CLUSTER, FASTCLUS, TREE
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Analyse de classification : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables d'analyse	effectue des analyses à l'aide des variables que vous avez affectées à cette fonction. Vous devez attribuer au moins une variable à cette fonction.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper des analyses en fonction d'une variable que vous avez déjà sélectionnée pour Variables de l'ensemble 1 ou Variables de l'ensemble 2.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque

Nom du rôle	Description
	fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Libellé d'identification	spécifie les valeurs à utiliser pour identifier les lignes dans historique des classifications et dans la table d'arborescence en sortie. Si aucune variable n'est attribuée à cette fonction, chaque ligne est dénotée par $OBSn$, n étant le numéro de la ligne.
Copier des variables	copie les valeurs de ces variables dans la table de l'arbre de sortie. Les observations de cette table représentant des classifications de plusieurs observations de la table d'entrée comportent des valeurs incomplètes pour les variables de la copie.

Analyse de classification : Définir les options de classification

Dans la liste, cliquez sur **Classification** pour accéder à ces options.

Nom d'option	Description
Méthode de classification	spécifie les méthodes de classification à utiliser dans l'analyse. ■ Distance moyenne minimale calcule la moyenne du groupe à l'aide de la méthode de distance minimale entre centres de gravité des classes, qui utilise les moyennes arithmétiques. ■ Méthode Centroïde utilise la distance minimale entre centres de gravité des classes (UPGMC). ■ Algorithme des nuées dynamiques utilise des distances euclidiennes, les centres de classification étant basés sur l'estimation des moindres carrés. Cette méthode suppose que les centres de classification sont les moyennes des observations qui sont attribuées à chaque classe lorsque l'algorithme est exécuté à des fins de convergence. Chaque itération réduit le critère de stratification jusqu'à ce que la convergence soit réalisée. ■ Méthode de variance minimale de Ward calcule la somme des carrés des erreurs à l'aide de la trace W.

Nom d'option	Description
Options de la méthode des nuées dynamiques	spécifie les options supplémentaires pour l'algorithme des nuées dynamiques. ■ Nombre max. de classifications spécifie le nombre maximal de classes autorisé. ■ Nombre max. d'itérations spécifie le nombre maximal d'itérations pour le recalcul des valeurs initiales de classification. ■ Remplacement de la valeur initiale spécifie le mode de remplacement des valeurs initiales. □ complet : le remplacement complet d'une valeur initiale s'effectue lorsque la distance entre l'observation et la valeur initiale la plus proche est supérieure à la distance minimale entre les valeurs initiales, ou lorsque la distance la plus courte entre l'observation et les valeurs initiales autres que la valeur initiale la plus proche est supérieure à la distance la plus courte entre la valeur initiale la plus proche et les autres valeurs initiales. □ partiel : le remplacement partiel d'une valeur initiale s'effectue lorsque la distance entre l'observation et la valeur initiale la plus proche est supérieure à la distance minimale entre les valeurs initiales □ néant : supprime le remplacement des valeurs initiales. □ aléatoire : le remplacement aléatoire d'une valeur initiale sélectionne un échantillon pseudo-aléatoire simple des observations complètes comme valeurs initiales de classification. ■ Indiquer la valeur initiale aléatoire : spécifie un entier positif comme valeur de début du générateur de nombres pseudo-aléatoires. Remarque : cette option n'est disponible que si vous avez sélectionné Aléatoire comme option de Remplacement de la valeur initiale.

Analyse de classification : Générer des courbes et des diagrammes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Remarque : les courbes et les diagrammes ne sont pas disponibles avec la méthode de classification de l'algorithme des nuées dynamiques.

Une vue graphique du processus de classification peut être très utile pour l'interprétation des classes. Les courbes sont générées à partir d'un sous-ensemble de la source de données en entrée. La façon dont le sous-ensemble est extrait varie selon que la statistique Critère de classification cubique (CCC) peut être calculée ou non. Si la statistique CCC peut être calculée, le sous-ensemble repose sur les valeurs non manquantes de la statistique CCC. Si elle ne peut pas être calculée, le sous-ensemble est extrait en fonction des valeurs non manquantes de la loi de pseudo- F .

Nom d'option	Description
Loi de pseudo F du rang signé	affiche un nuage de points pour le pseudo-F du rang signé par rapport au nombre de classes. Remarque : si vous sélectionnez l'option Loi du t-carré et de pseudo F du rang signé sur le volet Résultats, un graphique de la loi de pseudo F est toujours généré.
Loi du t-carré	affiche un nuage de points pour la loi du pseudo-t^2 par rapport au nombre de classes. Remarque : si vous sélectionnez l'option Loi du t-carré et de pseudo F du rang signé sur le volet Résultats, un graphique de la loi du t carré est toujours généré.
Statistiques CCC	affiche un nuage de points pour la statistique Critère de classification cubique (CCC) par rapport au nombre de classes. Remarque : cette option n'est pas disponible si la table d'entrée est de type DISTANCE.
Orientation du dendrogramme	affiche un dendrogramme des classes. Les feuilles représentent les objets classés, et la classification contenant tous les objets constitue la racine. Une classification contenant au moins deux objets mais pas tous les objets est une branche. Le noeud est le terme général désignant l'ensemble des feuilles, des branches et des racines. Si une classification A est la réunion des classes B et C, A est le parent de B et C, et B et C sont les enfants de A. Par conséquent, une feuille est un noeud sans enfant, et une racine un noeud sans parent. Si chaque classification ne comporte pas plus de deux enfants, le dendrogramme est une arborescence binaire. Sélectionnez Horizontal pour orienter le dendrogramme avec l'axe des y horizontal et la racine à gauche. Sélectionnez Vertical pour orienter le dendrogramme avec l'axe des y vertical et la racine au sommet, soit l'orientation par défaut.

Analyse de classification : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Enregistrer les données de sortie	génère un ensemble de données de sortie. Vous avez le choix entre ces options : ■ Données de l'arborescence des classifications crée une table de sortie pouvant être utilisée pour tracer le dendrogramme de la hiérarchie de classification. La table contient une observation pour chaque observation de la table d'entrée, plus une observation pour chaque classe de deux observations au moins (autrement dit, une observation pour chaque noeud de l'arborescence des classifications). Le nombre total d'observations de sortie est généralement $2n-1$, n étant le nombre d'observations en entrée. ■ Sortie sous forme d'arborescence crée une table de sortie pour la méthode de distance moyenne minimale, la méthode Centroid et la méthode de variance minimale de Ward. Cette table contient une observation pour chaque objet de l'arborescence et les variables CLUSTER et CLUSNAME indiquant l'appartenance de la classification à tout niveau spécifié dans l'arborescence. Vous pouvez également spécifier le nombre de classes ou niveaux à inclure dans la table. ☐ Pour spécifier le nombre de classes à inclure dans la table de sortie, activez la case à cocher Classe. Par défaut, le nombre de clusters est 10. Cependant, le nombre de classes peut être différent du nombre spécifié si l'une de ces conditions est remplie : ■ Le nombre de feuilles de l'arbre est inférieur au nombre de classes voulu. ■ Le nombre d'arbres non connectés dans la table est supérieur au nombre de classes voulu. ■ Un arbre à plusieurs facteurs ne contient pas de niveau avec le nombre de classes spécifié. ☐ Pour spécifier le nombre de niveaux à inclure dans la table de sortie, activez la case Élaguer, puis indiquez le niveau de l'arbre qui définit les classes disjointes dans la table de

Nom d'option	Description
	sortie. Seules les classes entre la racine et la hauteur du niveau spécifié sont affichées. ■ Classification des nuées dynamiques crée une table de sortie pour la méthode de l'algorithme des nuées dynamiques. Cette table de sortie contient toutes les données d'origine, plus les variables CLUSTER et DISTANCE générées par l'analyse de classification. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Afficher la sortie	affiche la sortie. Vous pouvez spécifier : ■ Afficher la sortie affiche des informations sur l'historique du processus de classification. Notamment le nombre de classes, les noms des classes jointes et le nombre d'observations de la nouvelle classification. Dans la zone de saisie Génération de classifications, indiquez le nombre de générations de l'historique à afficher. ■ Statistiques descriptives simples affiche les moyennes, les écarts-types, le coefficient d'asymétrie (Skewness), l'aplatissement (Excess kurtosis) et un coefficient de bimodalité. ■ CCC (Cubic clustering criterion) affiche le CCC (Cubic Clustering Criterion) et la valeur approximative attendue pour R^2 sous l'hypothèse nulle uniforme. Cette option s'applique uniquement aux données des coordonnées. Ces statistiques sont utiles pour déterminer le nombre de classes des données. Des valeurs CCC supérieures à 2 ou 3 indiquent que les classes sont correctes, des valeurs comprises entre 0 et 2 indiquent la présence de classes potentielles à utiliser avec précaution, et des valeurs négatives élevées peuvent indiquer des valeurs hors norme. Remarque : Cette option n'est pas disponible si la table d'entrée est de type DISTANCE. ■ Loi du t-carré et de pseudo F du rang signé affiche la loi du t-carré et de pseudo F. Ces statistiques peuvent être des indicateurs utiles du nombre de classes. Des valeurs relativement élevées de la loi de pseudo F du rang signé indiquent un point d'arrêt. Pour interpréter les valeurs de la loi du t-carré, la règle générale consiste à parcourir la colonne de haut en bas jusqu'à ce que vous trouviez la première valeur

Nom d'option	Description
	significativement plus élevée que la valeur précédente et à remonter la colonne d'une classe.
	Remarque : quand vous sélectionnez cette option, des graphes de la loi de pseudo F du rang signé et de t -carré sont générés si vous sélectionnez les options Loi de pseudo F du rang signé et loi du t -carré sur le volet Graphes .
	Remarque : Afficher la sortie est la seule option d'affichage disponible pour la méthode de l'algorithme des nuées dynamiques.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Comparer les données

<i>À propos de l'Assistant Comparer les données</i>	127
<i>Comparer les données : Sélectionner les données</i>	128
<i>Comparer les données : Sélectionner les variables à comparer</i>	128
<i>Comparer les données : Spécifier les critères de comparaison</i>	129
<i>Comparer les données : Spécifier les options de sortie</i>	129

À propos de l'Assistant Comparer les données

L'Assistant Comparer les données permet de créer un rapport de comparaison entre deux tables ou entre deux variables issues ou non de la même table.

Vous pouvez utiliser l'Assistant Comparer les données pour comparer les modifications apportées à une table SAS. Par exemple, supposons que vous ayez copié une table et que vous souhaitiez connaître les différences entre la copie et la table d'origine. L'Assistant Comparer les données permet d'afficher un résumé des modifications apportées à la table d'origine.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	COMPARE
Produits SAS requis	Base SAS
Autres produits SAS recommandés	Néant

Comparer les données : Sélectionner les données

- 1 Indiquez si vous souhaitez comparer des variables de deux tables distinctes ou d'une même table.
- 2 Sélectionnez les tables de base et de comparaison à utiliser.
- 3 Cliquez sur **Suivant** pour sélectionner les variables à comparer.

Comparer les données : Sélectionner les variables à comparer

Par défaut, toutes les variables possédant les mêmes nom et type sont consignées dans la table. Vous pouvez inclure des variables supplémentaires, modifier les variables à comparer et supprimer des variables.

Pour spécifier les variables à comparer :

- 1 Avant de sélectionner les variables à comparer, vous devez sélectionner les données.
- 2 Dans la liste déroulante **Faire correspondre les données par**, sélectionnez les critères de comparaison des variables. Vous pouvez choisir de faire correspondre les données par observations ou par variables ID. Si vous choisissez de les faire correspondre par variables ID, la variable doit exister à la fois dans les tables de base et de comparaison. Si vous comparez les variables d'une même table, vous ne pouvez les faire correspondre que par observations.
- 3 Pour ajouter une nouvelle paire de variables dans la table, cliquez sur **Nouvelle**. Une nouvelle ligne est ajoutée dans la table. Utilisez les listes déroulantes de chaque cellule pour sélectionner les variables à comparer.

Si l'une des paires de variables est incorrecte, utilisez la liste déroulante de la cellule pour sélectionner le nom de variable approprié. Vous pouvez également cliquer sur **Modifier**.

Pour supprimer une paire de variables, sélectionnez la ligne dans la table et cliquez sur **Supprimer**.

- 4 Cliquez sur **Suivant** pour spécifier les critères de comparaison.

Comparer les données : Spécifier les critères de comparaison

- 1 Avant de spécifier les critères de comparaison, vous devez sélectionner les variables à comparer.
- 2 Spécifiez la valeur du critère d'égalité dans la zone **Critère d'égalité**.
- 3 Sélectionnez la **Méthode d'estimation de l'égalité** dans la liste déroulante de même nom. Les valeurs numériques sont estimées inégales si la magnitude de leur différence, telle qu'elle est déterminée par la méthode sélectionnée, est supérieure à la valeur du critère d'égalité. Les méthodes disponibles sont les suivantes :
 - **Absolue** - compare la différence absolue des valeurs à la valeur du critère d'égalité. Si la valeur absolue de y moins x est supérieure au critère d'égalité, les valeurs sont estimées inégales.
 - **Exacte** - teste l'égalité exacte. Si la valeur de y n'est pas égale à la valeur de x , les valeurs sont estimées inégales.
 - **Pourcentage** - compare la différence de pourcentage absolue à la valeur du critère d'égalité.
 - **Relative** - compare la différence relative absolue à la valeur du critère d'égalité.
- 4 Indiquez la façon de traiter les valeurs manquantes. Vous avez le choix entre ces options : **Traiter une valeur manquante de la table de BASE comme étant égale à n'importe quelle valeur** ou **Traiter une valeur manquante de la table de COMPARAISON comme étant égale à n'importe quelle valeur**.
- 5 Cliquez sur **Suivant** pour spécifier les options de sortie.

Comparer les données : Spécifier les options de sortie

Par défaut, la tâche Comparer les données génère un rapport HTML comportant les statistiques descriptives relatives aux tables, aux variables, aux observations et à la comparaison des valeurs, ainsi que les résultats de la comparaison des valeurs.

Vous pouvez également choisir de créer une table de sortie pour les observations et une table de sortie pour les statistiques descriptives.

- 1 Avant de spécifier les options de sortie, vous devez spécifier les critères de comparaison.
- 2 Sélectionnez **Inclure la table de sortie** pour créer une table de sortie contenant une ligne pour chaque observation correspondante. Cette table contient une colonne pour chaque variable de l'observation, une colonne pour le type d'observation (`_TYPE_`) et une colonne pour le numéro d'observation (`_OBS_`). Les valeurs de la colonne `_TYPE_` peuvent appartenir à l'un des types suivants :
 - **BASE** - les valeurs de cette observation sont issues d'une observation de la table de base.
 - **COMPARE** - les valeurs de cette observation sont issues d'une observation de la table de comparaison.
 - **DIF** - les valeurs de cette observation sont les différences entre les valeurs de la table de base et celles de la table de comparaison.
 - **PERCENT** - les valeurs de cette observation sont les différences de pourcentage entre les valeurs de la table de base et celles de la table de comparaison.

Vous avez la possibilité d'inclure ces observations dans la table de sortie :

- **Écrire une observation pour chaque observation de la table de base** écrit dans la table de sortie les observations de la table de base. La valeur de la colonne `_TYPE_` de la table de sortie est réglée sur **BASE**.
 - **Écrire une observation pour chaque observation de la table de comparaison** écrit dans la table de sortie les observations de la table de comparaison. La valeur de la colonne `_TYPE_` de la table de sortie est réglée sur **COMP**.
 - **Inclure la valeur de la différence** écrit dans la table de sortie les différences entre les valeurs de la table de base et les valeurs de la table de comparaison. La valeur de la colonne `_TYPE_` de la table de sortie est réglée sur **DIF**.
 - **Inclure les valeurs pour les différences en pourcentage** écrit dans la table de sortie les différences en pourcentage entre les valeurs de la table de base et les valeurs de la table de comparaison. La valeur de la colonne `_TYPE_` de la table de sortie est réglée sur **PERCENT**.
 - **Supprimer les observations lorsque les valeurs sont égales** n'insère pas d'observations dans la table de sortie en cas d'égalité de toutes les valeurs des variables.
- 3 Sélectionnez **Inclure la table de sortie pour les statistiques descriptives** pour créer un table de sortie contenant une ligne pour chaque statistique de résumé correspondant à une paire de variables. La table de sortie contient les colonnes suivantes :
 - `_VAR_` - contient le nom de la variable issue de la table de base.
 - `_WITH_` - contient le nom de la variable issue de la table de comparaison.
 - `_TYPE_` - contient le nom de statistique de l'observation.

- **_BASE_** - contient la valeur de la statistique calculée à partir des valeurs de la variable dans la table de base avec des observations correspondantes dans la table de comparaison.
 - **_COMP_** - contient la valeur de la statistique calculée à partir des valeurs de la variable dans la table de comparaison avec des observations correspondantes dans la table de base.
 - **_DIF_** - contient la valeur de la statistique calculée à partir des différences des valeurs de la variable dans la table de base et la variable correspondante dans la table de comparaison.
 - **_PCTDIF_** - contient la valeur de la statistique calculée à partir des différences de pourcentage des valeurs de la variable dans la table de base et la variable correspondante dans la table de comparaison.
- 4 Spécifiez le nombre maximal de différences à imprimer.
 - 5 Cliquez sur **Terminer** pour générer la sortie.

Courbe de niveau

<i>À propos de la tâche Courbe de niveau</i>	133
<i>Courbe de niveau : Sélectionner un type de graphe</i>	134
<i>Courbe de niveau : Attribuer des variables aux fonctions d'analyse</i>	135
<i>Courbe de niveau : Définir les options d'apparence</i>	136
Définir les options de la courbe de niveau	136
Options d'apparence supplémentaires	137
<i>Indiquer des titres et des notes de bas de page</i>	137
<i>Afficher les propriétés</i>	138

À propos de la tâche Courbe de niveau

La tâche Courbe de niveau crée des courbes de niveau pleines, avec motifs ou lissées qui représentent les relations mathématiques entre trois variables numériques. Les courbes représentent les relations tridimensionnelles dans deux dimensions. Les traits ou aires d'une courbe de niveaux représentent les niveaux d'importance (z) correspondant à une position (x, y) sur un plan.

Vous pouvez utiliser cette tâche pour illustrer une occurrence naturelle. Par exemple, vous pouvez créer une courbe de niveau qui illustre le pourcentage d'argile contenu dans les échantillons de sol à différents endroits du site de test. Les axes x et y du graphe représentent un graphique de hauteur de surface à différents endroits x et y. Les courbes de niveau représentent les endroits du plan dont le pourcentage d'argile correspond à celui spécifié dans la légende.

Pour plus d'informations, voir "[Sélectionner un format de sortie pour les graphiques](#)" page 6.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GCONTOUR, G3GRID

Nom du prérequis	Noms des procédures et des produits
Produits SAS requis	SAS Base, SAS/GRAPH
Autres produits SAS recommandés	Néant

Courbe de niveau : Sélectionner un type de graphe

Dans la liste, cliquez sur **Courbe de niveau** pour accéder à ces options.

Nom d'option	Description
Courbe de niveau	utilise des lignes pour représenter la variable (ou le niveau) Z.
Courbe de niveau pleine	remplit chaque courbe de niveau avec une couleur différente. Chaque niveau de contour est représenté par une couleur.
Courbe de niveau pleine avec contour	remplit chaque courbe de niveau avec une couleur différente, puis entoure chaque zone pour indiquer clairement à quel endroit se trouvent les courbes de niveau.
Courbe de niveau avec motif	représente les niveaux au moyen de rectangles avec motifs. Le motif de chaque rectangle est déterminé par le calcul de la moyenne des valeurs de la variable Z des quatre angles du rectangle, et par l'attribution du motif du niveau le plus proche de cette moyenne.
Courbe de niveau lissée	combine les cellules adjacentes avec le même motif afin de former une seule zone de motif.

Courbe de niveau : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Horizontal	La colonne que vous attribuez à cette fonction est la variable d'axe horizontal ou X du graphique.
Vertical	La colonne que vous attribuez à cette fonction est la variable d'axe vertical ou Y du graphique.
Profondeur	La colonne que vous attribuez à cette fonction est la variable de troisième dimension ou d'axe z du graphe.
Grouper les graphiques par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de graphes créés. Un graphique distinct est créé pour chaque valeur unique.

Courbe de niveau : Définir les options d'apparence

Définir les options de la courbe de niveau

Dans la liste, sous l'en-tête Apparence, cliquez sur **Niveau** pour accéder à ces options.

Nom du rôle	Description
Spécifier le nombre de niveaux	Spécifie le nombre de niveaux, c'est-à-dire le nombre de régions entre lesquelles diviser les valeurs de la variable attribuée à la fonction Profondeur (ou variable de l'Axe z).
Générer automatiquement des libellés pour les niveaux	Activez la case à cocher pour générer automatiquement des libellés pour les niveaux du contour. Le libellé de chaque niveau du contour est la valeur de Profondeur pour ce niveau. La légende s'affichera tout de même. Remarque : cette option n'est pas prise en charge par les formats de graphique ActiveX, Image ActiveX, Java et Image Java.
Prétraiter les données avec des options d'interpolation	Si vous exécutez la tâche Courbe de niveau et qu'une erreur dans le journal indique que "moins de la moitié des cellules contiennent des valeurs", cochez cette case avant de relancer la tâche. Le logiciel crée alors une table dont les valeurs des variables horizontales (x et y) forment une grille complète, et interpole la valeur des variables verticales (z) pour chaque point du plan x-y. Vous pouvez choisir la méthode d'interpolation en sélectionnant l'une des options d'interpolation suivantes : ■ Jointure pour utiliser une interpolation linéaire au sein d'un ensemble de régions triangulaires. Cette méthode d'interpolation crée des valeurs dans la plage des valeurs initiales de la variable verticale, mais la surface résultant de l'interpolation peut ne pas être lisse. ■ Sans mise à l'échelle pour empêcher la mise à l'échelle des variables X et Y avec la même étendue avant l'interpolation. ■ Partielle afin d'utiliser un lissage pour estimer les dérivées de l'interpolation. Une courbe spline à deux

Nom du rôle	Description
	variables est ajustée par rapport aux voisins les plus proches et est utilisée pour estimer les dérivées nécessaires. Les contours ne sont pas aussi lisses qu'avec l'option Spline mais cette méthode est moins exigeante en ressources.
	■ Spline pour utiliser une spline à deux variables pour l'interpolation. Cette méthode peut demander beaucoup de temps s'il y a plus de 100 points de données.

Options d'apparence supplémentaires

Certaines options sont communes aux tâches graphiques. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir les options d'un axe" page 729](#)
- ["Ajouter des lignes de référence" page 732](#)
- ["Définir les options de légende" page 733](#)
- ["Personnaliser l'aire" page 734](#)
- ["Définir les options avancées" page 739](#)

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Copier des fichiers

<i>À propos de la tâche Copier des fichiers</i>	139
<i>Copier des fichiers de l'ordinateur local vers un SAS Server</i>	140

À propos de la tâche Copier des fichiers

La tâche Copier des fichiers permet de transférer des fichiers de votre machine locale vers un SAS Server. Vous pouvez également utiliser cette tâche pour copier des fichiers du SAS Server vers votre ordinateur local. La tâche Copier des fichiers fonctionne d'une façon similaire vers une application FTP. Cependant, cette tâche repose sur les protocoles SAS pour réaliser les transferts de fichiers et ne requiert pas un FTP Server.

Cette tâche utilise SAS Integration Technologies pour copier les données de fichiers sur le réseau. La tâche ne repose pas sur PROC DOWNLOAD ni PROC UPLOAD de SAS/CONNECT.

Remarque : cette tâche n'est disponible que dans SAS Enterprise Guide. Vous ne pouvez pas exécuter le code généré par cette tâche en dehors de SAS Enterprise Guide. Si vous exportez un projet SAS Enterprise Guide ou un flux de processus sous forme de programme SAS, les étapes des tâches Copier des fichiers ne sont pas incluses.

Copier des fichiers de l'ordinateur local vers un SAS Server

- 1 Sélectionnez le serveur à utiliser comme source ou destination des fichiers à copier.
- 2 Sélectionner le type de transfert. Vous pouvez télécharger des fichiers de votre ordinateur local vers le SAS Server ou télécharger des fichiers du SAS Server vers votre ordinateur local.
- 3 Spécifiez le chemin complet des fichiers que vous voulez copier. Dans la case **Fichiers source à copier**, vous pouvez utiliser les caractères génériques standards pour faire correspondre plusieurs fichiers. Un astérisque (*) correspond à tous les caractères, dans n'importe quel nombre, avant le prochain caractère non générique. Un point d'interrogation (?) fait correspondre n'importe quel caractère unique qui occupe cette position dans le nom du fichier.

Voici quelques exemples de caractères génériques et les modèles auxquels ils correspondent :

`/u/dept/finance/*.xls`

Correspond à tous les fichiers XLS dans le dossier /u/dept/finance.

`C:\Data\??Sep2012.csv`

Correspond à tous les fichiers dont les noms commencent par deux caractères et se terminent par "Sep2012.csv". Exemples de correspondances : "01Sep2012.csv" et "15Sep2012.csv". Le nom de fichier "199Sep2011.csv" ne correspondrait pas car le préfixe est composé de trois caractères, et non de deux.

`C:\Data\15???2012.*`

Correspond à tous les fichiers dont les noms commencent par "15", suivi de trois caractères, suivis de "2012", suivi d'une extension de fichier. Exemples de correspondances : "15Sep2012.csv" et "15Oct2012.xls".

Remarque : Les caractères génériques ne sont autorisés que dans le nom de fichier et non dans le chemin du dossier. Par exemple, `c:\data\Sept*\report.xls` n'est pas une spécification de fichier source valide pour cette tâche.

- 4 Indiquez si vous voulez copier les fichiers sur votre ordinateur local ou sur un SAS Server. Lorsque vous spécifiez un dossier cible, souvenez-vous de ces règles :
 - Le dossier doit déjà exister sur le système cible. La tâche Copier des fichiers ne crée pas de nouveaux dossiers.

- Vous devez avoir les autorisations de fichiers appropriées pour écrire un contenu vers le dossier.
 - Vous devez spécifier un chemin absolu pour le dossier cible (même lors du téléchargement vers le serveur des fichiers et l'espace de travail SAS est configuré avec un chemin racine de départ de SASUSER ou un autre chemin racine non système).
- 5 Indiquez si vous voulez utiliser les macro-variables dans les spécifications de fichiers ou de dossiers. Activez la case à cocher **Résoudre les macro-variables SAS dans les chemins source et de destination** pour que la tâche Copier des fichiers se comporte de façon plus dynamique en utilisant des macro-variables SAS et des expressions dans le nom de fichier ou le chemin du dossier cible. Lorsque cette option est activée, la tâche Copier des fichiers tente d'évaluer les macro-variables et les fonctions macro de ces champs avant de démarrer l'opération de copie.

Voici quelques exemples de valeurs qui contiennent des références macros et la manière dont elles peuvent se résoudre.

`/u/&SYSUSERID/Data`

Est résolu en un dossier qui contient l'identifiant de l'utilisateur SAS dans le SAS Workspace Server.

`C:\Data%\sysfunc(substr(&SYSUSERID,,3))`

Est résolu en un dossier qui contient seulement une partie de l'identifiant SAS dans le SAS Workspace Server.

`/u/Data/&outputFile.xls`

Est résolu en un nom de fichier XLS où la racine du nom du fichier est stockée dans la macro-variable &OUTPUTFILE. À l'aide de cette macro-variable suppose qu'un processus précédent (peut-être un programme SAS) a défini la macro-variable pour correspondre à un nom de fichier existant.

ATTENTION

Les macro-variables doivent exister dans votre session SAS (soit en tant que macro-variables SAS intégrées soit définies dans un processus précédent). Si les macro-variables ne sont pas spécifiées correctement ou si vous utilisez une syntaxe incorrecte pour une fonction macro, la tâche Copier des fichiers peut générer des erreurs ou des résultats inattendus.

Pour vérifier que les expressions de macros se résolvent correctement, cliquez sur **Tester les valeurs de macro**. Un court programme SAS s'exécute pour résoudre les expressions de macros, et les résultats s'affichent dans une nouvelle fenêtre. Rappel : lorsque vous testez ces valeurs, les expressions de macros sont résolues en valeurs qui reflètent l'état en cours des macro-variables ou les résultats des fonctions macros que vous utilisez. Ces valeurs peuvent être différentes lorsque vous exécutez la tâche plus tard.

- 6 Indiquez s'il faut remplacer des fichiers existants dans le dossier cible. Par défaut, la tâche Copier des fichiers ne remplace pas les fichiers existants. Si un fichier source correspond au nom d'un fichier existant, la tâche ne copie pas ce fichier et ajoute une note au journal.

- 7 Indiquez si la tâche détecte automatiquement et recopie (à l'aide du style de fin de ligne approprié) les fichiers texte qui sont transférés. Les environnements UNIX et Windows utilisent différentes conventions pour les caractères de fin de ligne dans les fichiers texte. Les environnement UNIX utilisent un caractère d'interligne (LF), et les environnements Windows utilisent une interligne (LF) suivie d'un retour chariot (CR). L'absence de caractère CR peut rendre les fichiers texte UNIX difficiles à lire dans les applications d'édition de texte standard telles que le Bloc-notes.

Corrélations

<i>À propos de la tâche Corrélations</i>	143
<i>Corrélation : Attribuer des variables aux fonctions d'analyse</i>	144
<i>Corrélation : Définir les options de corrélation</i>	145
<i>Corrélation : Définition des options des résultats</i>	147
<i>Corrélations : Créer la table de sortie</i>	148
<i>Indiquer des titres et des notes de bas de page</i>	148
<i>Afficher les propriétés</i>	149

À propos de la tâche Corrélations

La corrélation de données est une procédure statistique permettant de décrire une relation entre les variables numériques. La relation est décrite par le calcul des coefficients de corrélation pour les variables.

Vous pouvez utiliser la tâche Corrélations pour obtenir des informations supplémentaires sur les ventes d'un produit. Par exemple, vous pouvez vous en servir pour déterminer s'il existe une relation entre la taille d'un client (exprimée en ventes) et la quantité de produits qu'il achète à votre société.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	CORR
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Corrélation : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables d'analyse	effectue des analyses à l'aide des variables que vous avez affectées à cette fonction. Vous devez attribuer au moins une variable à cette fonction.
Corréler avec	met en corrélation chaque variable affectée à cette fonction avec les variables d'analyse.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper des analyses en fonction d'une variable que vous avez déjà sélectionnée dans Variables d'analyse ou Corréler avec.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Variables partielles	spécifie les variables à utiliser dans la corrélation partielle de Pearson, la corrélation partielle de classement de Spearman, ou le tau-b partiel de Kendall, en fonction du type de corrélation choisi.

Nom du rôle	Description
Poids relatif	spécifie la valeur de la variable dans chaque ligne pour calculer les coefficients de corrélation produit-moment. Vous ne pouvez attribuer qu'une colonne à cette fonction.

Corrélation : Définir les options de corrélation

Dans la liste, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Types de corrélation	spécifie l'analyse de corrélation. ■ Pearson - calcule la corrélation produit-moment de Pearson. Il s'agit de la mesure paramétrique de l'association de deux variables aléatoires continues. La plage des corrélations s'étend de -1 à 1. ■ Hoeffding calcule la mesure de dépendance D de Hoeffding. Il s'agit d'une mesure non paramétrique d'association qui détecte des départs plus généraux d'indépendance. Cette statistique D est 30 fois plus grande que la définition usuelle et couvre la plage comprise entre -0,5 et 1 de telle sorte que seules les valeurs positives importantes indiquent une dépendance. ■ Kendall - calcule le coefficient tau-b de Kendall. Il s'agit d'une mesure non paramétrique d'association basée sur le nombre des concordances et des discordances dans les observations appariées. La concordance intervient lorsque les observations appariées varient ensemble, une discordance intervient lorsque les observations appariées varient différemment. Le coefficient tau-b de Kendall s'étend de -1 à 1. ■ Spearman - calcule la corrélation d'ordre de classement de Spearman. Il s'agit d'une mesure non paramétrique d'association basée sur le classement des valeurs. La plage des corrélations s'étend de -1 à 1. Remarque : les types de corrélation Hoeffding, Kendall et Spearman ne sont pas disponibles si vous attribuez une variable à la fonction Poids relatif. Le type Hoeffding est

Nom d'option	Description
	également indisponible si vous attribuez une variable à la fonction Variables partielles .
Options de la corrélation de Pearson	spécifie les options pour la corrélation de Pearson : ■ Coefficient alpha de Cronbach - calcule le coefficient alpha de Cronbach. Les coefficients qui utilisent les valeurs brutes et les valeurs standardisées (en ajustant les variables pour obtenir une variance unitaire de 1) sont calculés séparément. Pour chaque variable de corrélation, la corrélation est calculée entre la variable et toutes les autres variables restantes. Le coefficient alpha de Cronbach est calculé en utilisant uniquement les variables restantes. Remarque : cette option n'est pas disponible si vous attribuez une variable à la fonction Corréler avec. ■ Covariances - calcule une matrice de covariance. Si vous attribuez une variable partielle, une matrice de covariance partielle est calculée. ■ Sommes des carrés et produits croisés - calcule la somme des carrés et des produits croisés. Si vous attribuez une variable partielle, la matrice SSCP non partielle est calculée. ■ Sommes corrigées des carrés et produits croisés - calcule les sommes corrigées des carrés et des produits croisés. Si vous attribuez une variable partielle, calcule les matrices CSSCP partielle et non partielle. ■ Supprimer les corrélations de Pearson des résultats - exclut les corrélations de Pearson des résultats.
Options de Fisher	utilise les options de Fisher pour demander des limites de confiance et des p-values sous une hypothèse nulle spécifiée pour les coefficients de corrélation en utilisant la transformation z de Fisher. Remarque : ces options sont disponibles si vous sélectionnez Pearson ou Spearman comme type de corrélation. ■ Indiquer la valeur de alpha spécifie le niveau des limites de confiance pour la corrélation. La valeur de cette option doit être comprise entre 0 et 1. La valeur par défaut est 0,05. ■ Type d'intervalle de confiance permet de spécifier le type de l'intervalle de confiance. La valeur par défaut est une limite de confiance bilatérale de l'alternative bilatérale. ■ Indiquer la valeur de rho0 dans l'hypothèse 0 permet d'indiquer la valeur de rho0 dans l'hypothèse nulle. La valeur par défaut est 0. ■ Utiliser l'ajustement biaisé pour la construction des niveaux de confiance - indique si l'ajustement biaisé est utilisé pour la construction de l'intervalle de confiance. Par défaut, cette case à cocher est activée et une nouvelle

Nom d'option	Description
	estimation de corrélation est produite à l'aide de l'ajustement biaisé.
Diviseur pour variance	spécifie le diviseur utilisé pour le calcul des variances, des écarts-types et des covariances. Par défaut, le diviseur utilisé pour le calcul de la variance est le degré de liberté.
Omettre les lignes contenant des valeurs manquantes pour les variables faisant l'objet de la corrélation	utilise uniquement les lignes dans lesquelles aucune des variables d'analyse ne contient de valeur manquante. Par défaut, la tâche Corrélations utilise toutes les paires de valeurs non manquantes pour chaque paire de variables d'analyse. Cela signifie que certaines corrélations sont calculées avec plus d'observations que d'autres.

Corrélation : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Graphes	génère un nuage de points pour chaque paire de corrélations.
Résultats à afficher	spécifie les options pour l'analyse de corrélation. Dans les résultats, vous pouvez inclure ces valeurs : ■ présenter les statistiques descriptives pour chaque variable d'analyse, y compris la moyenne, l'écart-type, la somme, le minimum et le maximum. ■ afficher les probabilités de signification associée à chaque corrélation. ■ classer les coefficients de corrélations du plus fort au plus faible. Pour chaque variable d'analyse, les variables Corréler avec sont répertoriées dans l'ordre de la valeur absolue décroissante du coefficient de corrélation. Par défaut, les coefficients de corrélation ne sont pas classés. Si vous choisissez de classer les coefficients de corrélation, vous pouvez également indiquer le nombre de corrélations à

Nom d'option	Description
	afficher pour chaque variable. Si vous spécifiez un nombre de corrélations inférieur au maximum, la tâche Corrélations classe les corrélations de la plus élevée à la plus petite en valeur absolue et renvoie les n meilleures corrélations, où n est le nombre de corrélations que vous avez spécifié.

Corrélations : Créer la table de sortie

Dans la liste, cliquez sur **données de sortie** pour accéder à ces options.

Activez la case à cocher **Enregistrer les données de sortie**. Les statistiques de corrélation incluses dans la sortie finale dépendent des types de corrélation que vous avez sélectionnés dans la zone **Types de corrélation**.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Créer un format

<i>À propos de la tâche Créer un format</i>	151
<i>Créer un format : Définir les options d'un format</i>	152
<i>Créer un format : Définir le nouveau format</i>	153
Comment définir un nouveau format	153
Exemple 1 : format alphanumérique	154
Exemple 2 : format numérique	155
<i>Afficher les propriétés</i>	156

À propos de la tâche Créer un format

La tâche Créer un format crée un format numérique ou alphanumérique. Un format numérique spécifie le texte qui apparaît lors de l'affichage ou de l'impression des valeurs enregistrées d'une variable numérique. Un format alphanumérique réalise la même opération avec une variable alphanumérique. Cette tâche permet de formater la sortie.

Vous pouvez utiliser cette tâche pour :

- imprimer des valeurs numériques sous forme de valeurs alphanumériques (par exemple, imprimer 1 pour MASCULIN et 2 pour FÉMININ)
- imprimer une chaîne de caractères sous forme d'une autre chaîne de caractères (par exemple, imprimer OUI pour YES)
- imprimer des valeurs numériques en utilisant un modèle (par exemple, imprimer 9458763450 sous la forme 945-876-3450).

Remarque : cette tâche n'est disponible que dans SAS Enterprise Guide.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	FORMAT
Produits SAS requis	Base SAS
Autres produits SAS recommandés	Néant

Créer un format : Définir les options d'un format

Dans la liste, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Nom du format	spécifie le nom du format que vous créez. Ce nom sera utilisé comme référence au format chaque fois que vous l'appliquerez. Les noms de format doivent être conformes à ces règles : ■ Ils peuvent uniquement contenir des lettres, des chiffres et des traits de soulignement ; les autres caractères sont interdits. GENDER et M_OR_F sont des noms de formats valides, mais M-OR-F et M#F ne sont pas valides. ■ Ils ne peuvent pas commencer ou se terminer par un chiffre. HEX2DEC est un nom de format valide, mais 4MAT ne l'est pas. ■ Leur longueur maximale dépend du type du format. Le nom d'un format alphanumérique ne peut contenir plus de 31 caractères. Le nom d'un format numérique ne peut contenir plus de 32 caractères. Remarque : si la version de SAS exécutée est antérieure à la version SAS 9, la taille maximale d'un nom du format caractère est de sept caractères et la taille maximale d'un nom de format numérique est de huit caractères. L'utilisation d'un nom de format plus long peut entraîner des résultats inattendus. Le nom HOMMEOUFEMME est trop long, quel que soit le type du format. STATCODE est un nom valide pour le format numérique, mais est trop long pour le format alphanumérique.
Spécifier la largeur du format	spécifie la largeur du format. Par défaut, la largeur correspond à la valeur formatée la plus longue.

Nom d'option	Description
Emplacement d'enregistrement du format	sélectionnez le serveur et la bibliothèque dans laquelle vous voulez stocker le nouveau format. Par défaut, le format sera stocké dans la bibliothèque EGTASK sur le serveur sélectionné, si cette bibliothèque existe. Si cette bibliothèque n'existe pas, le format est stocké dans la bibliothèque SASUSER. Dans les deux cas, le format est conservé lorsque vous quittez l'application et reste disponible pour toute utilisation ultérieure. Remarque : si vous choisissez de stocker le format dans la bibliothèque WORK, il sera supprimé lorsque vous quitterez l'application. Un format peut exister et être néanmoins indisponible. Un format n'est disponible que si la bibliothèque dans laquelle il est stocké se trouve dans le chemin de recherche du format.
Type de format	indiquez si vous créez un format alphanumérique ou numérique. Pour éviter l'imprécision inhérente à l'enregistrement en virgule flottante, un format numérique est appliqué avec un facteur flou. Par défaut, le facteur flou est $1E-12$. Pour changer le facteur flou, activez la case à cocher Spécifier le facteur flou et indiquez une nouvelle valeur dans la zone.

Créer un format : Définir le nouveau format

Comment définir un nouveau format

Dans la liste, cliquez sur **Définir des formats** pour accéder à cette page.

Utilisez cette page pour mettre en correspondance les valeurs stockées d'une variable dans le texte qui apparaîtra quand vous afficherez ou imprimerez ces valeurs. La définition de format comporte deux parties. La colonne **Libellé** permet de spécifier le texte qui sera affiché ou imprimé et la colonne **Intervalles** permet de spécifier un ou plusieurs intervalles de valeurs stockées qui seront transformés en texte. Une zone en rouge sur un intervalle indique un conflit entre cet intervalle et un autre. Par exemple, un conflit survient si vous faites correspondre tous les chiffres compris entre 0 et 5 avec le libellé LOW, et tous les chiffres entre 3 et 7 avec le libellé MEDIUM.

Pour définir un nouveau format

- 1 À côté de la zone **Définition du format**, cliquez sur **Nouveau**. Une nouvelle ligne est ajoutée dans les zones **Définition du format** et **Définition de l'intervalle**.
- 2 Dans la colonne **Libellé**, entrez le texte à afficher ou à imprimer
- 3 Dans la zone **Définitions de l'intervalle**, indiquez les valeurs stockées qui seront mises en correspondance dans ce texte. Pour indiquer une définition d'intervalle :
 - a Sélectionnez un type pour la définition de l'intervalle (**Discrète** ou **Intervalle**).
 - b Entrez les valeurs à mettre en correspondance avec le libellé dans la définition du format. Chaque valeur discrète ou intervalle doit être saisi en tant qu'élément séparé.

Exemple 1 : format alphanumérique

Vous définissez un format alphanumérique pour une variable nommée GENDER comportant les quatre valeurs stockées suivantes : f, F, m, M. Vous souhaitez que le logiciel affiche "femme" pour les valeurs stockées f et F, et "homme" pour les valeurs stockées m et M. La page **Définir les formats** doit se présenter de la façon suivante :

Créer un format

Options
Définir des formats
Propriétés

Définir des formats

Définition de format :

Libellé	Intervalles
féminin	"f" "F"
masculin	"m" "M"

Définitions d'intervalle pour "masculin":

Type	Valeurs
Discrete	"m"
Discrete	M

Aperçu du code Exécuter Enregistrer Annuler Aide

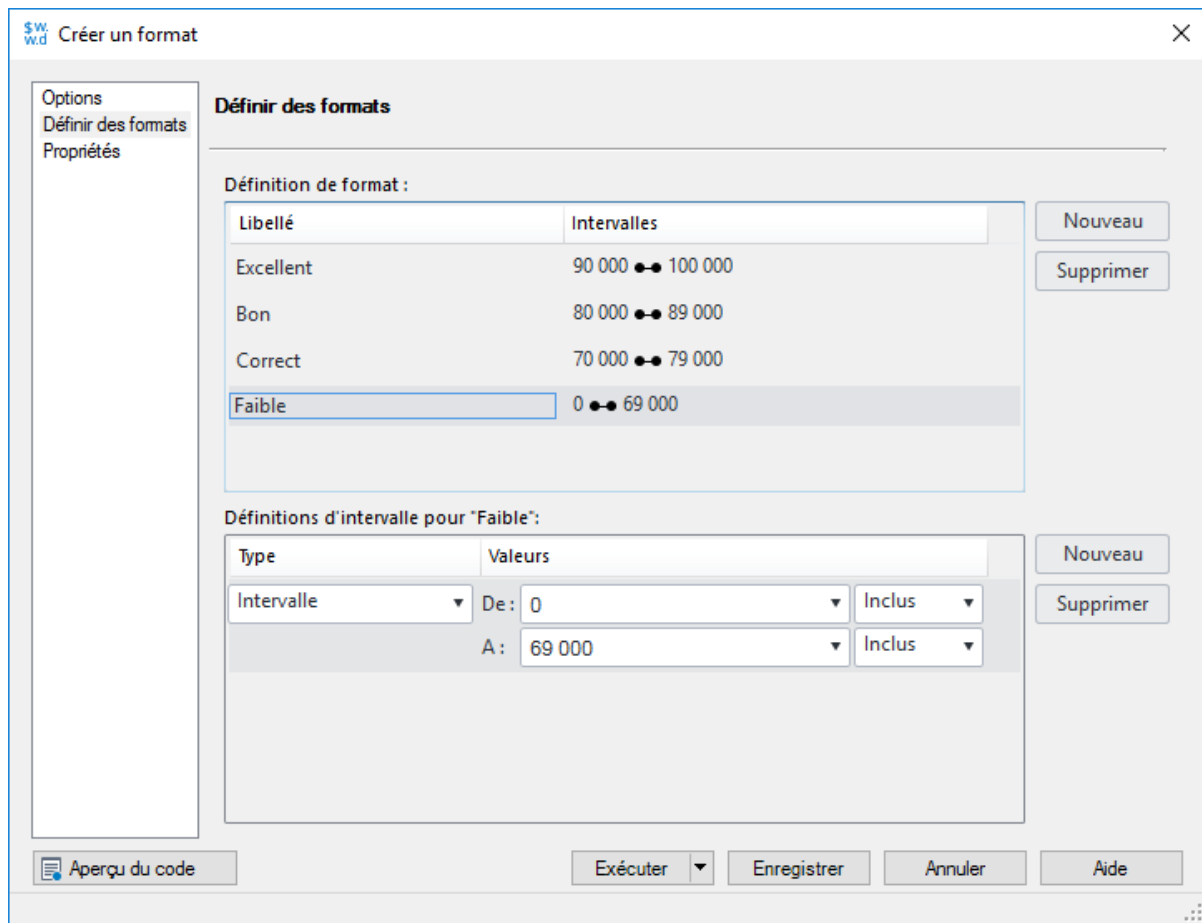
Veuillez saisir un nom de format valide. Vous trouverez dans l'aide des informations sur les noms de format.

Exemple 2 : format numérique

Vous définissez un format numérique pour une variable nommée SALES. Vous voulez que le logiciel affiche une notation en fonction de l'intervalle des valeurs de ventes.

- Excellent = 90 000–100 000 (inclus)
- Bon = 80 000–89 000 (inclus)
- Correct = 70 000–79 000 (inclus)
- Faible = 0–69 000 (inclus)

La page **Définir des formats** doit se présenter comme celle-ci :



Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,

- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Créer un format à partir d'une table

À propos de la tâche Créer un format à partir d'une table	159
Créer un format à partir d'une table	161

À propos de la tâche Créer un format à partir d'une table

La tâche Créer un format à partir d'une table permet de créer un format SAS en utilisant les données enregistrées dans une table SAS. Un format SAS permet de mapper une valeur de données brutes en valeur formatée. Vous pouvez créer des formats SAS pour les données alphanumériques ou numériques.

Voici des exemples de situations où utiliser la tâche Créer un format à partir d'une table :

- Votre société a l'intention d'organiser, l'année prochaine, des groupes orientés clients dans chaque région (Nord, Sud, Est, Ouest) des États-Unis. Vos données de ventes 2009 répertorient l'adresse et l'état de chaque client. Mais vous voulez que ces données soient classées par catégories selon leur région. Dans la table REGIONS, vous avez déjà fait correspondre les états avec la région correspondante :

État	Région
Arizona	Ouest
Californie	Ouest
Floride	Sud

État	Région
Vermont	Nord
Virginie	Est

En utilisant la table REGIONS comme entrée dans la tâche Créer un format à partir d'une table, vous pouvez créer un format SAS que vous pourrez appliquer ultérieurement sur les valeurs d'état dans vos données relatives aux ventes.

- Récemment, votre société a mené plusieurs essais cliniques. Les données concernant les patients incluent l'âge de chaque participant. Dans la table AGES, vous avez déjà divisé les tranches d'âge en catégories discrètes :

Age	Faible	Élevée
Nourisson	0	2
Enfant en bas âge	3	4
Enfant	5	9
Pré-adolescents	10	12
Adolescents	13	17
Adulte	18	100

Remarque : cette tâche n'est disponible que dans SAS Enterprise Guide.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	FORMAT
Produits SAS requis	Base SAS
Autres produits SAS recommandés	Aucun

Créer un format à partir d'une table

- 1 Sélectionnez la table d'entrée. Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Table source**.

- 2 Saisissez un nom pour le format.

Les noms de formats doivent respecter les règles suivantes :

- Ils peuvent uniquement contenir des lettres, des chiffres et des traits de soulignement ; les autres caractères sont interdits. GENDER et M_OR_F sont des noms de formats valides, mais M-OR-F et M#F ne sont pas valides.
- Ils ne peuvent pas commencer ou se terminer par un chiffre. HEX2DEC est un nom de format valide, mais 4MAT ne l'est pas.
- Leur longueur maximale dépend du type du format. Le nom d'un format alphanumérique ne peut contenir plus de 31 caractères. Le nom d'un format numérique ne peut contenir plus de 32 caractères.

Remarque : si la version de SAS exécutée est antérieure à la version SAS 9, la taille maximale d'un nom du format caractère est de sept caractères et la taille maximale d'un nom de format numérique est de huit caractères. L'utilisation d'un nom de format plus long peut entraîner des résultats inattendus. Le nom HOMMEOUFEMME est trop long, quel que soit le type du format. STATCODE est un nom valide pour le format numérique, mais est trop long pour le format alphanumérique.

- 3 Indiquez si vous créez un format alphanumérique ou numérique. Pour éviter l'imprécision inhérente à l'enregistrement en virgule flottante, un format numérique est appliqué avec un facteur flou. Par défaut, le facteur flou est **1E-12**.
- 4 Sélectionnez la bibliothèque dans laquelle vous voulez stocker le nouveau format. Par défaut, le format est stocké dans la bibliothèque Work sur le serveur sélectionné. Cependant, si vous le stockez dans la bibliothèque Work, le format est supprimé lorsque vous quittez SAS Enterprise Guide. Si vous voulez le rendre disponible pour une utilisation future, sélectionnez une autre bibliothèque de sortie.

Remarque : un format peut exister et être néanmoins indisponible. Un format n'est disponible que si la bibliothèque dans laquelle il est stocké se trouve dans le chemin de recherche du format.

- 5 Indiquez si les valeurs de la table d'entrée sont des valeurs discrètes ou bien un intervalle de valeurs.

Pour créer un format à partir de valeurs discrètes :

- a Dans le champ **Types de valeurs**, sélectionnez **Valeurs discrètes/de recherche**.
- b Dans la zone **Variables** de la liste déroulante **Valeurs discrètes**, sélectionnez la variable dans la table qui contient les valeurs à formater. Les valeurs de la variable sélectionnée apparaissent dans la colonne **Valeur de recherche** dans la table de l'aperçu.
- c Dans la liste déroulante **Libellés**, sélectionnez la variable dans la table qui contient les libellés pour les valeurs. Les valeurs de la variable sélectionnée apparaissent dans la colonne **Libellé** dans la table de l'aperçu.

Pour créer un format à partir d'un intervalle de valeurs :

- a Dans la case **Types de valeurs**, sélectionnez **Intervalles**.
 - b Dans la zone **Variables** de la liste déroulante **Valeurs inférieures**, sélectionnez la variable dans la table qui contient les valeurs de début pour l'intervalle. Les valeurs de la variable sélectionnée apparaissent dans la colonne **Début** de la table de l'aperçu.
 - c Dans la liste déroulante **Valeurs supérieures**, sélectionnez la variable dans la table qui contient les valeurs de fin de l'intervalle. Les valeurs de la variable sélectionnée apparaissent dans la colonne **Fin** de la table de l'aperçu.
 - d Dans la liste déroulante **Libellés**, sélectionnez la variable dans la table qui contient les libellés pour les valeurs. Les valeurs de la variable sélectionnée apparaissent dans la colonne **Libellé** dans la table de l'aperçu.
- 6 Spécifiez la longueur maximale du libellé. Par défaut, le libellé peut comporter huit caractères.
 - 7 Spécifiez un libellé pour d'autres valeurs. Lorsque vous appliquez ce format à une source de données, SAS Enterprise Guide peut trouver des valeurs dans les données qui n'ont pas été prises en compte. Vous pouvez créer un libellé unique pour toutes les valeurs de la source de données qui n'ont pas une valeur formatée correspondante. Par défaut, le libellé de ces valeurs est **Autre**.
 - 8 Indiquez si vous souhaitez inclure un rapport de synthèse dans vos résultats. Par défaut, il n'est pas créé lorsque vous exécutez la tâche.

Créer une table de caractéristiques cartographiques

<i>À propos de la tâche Créer une table de caractéristiques cartographiques</i>	163
<i>Créer une table de caractéristiques cartographiques : Attribuer des variables aux fonctions d'analyse</i>	164
<i>Créer une table de caractéristiques cartographiques : Définir les options</i>	165
Définir les options de la projection des données cartographiques	165
Définir les options de la table des caractéristiques	166
<i>Créer une table de caractéristiques cartographiques : Spécifier les résultats</i>	167
<i>Indiquer des titres et des notes de bas de page</i>	167
<i>Afficher les propriétés</i>	167

À propos de la tâche Créer une table de caractéristiques cartographiques

La tâche Créer une table de caractéristiques cartographiques crée une table de caractéristiques pour des tables cartographiques personnalisées ou pour d'autres tables que vous utilisez avec la tâche Représentation graphique. Elle vous permet également de projeter une table cartographique, c'est-à-dire de convertir les coordonnées sphériques d'une table cartographique en coordonnées cartésiennes. Cette tâche vous permet d'utiliser l'une des nombreuses techniques de projection cartographique pour projeter les coordonnées d'une table cartographique traditionnelle sur un plan bidimensionnel tout en essayant de réduire les déformations de surface, de distance, de direction et de propriétés de forme de la sphère d'origine.

Remarque : cette tâche requiert la macro %ADDFEAT, fournie avec SAS Version 8.2 et versions ultérieures. Sinon, veuillez contacter le Support Technique SAS qui vous fournira des instructions pour obtenir cette macro.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GPROJECT
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	CARTES

Créer une table de caractéristiques cartographiques : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Remarque : la source de données en entrée de cette tâche doit impérativement comporter une colonne X et une colonne Y contenant les coordonnées de données spatiales de la carte. La bibliothèque Maps contient des tables de correspondances auxquelles une table de caractéristiques est déjà associée. Par exemple, la table US2 est la table de caractéristiques de la carte U.S. (États-Unis).

La bibliothèque Maps étant protégée, vous ne pouvez pas en utiliser les tables comme données d'entrée de la tâche Créer une table de caractéristiques cartographiques. Toutefois, vous pouvez copier les tables de la bibliothèque Maps vers un emplacement non protégé pour vous en servir ensuite dans cette tâche.

Nom du rôle	Description
Variables d'identification	Les variables que vous attribuez à cette fonction sont utilisées pour identifier les sections dans la carte. Une variable d'identification au moins est requise pour à la fois

Nom du rôle	Description
	projeter les données et/ou créer une table des caractéristiques.
Variables d'attribut à conserver	Ce sont les variables de la table d'entrée de cartographie (géométrie) que vous voulez également inclure dans la table des caractéristiques que vous allez créer. Remarque : cette fonction de données n'apparaît que si l'option Créer une table de caractéristiques cartographiques est sélectionnée.
Variable de longitude à projeter	Il s'agit de la variable de la table d'entrée qui contient le point de données de longitude non projeté (sphérique). Remarque : cette fonction de données n'apparaît que si l'option Projeter les données cartog, (des coordonnées sphériques aux coordonnées cartésiennes) est sélectionnée.
Variable de latitude à projeter	Il s'agit de la variable de la table d'entrée qui contient le point de données de latitude non projeté (sphérique). Remarque : cette fonction de données n'apparaît que si l'option Projeter les données cartog, (des coordonnées sphériques aux coordonnées cartésiennes) est sélectionnée.

Créer une table de caractéristiques cartographiques : Définir les options

Définir les options de la projection des données cartographiques

Dans la liste, cliquez sur **Options** pour accéder à ces options.

- 1 Activez la case à cocher **Projeter les données cartog, (des coordonnées sphériques aux coordonnées cartésiennes)**.
- 2 Sélectionnez la méthode de projection appropriée dans la liste déroulante **Méthode de projection**.

- **Projection équivalente d'Albers** est une projection conique de la surface d'une sphère, qui la coupe au niveau de deux parallèles standard de latitude. L'axe du cône coïncide avec le prolongement de l'axe polaire de la sphère. Chaque section de la carte résultante présente un ratio constant par rapport à l'ère de la sphère. En général, la déformation tend à augmenter en direction des pôles de latitude, en dehors des deux parallèles standard. La projection d'Albers est particulièrement utile pour représenter des zones de grande et petite étendue d'est en ouest. Cependant, les parallèles standard doivent être situées du même côté de l'équateur.
 - Dans la **projection conforme de Lambert**, les méridiens de longitude sont des lignes qui rayonnent à partir du sommet d'un cône sécant et les parallèles de latitude sont des cercles concentriques. Elle est idéale pour des cartes de navigation et des représentations présentant une extension est-ouest relativement faible. Cependant, les parallèles standard doivent être situées du même côté de l'équateur.
 - **Projection gnomonique (azimutale)** est une projection plane directe de la surface de la sphère sur un plan imaginaire tangent à la sphère au pôle de projection de la représentation. Par défaut, le pôle de projection est placé au centre de la représentation des données projetées. La déformation augmente avec la distance au pôle de la représentation. Cette projection est bien adaptée à la représentation des zones présentant une faible extension est-ouest.
- 3 Sélectionnez les **Unités des coordonnées** pour les valeurs non projetées de longitude et latitude de la table cartographique d'entrée. Radians est la sélection par défaut.
 - 4 Sélectionnez la **Direction de longitude croissante**. La valeur par défaut est Ouest.
 - 5 Si vous voulez conserver les observations pour lesquelles les valeurs projetées des variables X et Y sont identiques à celles des observations précédentes, activez la case à cocher **Garder les obs. identiques successives**. Par défaut, les observations identiques successives sont supprimées.

Définir les options de la table des caractéristiques

Dans la liste, cliquez sur **Options** pour accéder à ces options.

Pour créer une nouvelle table des caractéristiques avec la table d'entrée des données cartographiques :

- 1 Sélectionnez la case à cocher **Créer une table de caractéristiques cartographiques**.
- 2 Si vous voulez changer le nom de la variable de géométrie, indiquez le nouveau nom dans la zone **Nom de la nouvelle var. de géométrie cartog.** Ce nom ne doit pas commencer par un nombre ni contenir des espaces, des signes de ponctuation ou des caractères spéciaux.

Créer une table de caractéristiques cartographiques : Spécifier les résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Par défaut, cette tâche stocke la carte du projet et les tables de sortie des caractéristiques dans votre bibliothèque SASUSER. Pour désigner un autre emplacement de stockage, utilisez la boîte de dialogue **Enregistrer sous**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Assistant Créer les données d'une série chronologique

<i>À propos de l'Assistant Créer les données d'une série chronologique</i>	169
<i>Assistant Créer les données d'une série chronologique : Sélectionner les données</i>	170
<i>Assistant Créer les données d'une série chronologique : Attribuez des variables aux fonctions</i>	170
<i>Assistant Créer les données d'une série chronologique : Spécifiez des options pour les valeurs de cumul, manquantes et zéro.</i>	171
<i>Assistant Créer les données d'une série chronologique : Sélectionner une autre sortie</i>	172
<i>Assistant Créer les données d'une série chronologique : Indiquez le titre et les notes de bas de page</i>	172

À propos de l'Assistant Créer les données d'une série chronologique

L'Assistant **Créer les données d'une série chronologique** vous aide à convertir des données transactionnelles en séries chronologiques à intervalles fixes. Les données transactionnelles sont des données horodatées collectées au fil du temps de façon irrégulière ou selon une fréquence variable.

Par exemple, vous disposez des données de vente du site web de votre société. Les ventes sont horodatées mais ne sont pas collectées à un intervalle donné. Vous pouvez utiliser cette tâche pour convertir les données en une table de série chronologique.

Assistant Créer les données d'une série chronologique : Sélectionner les données

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Cliquez sur **Suivant** pour attribuer des variables aux fonctions.

Assistant Créer les données d'une série chronologique : Attribuez des variables aux fonctions

- 1 Avant d'attribuer des variables aux fonctions, vous devez sélectionner les données.
- 2 Affectez des variables à ces fonctions requises :
 - **ID temps** - la colonne que vous attribuez à cette fonction est l'ID temps des données. L'ID temps est une variable de la table d'entrée qui contient la valeur date, date-heure ou heure SAS de chaque observation. Elle permet de déterminer la fréquence et l'ordre des données et d'extrapoler les valeurs ID temps des prévisions. Vous ne pouvez attribuer qu'une variable à cette fonction et il doit s'agir d'une variable date, d'une variable date-heure ou d'une variable numérique contenant des valeurs date ou date-heure.
 - **Série chronologique pour** - la colonne que vous attribuez à cette fonction est la variable de modélisation et de prévision. Vous devez attribuer au moins une variable numérique à cette fonction. Utilisez les flèches vers le haut et vers le bas pour indiquer l'ordre des variables.
- 3 (Facultatif) Attribuez une variable à la fonction **Série chronologique distincte pour les valeurs de (Variable de groupe)**. Cette variable groupe les observations dont la variable de sous-groupe (BY) contient la même valeur. L'attribution d'une variable de sous-groupe (BY) permet d'obtenir une analyse distincte pour chaque groupe d'observations. Vous pouvez attribuer des variables alphanumériques et numériques à cette fonction.

L'ordre des variables de sous-groupe (BY) décrit la structure de la hiérarchie. Utilisez les flèches vers le haut et vers le bas pour indiquer l'ordre des variables.

- 4 Cliquez sur **Suivant** pour spécifier les options pour les valeurs de cumul, manquantes et zéro.

Assistant Créer les données d'une série chronologique : Spécifiez des options pour les valeurs de cumul, manquantes et zéro.

- 1 Avant de spécifier ces options, vous devez attribuer des variables aux fonctions.
- 2 Si la tâche a détecté l'intervalle de la table d'entrée, l'intervalle par défaut apparaît dans la liste déroulante **Intervalle**. Vous pouvez modifier l'intervalle, mais la liste déroulante ne contient que les intervalles appropriés à la variable que vous avez attribuée à la fonction **ID temps**.

La détection d'un intervalle dans des sources de données de grande taille peut prendre du temps. Si aucun intervalle n'a été détecté lorsque vous parvenez à cette étape, vous pouvez en choisir un dans la liste déroulante **Intervalle**. Tant qu'aucun intervalle n'a été détecté, la liste contient toutes les valeurs possibles pour les options d'intervalle. Une fois l'intervalle déterminé, un message d'avertissement s'affiche en cas de conflit entre le type de variable que vous avez attribuée à la fonction ID temps et la valeur de l'intervalle. La liste des valeurs d'intervalle change pour n'afficher que les valeurs appropriées au type de variable. Pour plus d'informations sur chacune des statistiques, voir Glossaire de termes statistiques.

- 3 Sélectionnez la variable de la série chronologique dans le panneau **Série chronologique**. Il s'agit des variables que vous avez attribuées à la fonction **Série chronologique pour**.
- 4 Indiquez la statistique permettant d'agrégier les données aux différents niveaux de la hiérarchie dans la liste déroulante **Statistiques cumulées**.
- 5 Indiquez le traitement à appliquer aux valeurs manquantes dans la liste déroulante **Considérer les valeurs manquantes comme**.
- 6 Répétez les étapes 3 à 5 pour chacune des variables de la série chronologique.
- 7 Cliquez sur **Suivant** pour spécifier les options de sortie.

Assistant Créer les données d'une série chronologique : Sélectionner une autre sortie

- 1 Avant de spécifier les options de sortie, vous devez spécifier les options pour les valeurs de cumul, manquantes et zéro.
- 2 Activez la case à cocher **Série chronologique graphique** si vous voulez inclure un graphique des données de la série chronologique dans la sortie.
- 3 Activez la case à cocher **Enregistrer la série chronologique dans une table**.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

- 4 Cliquez sur **Suivant** pour indiquer des titres et des notes de bas de page.

Assistant Créer les données d'une série chronologique : Indiquez le titre et les notes de bas de page

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

- 1 Avant d'indiquer le titre et les notes de bas de page, vous devez spécifier les options de sortie.
- 2 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie. Vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

- 3 Cliquez sur **Terminer** pour exécuter la tâche.

Attributs de la table

<i>À propos de l'Assistant Attributs de table</i>	175
<i>Attributs de la table : Sélectionner les données</i>	176
<i>Attributs de la table : Spécifier la sortie</i>	176

À propos de l'Assistant Attributs de table

L'Assistant Attributs de la table permet de créer un rapport indiquant le nombre d'observations, l'emplacement et la date de création de la table, ainsi que les noms des variables, les libellés, les types et les formats.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	DATASETS
Produits SAS requis	Base SAS
Autres produits SAS recommandés	Néant

Attributs de la table : Sélectionner les données

- 1 Cliquez sur **Ajouter**. La boîte de dialogue **Ouvrir** apparaît.
- 2 Sélectionnez la ou les tables à utiliser. Pour sélectionner plusieurs tables, appuyez sur la touche CTRL. Cliquez sur **Ouvrir**. Les tables sélectionnées apparaissent dans la zone **Données**.
- 3 Cliquez sur **Suivant** pour spécifier les options de sortie.

Attributs de la table : Spécifier la sortie

- 1 Avant de définir les options de sortie, vous devez sélectionner les données.
- 2 Indiquez l'emplacement de la table de sortie. Par défaut, la table de sortie est enregistrée dans la bibliothèque SASUSER. Pour modifier l'emplacement ou le nom de la table de sortie, cliquez sur **Parcourir**. La table de sortie résultante est également ajoutée au projet.
- 3 Sélectionnez le type de sortie. Vous avez le choix entre ces options :
 - **Rapport par défaut** contient la sortie de la procédure DATASETS. Ce rapport comprend les attributs de données suivants : date de création et de dernière modification de la table, nombre d'observations, codage, informations liées au moteur ou à l'hôte (le cas échéant) et liste alphabétique des variables et de leurs attributs.
 - **Rapport amélioré** contient la sortie de la procédure REPORT. Ce rapport affiche les attributs de la table et des variables. À partir de ce rapport, vous pouvez déterminer le type de table, ses dates de création et de modification, le nombre d'observations, les libellés des variables et les types de variable. Vous pouvez trier les lignes de la table de variables par ordre croissant ou décroissant selon les options suivantes :
 - ☐ **Nom de la variable**
 - ☐ **Ordre des variables dans la table**
 - ☐ **Type de variable**
 - ☐ **Format de la variable**
 - ☐ **Libellé de la variable**

- 4 Cliquez sur **Terminer** pour générer la sortie.

Supprimer les tables et les formats

<i>À propos de la tâche Supprimer les tables et les formats</i>	179
<i>Supprimer des tables, des vues et des formats personnalisés</i>	180
<i>Afficher les propriétés</i>	180

À propos de la tâche Supprimer les tables et les formats

La tâche Supprimer les tables et les formats permet de supprimer définitivement les tables, vues et formats. Vous pouvez supprimer plusieurs éléments dans plusieurs bibliothèques et catalogues différents à la fois. Vous pouvez supprimer plusieurs éléments en une seule fois dans des bibliothèques et des catalogues différents.

Remarque : cette tâche est disponible dans SAS Enterprise Guide uniquement.

Pour exécuter cette tâche, les prérequis sont :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	CATALOG, SQL
Produits SAS requis	Base SAS
Autres produits SAS recommandés	Aucun

Supprimer des tables, des vues et des formats personnalisés

- 1 Dans la liste déroulante **Serveur**, sélectionnez le serveur sur lequel sont stockés les tables, les vues et les formats personnalisés.
- 2 Dans le panneau **Bibliothèques et catalogues**, sélectionnez une bibliothèque ou un catalogue. Les tables, les vues ou les formats personnalisés de la bibliothèque ou du catalogue sélectionné apparaissent dans la vue Liste. Par défaut, ces éléments sont triés selon leur nom. Cliquez sur l'en-tête de colonne pour les trier selon leur date de modification ou leur type.
- 3 Activez la ou les cases à cocher des tables, vues ou formats personnalisés à supprimer. Note

Remarque : sous le panneau **Bibliothèques et catalogues** apparaissent le nombre total de tables, de vues et de formats personnalisés sélectionnés, ainsi que le nombre d'éléments sélectionnés dans le dossier en cours.

- 4 Répétez les étapes 2 et 3 jusqu'à ce que vous ayez sélectionné tous les éléments à supprimer.
- 5 Cliquez sur **Exécuter** pour supprimer définitivement les éléments sélectionnés.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,

- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Analyse discriminante

<i>À propos de la tâche d'analyse discriminante</i>	183
<i>Analyse discriminante : Attribuer des variables aux fonctions d'analyse</i>	184
<i>Analyse discriminante : Définir les options de l'analyse</i>	185
<i>Analyse discriminante : Classifier les nouvelles données</i>	187
Indiquer une table d'entrée	187
Indiquer une variable de test	188
<i>Analyse discriminante : Définition des options des résultats</i>	189
<i>Indiquer des titres et des notes de bas de page</i>	190
<i>Afficher les propriétés</i>	190

À propos de la tâche d'analyse discriminante

Appliquée à un ensemble d'observations contenant une ou plusieurs variables quantitatives et une variable de classification qui définit les groupes d'observations, la tâche Fonction discriminante développe un critère discriminant qui classe chaque observation dans l'un des groupes. Le critère discriminant déduit de cette table peut être appliqué à une seconde table pendant la même exécution de la fonction discriminante.

Cette tâche peut être utilisée pour classer des observations naturelles. Vous pouvez, par exemple, créer une table SAS qui contient l'espèce, le poids, trois longueurs différentes, la hauteur et la largeur de 159 poissons. Vous utilisez ensuite ces données pour rechercher une fonction discriminante sur la base de ces six variables afin de déterminer le meilleur classement pour chaque poisson dans une des sept espèces.

Remarque : cette tâche est disponible dans SAS Enterprise Guide uniquement.

Pour exécuter cette tâche, les prérequis sont :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	DISCRIM
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Analyse discriminante : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables d'analyse	effectue des analyses sur chaque variable de la liste. Vous devez attribuer au moins une variable à cette fonction.
Variable de classification	définit les groupes pour l'analyse. Les niveaux de classification sont déterminés par les valeurs formatées de cette variable. Vous devez attribuer une seule variable à cette fonction.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper des analyses en fonction d'une variable qui a déjà été sélectionnée pour les fonctions Variable d'analyse ou Variable de classification.

Nom du rôle	Description
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Libellé d'identification	affiche les valeurs de la variable de classification plutôt que le numéro d'observation. L'attribution d'une variable à cette fonction n'est effective que si vous activez la case à cocher Résultats de la classification de substitution pour chaque observation . Vous ne pouvez attribuer qu'une seule variable à cette fonction.
Poids relatif	utilise les valeurs de cette variable pour pondérer les valeurs correspondantes de la variable d'analyse. Vous ne pouvez attribuer qu'une seule variable à cette fonction.

Analyse discriminante : Définir les options de l'analyse

Dans la liste, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Résultats de synthèse de la classification à validation croisée	Classe chaque observation de la table d'entrée en utilisant une fonction discriminante calculée à partir des autres observations de la table, et sans l'observation en cours de classification. Cette option ne s'applique que si une méthode paramétrique est utilisée (méthode par défaut).
Afficher les résultats de validation croisée pour chaque observation	affiche les résultats de la classification de validation croisée pour chaque observation. Remarque : cette option n'est disponible que si vous activez la case à cocher Résultats de synthèse de la classification de validation croisée .
Résultats de la classification de resubstitution pour chaque observation	affiche les résultats de classification de resubstitution pour chaque observation.

Nom d'option	Description
Estimations des taux d'erreur sur des prob. a posteriori	affiche les estimations des taux d'erreur sur des probabilités a posteriori du critère de classification basé sur les résultats de classification.
Test univarié d'égalité des moyennes de classe	affiche les statistiques à une variable qui sont utilisées pour tester l'hypothèse que les moyennes de classes sont égales dans la population pour chaque variable.
Tests multivariés d'égalité des moyennes de classe	affiche les statistiques multivariées qui sont utilisées pour tester l'hypothèse que les moyennes de classes sont égales dans la population.
Afficher les résultats de la classification de resubstitution	affiche les résultats de classification de resubstitution pour chaque observation. Vous ne pouvez spécifier cette option que lorsque la table d'entrée est une table SAS ordinaire.
Afficher les résultats de classification à validation croisée	affiche les résultats de la classification de validation croisée pour les observations mal classées uniquement.
Réaliser une analyse discriminante canonique	effectue une analyse discriminante canonique.
Limite le nombre de variables canoniques à calculer	indique le nombre de variables canoniques à calculer. Le nombre que vous spécifiez doit être inférieur ou égal au nombre de variables. Si vous définissez cette option à 0, la sortie inclut les corrélations canoniques mais pas les coefficients canoniques, structures ou moyennes. Soit v le nombre de variables d'analyses, et c le nombre de variables de classification. Si vous ne limitez pas le nombre de variables canoniques à calculer, seules les variables canoniques $\min(v, c - 1)$ sont générées. Si vous incluez une table de sortie dans vos résultats, v variables canoniques sont générées. Dans ce cas, les $v - (c - 1)$ dernières variables canoniques ont des valeurs manquantes.
Probabilités a priori	spécifie la méthode de calcul des probabilités a priori. Ces méthodes sont disponibles : ■ Égal indique que les probabilités a priori de l'appartenance au groupe doivent être égales. ■ Tailles d'échantillon proportionnelles indique que les probabilités a priori d'appartenance au groupe doivent être proportionnelles aux tailles d'échantillon.

Nom d'option	Description
	■ Saisissez la probabilité de chaque niveau permet de spécifier la probabilité a priori pour chaque niveau de la variable de classification. Si la somme des probabilités a priori n'est pas égale à un, ces probabilités changent d'échelle proportionnellement pour que la somme soit égale à un. Remarque : la somme des probabilités que vous spécifiez doit être égale à 1. Cliquez sur Réinitialiser les probabilités pour rétablir les probabilités à leurs valeurs d'origine.

Analyse discriminante : Classifier les nouvelles données

Indiquer une table d'entrée

Dans la liste, cliquez sur **Classifier les nouvelles données** pour accéder à ces options.

Vous pouvez sélectionner une table (appelée ici "table d'entrée") contenant des lignes qui vont être classifiées en fonction d'un critère discriminant extrait de la table que vous analysez. La table d'entrée doit contenir toutes les variables que vous avez attribuées à la fonction **Variables d'analyse**.

Par défaut, les résultats de classification de la table d'entrée sont inclus dans les résultats de la sortie. La table d'entrée est constituée des données d'entrée que vous avez utilisées à l'ouverture de la tâche.

Pour indiquer une autre table d'entrée, activez la case à cocher **Nouvelles données à classer** et cliquez sur **Parcourir**.

Si vous avez attribué une variable à la fonction **Variable de sous-groupe (BY)** et si vous voulez classifier chaque groupe BY dans les nouvelles données par fonction discriminante du groupe BY correspondant dans les données d'entrée d'origine, les nouvelles données que vous sélectionnez doivent être triées dans le même ordre que les données d'entrée que vous avez utilisées pour ouvrir la tâche.

Remarque : vous devez attribuer une variable à la fonction **Variable de classification** de façon à activer la case à cocher **Nouvelles données à classer**.

Remarque : vous devez activer la case à cocher **Nouvelles données à classer** pour spécifier une table d'entrée ou pour indiquer une variable de test.

Pour définir les options de la table d'entrée, utilisez ces options :

- Par défaut, les résultats de classification de la table d'entrée sont inclus dans les résultats de la sortie. Désactivez la case à cocher **Afficher les résultats de classification des nouvelles données** pour supprimer les résultats de classification de la table d'entrée.
- Sélectionnez **Table de sortie des résultats de classif. pour les nouvelles données** afin de créer une table de sortie permanente qui contiendra toutes les données de la table d'entrée, plus les probabilités postérieures et la classe dans laquelle chaque ligne est classifiée.
- Sélectionnez **Table de sortie pour estimations de densité de groupe** afin de créer une table de sortie permanente qui contiendra toutes les données de la table d'entrée, ainsi que les estimations de densité propres au groupe pour chacune des lignes.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Après avoir indiqué une table d'entrée, vous pouvez indiquer une variable de test.

Indiquer une variable de test

Dans la liste, cliquez sur **Classifier les nouvelles données** pour accéder à ces options.

Une fois que vous avez indiqué une table d'entrée contenant les données de classification, vous pouvez indiquer une variable de test.

Pour indiquer une variable de test :

- 1 Sélectionnez une variable dans la liste déroulante **Variable de classification dans les nouvelles données**. Cette variable est utilisée pour déterminer si une observation de la table d'entrée est mal classée. La variable que vous indiquez ici doit être du même type (alphanumérique ou numérique) que la variable attribuée à la fonction **Variable de classification**.

Si vous ne sélectionnez pas de variable ici, c'est la variable attribuée à la fonction **Variable de classification** qui est utilisée.

Remarque : vous pourrez utiliser cette option si la variable qui est attribuée à la fonction **Variable de classification** existe dans la table d'entrée mais a un nom différent de celui de la table d'origine. Par exemple, la variable attribuée à la fonction Variable de classification est 'Sex', mais la variable de classification dans la table d'entrée est 'Gender'.

- 2 Si nécessaire, sélectionnez un nom de variable dans la liste déroulante **Fréquence pour la variable de test**. La variable que vous sélectionnez ici a la même fonction pour la variable de test que la variable attribuée à la fonction **Nombre d'occurrences**.
- 3 Si nécessaire, sélectionnez un nom de variable dans la liste déroulante **Identification du libellé d'essai**. La variable que vous sélectionnez ici a la même fonction pour la variable de test que la variable attribuée à la fonction **Libellé d'identification**. La sélection d'une variable ici n'est effective que si la case à cocher **Afficher les résultats de classification des nouvelles données** est activée.

Analyse discriminante : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Enregistrer les données de sortie	spécifie les résultats à inclure dans une table de sortie. Vous avez le choix entre ces options : ■ La table des Probabilités a posteriori et classification de resubstitution contient vos données d'entrée, plus les possibilités a posteriori et la classe dans laquelle chaque observation est classée par validation croisée. ■ La table des Probabilités a posteriori et classif. de validation croisée contient vos données d'entrée, plus les possibilités a posteriori et la classe dans laquelle chaque observation est classée par validation croisée. ■ La table Statistiques descriptives et fonctions discriminantes contient diverses statistiques comme la moyenne, l'écart-type et la corrélation. ■ La table Grouper estimations de densité spécifique pour chaque obs contient vos données d'entrée, plus les estimations de densité spécifiques à un groupe pour chaque observation,.

Nom d'option	Description
	Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Supprimer les rapports	N'affiche pas les résultats.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Analyse de distribution

<i>À propos de la tâche d'analyse de distribution</i>	193
<i>Analyse de distribution : Attribuer des variables aux fonctions d'analyse</i>	194
<i>Analyse de distribution : Sélectionner les distributions</i>	195
<i>Analyse de distribution : Personnaliser la distribution</i>	196
<i>Analyse de distribution : Générer des graphes</i>	197
<i>Analyse de distribution : Ajouter un encart</i>	199
<i>Analyse de distribution : Indiquer les options des résultats</i>	199
<i>Indiquer des titres et des notes de bas de page</i>	200
<i>Afficher les propriétés</i>	200

À propos de la tâche d'analyse de distribution

La tâche Analyse de distribution fournit des outils d'agrégation de données, ainsi que des informations sur la distribution des variables numériques. Vous pouvez également l'utiliser pour créer différents types de graphe, notamment des histogrammes, des courbes de probabilités, des diagrammes quantile-quantile et des boîtes à moustaches.

Cette tâche permet également de créer les statistiques descriptives d'un produit. Par exemple, supposons que vous ayez stocké dans une table SAS les quotités de financement de 5840 prêts hypothécaires. Vous utiliserez la tâche Analyse de distribution pour créer les éléments suivants :

- une table des mesures récapitulatives, parmi lesquelles les estimations des moments, et une table des observations extrêmes
- un histogramme pour visualiser la distribution des rapports de quotité de financement. En outre, ce graphique met en évidence les caractéristiques de la distribution telles que le coefficient d'asymétrie (Skewness) et la pointe.

- une analyse sur la distribution des données. Cette tâche permet d'exécuter des tests de normalité et de créer des graphiques (une courbe de probabilité, par exemple).

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	UNIVARIATE, BOXPLOT
Produits SAS requis	SAS Base, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Analyse de distribution : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Nom du rôle	Description
Variables d'analyse	Indique les variables numériques faisant l'objet de l'analyse. Une variable d'analyse est requise et vous pouvez attribuer plusieurs variables à cette fonction.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : Vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez sélectionnée pour une fonction d'analyse.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Poids relatif	spécifie la variable à utiliser comme variable de poids relatif. Chaque valeur de la variable de cette fonction est utilisée pour pondérer les valeurs correspondantes des variables que vous

Nom du rôle	Description
	indiquez en tant que variables d'analyse. Vous ne pouvez attribuer qu'une seule variable à cette fonction.
Variables de classification	indique deux variables maximum qui sont utilisées pour grouper les données par niveaux de classification. Si vous choisissez de générer un histogramme, une courbe de probabilité ou un diagramme quantile-quantile, des graphes comparatifs sont créés. Si vous choisissez de générer une boîte à moustaches, des boîtes à moustaches sont créées les unes à côté des autres.

Analyse de distribution : Sélectionner les distributions

Dans la liste, sous l'en-tête Distributions, cliquez sur **Récapitulatif** pour accéder à ces options.

Vous pouvez ajuster plusieurs distributions théoriques. Après avoir sélectionné la distribution pour le calcul, sélectionnez le nom de la distribution dans la liste afin de spécifier les valeurs de paramètres et les propriétés de courbe pour cette distribution.

Utilisez les cases à cocher pour sélectionner une ou plusieurs de ces courbes de densité :

- **Normal** ajuste une densité normale avec une moyenne (μ) et un écart-type (σ).
- **Lognormal** ajuste une densité lognormale avec un paramètre d'échelle (ζ), un paramètre de seuil (θ) et un paramètre de forme (σ).
- **Exponentiel** ajuste une densité exponentielle avec un paramètre d'échelle (σ) et un paramètre de seuil (θ).
- **Weibull** ajuste une densité Weibull à trois paramètres avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et un paramètre de forme (c).
- **Bêta** ajuste une densité bêta avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et des paramètres de forme (α et β).
- **Gamma** ajuste une densité gamma avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et un paramètre de forme (α).
- **Kernel** ajuste l'estimation non-paramétrique de densité de noyau.

Sélectionnez le style des graphiques à utiliser dans la sortie. Vous avez le choix entre les graphiques SAS traditionnels et les graphiques statistiques ODS. Pour

plus d'informations sur les graphiques ODS, voir [SAS Output Delivery System: User's Guide](#).

Analyse de distribution : Personnaliser la distribution

Dans la liste, sous l'en-tête Distributions, cliquez sur le nom de la distribution que vous voulez personnaliser pour pouvoir accéder à ces options.

Remarque : ces options de distribution ne sont disponibles que si vous avez sélectionné la distribution au préalable sur la page **Récapitulatif**, ou bien si vous activez la case à cocher **Nom de la distribution** sur la page de cette distribution.

Pour les distributions normale, lognormale, exponentielle, Weibull, bêta et gamma, vous pouvez soit estimer les paramètres de distribution soit les spécifier. Par défaut, la tâche estime les valeurs de paramètre de toutes les variables d'analyse.

Pour définir les paramètres de chaque variable d'analyse :

- 1 Désactivez la case à cocher **Appliquer la distribution à toutes les variables**.
- 2 Activez la case à cocher de la variable dans la zone **Variables d'analyse**.
- 3 Spécifiez les valeurs des paramètres.

Remarque : pour revenir aux valeurs des paramètres estimés, cliquez sur **Utiliser les estimations**.

Pour les estimations non paramétriques de densité du noyau, vous pouvez indiquer la largeur de bande (c) et la fonction du noyau (k). Par défaut, la tâche estime une largeur de bande qui minimise l'erreur au carré intégrée de la moyenne approximative (MISE). Si vous ne souhaitez pas estimer la largeur de bande, désactivez la case à cocher **Minimiser MISE (Mean Integrated Squared Error)**. Pour spécifier le paramètre de largeur de bande normalisée de la densité du noyau, saisissez la largeur de bande dans le champ de saisie. Vous pouvez saisir jusqu'à cinq valeurs séparées par un espace. Par exemple, vous pouvez entrer ces valeurs dans la zone de texte : 0,5 1,0 1,5.

Remarque : si vous ne désactivez pas la case à cocher **Minimiser MISE (Mean Integrated Squared Error)**, la valeur MISE est considérée comme l'une des cinq valeurs de largeur de bande possibles.

Vous pouvez également définir la fonction du noyau (normale, quadratique ou triangulaire) utilisée pour calculer l'estimation de la densité du noyau. La fonction du noyau par défaut est **Normale**.

Vous pouvez définir des fonctions de noyau pour un maximum de cinq estimations. Si vous définissez plusieurs fonctions, leur ordre doit toujours être le suivant : normale, quadratique, triangulaire. Pour ajouter des fonctions ou en modifier l'ordre, vous devez modifier le code de la tâche.

Si vous définissez plus de fonctions de noyau que de largeurs de bande, la dernière largeur de bande est reprise pour les estimations restantes. Par exemple, si vous définissez les largeurs de bande 0,5 et 1 et les fonctions de noyau normale, quadratique et triangulaire, la largeur de bande 1 est utilisée pour la fonction triangulaire.

Si vous définissez plus de largeurs de bande que de fonctions de noyau, la dernière fonction de noyau est reprise pour les largeurs de bande restantes. Par exemple, si vous indiquez 0,5, 1 et 1,5 pour définir les largeurs de bande, et normale, quadratique et triangulaire pour définir les fonctions de noyau, la fonction quadratique est utilisée pour la largeur de bande 1,5.

Analyse de distribution : Générer des graphes

Dans la liste, sous l'en-tête Apparence, cliquez sur **Graphes** pour accéder à ces options.

Remarque : si une variable de poids relatif est attribuée à une fonction d'analyse, vous ne pouvez pas créer de graphes.

Nom d'option	Description
Courbe d'histogramme	crée un histogramme et superpose les courbes de densité des distributions théoriques continues et des estimations de densité du noyau. Pour sélectionner une distribution, cliquez sur Récapitulatif dans la liste. Si vous attribuez des variables de classification, un histogramme comparatif, composé des histogrammes de la composante, est créé. Les histogrammes de la composante vous permettent de comparer la distribution d'une variable d'analyse avec les niveaux de classification.
Courbe de probabilité	crée une courbe de probabilité pour une distribution théorique continue et superpose une ligne de référence qui correspond à la fois au paramètre de position indiqué ou estimé et au paramètre d'échelle de la distribution théorique. La courbe de probabilité s'utilise pour comparer les valeurs classées d'une variable avec les centiles d'une distribution théorique indiquée. Si la distribution des données correspond à la distribution théorique, les points de la courbe forment un modèle linéaire.

Nom d'option	Description
	Si vous attribuez des variables de classification, une courbe comparative de probabilité, composée des courbes de probabilité de la composante, est créée. Si vous ne sélectionnez pas une distribution, la distribution normale est utilisée.
Diagramme des quantiles	crée une courbe quantile-quantile (courbe Q-Q) pour une distribution théorique continue et superpose une ligne de référence qui correspond à la fois au paramètre d'emplacement spécifié ou estimé et au paramètre d'échelle de la distribution théorique. La courbe Q-Q s'utilise pour comparer les valeurs classées d'une variable aux quantiles d'une distribution théorique indiquée. Si la distribution des données correspond à la distribution théorique, les points de la courbe forment un modèle linéaire. Si vous attribuez des variables de classification, une courbe quantile-quantile comparative, composée des courbes quantile-quantile de la composante, est créée. Si vous ne sélectionnez pas une distribution, la distribution normale est utilisée. Remarque : les courbes Q-Q sont similaires aux courbes de probabilité. Les courbes Q-Q sont préférables pour réaliser une évaluation graphique de paramètres de distribution et d'indices de capabilité, alors que les courbes de probabilité sont mieux adaptées à une estimation graphique des centiles.
Boîte à moustaches	crée une boîte à moustaches. Si vous attribuez des variables de classification, des graphiques de distribution sont créés les uns à côté des autres. Une boîte à moustaches affiche la moyenne, les quartiles, l'observation minimale et l'observation maximale pour un groupe d'observations.
Graphes basés sur du texte	Crée les graphes suivants : ■ une arborescence ou un histogramme (selon le nombre d'observations) ; ■ une boîte à moustaches ; ■ une courbe de probabilité normale ; ■ un graphe adjacent si vous avez attribué une variable à la fonction Grouper l'analyse par.
Couleur de l'axe, Couleur du fond, Largeur de l'axe	personnalise l'apparence du graphique. Remarque : ces options ne sont pas disponibles pour la courbe d'arborescence car il s'agit d'une simple courbe de texte.

Analyse de distribution : Ajouter un encart

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Encart** pour accéder à ces options.

Vous pouvez améliorer un graphe en ajoutant une case ou une table (appelée encart) contenant les statistiques descriptives directement sur le graphe. Si vous créez un histogramme comparatif, un encart apparaît dans chaque histogramme.

Analyse de distribution : Indiquer les options des résultats

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Par défaut, le rapport Analyse de distribution contient les tables suivantes : une table d'intervalles de confiance de base, une table de mesures de base, une table de tests de position et une table de moments.

Nom d'option	Description
Supprimer stat. descriptives et indices de capacité	supprime toutes les statistiques descriptives et les tables d'indices de capacité de la sortie. La sélection de cette option ne supprime pas ces tables : ■ tables créées par la distribution sélectionnée. Pour supprimer ces tables, activez la case à cocher Supprimer les tables de distribution dans le panneau correspondant à cette distribution. ■ Tables créées par l'instruction INTERVALS. Pour supprimer ces tables, vous pouvez spécifier l'option NOPRINT. Pour plus d'informations sur l'instruction INTERVALS, consultez l'aide relative à la procédure CAPABILITY.
Enregistrer les statistiques de sortie dans une table	crée une table de sortie. Les statistiques incluses dans la sortie finale dépendent des tables qui sont sélectionnées dans la zone Tables. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de

Nom d'option	Description
	sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.
	SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir .

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,

- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Anneau

<i>À propos de la tâche Anneau</i>	203
<i>Anneau : Sélectionner un type de graphique</i>	204
<i>Anneau : Attribuer des variables aux fonctions d'analyse</i>	204
<i>Anneau : Définir les options d'apparence</i>	205
Définir les options de l'anneau	205
Définir la présentation	206
Définir les options avancées	207
Options d'apparence supplémentaires	207
<i>Indiquer des titres et des notes de bas de page</i>	208
<i>Afficher les propriétés</i>	208

À propos de la tâche Anneau

La tâche Anneau crée un diagramme simple, ou un groupe de diagrammes, qui représente la contribution relative des différentes parties à un ensemble. Les données se présentent sous la forme de "secteurs", ou parts, d'un cercle. Chaque secteur représente une catégorie de données. La taille d'un secteur représente la contribution des données aux statistiques totales du graphique. Par exemple, un anneau peut afficher les ventes de chaque magasin sous forme de fraction des ventes totales de la chaîne. Un anneau est similaire à un diagramme circulaire, à ceci près que l'anneau présente un trou en son centre.

Pour plus d'informations, voir "[Sélectionner un format de sortie pour les graphiques](#)" page 6.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GCHART

Nom du prérequis	Noms des procédures et des produits
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Anneau : Sélectionner un type de graphique

Dans la liste, cliquez sur **Anneau** pour accéder à ces options.

Nom d'option	Description
Anneau simple	crée un graphique qui affiche la contribution relative des parties à l'ensemble. Les données se présentent sous la forme de "secteurs", ou parts, d'un cercle. Chaque secteur représente une catégorie de données. La taille d'un secteur représente la contribution des données aux statistiques totales du graphique.
Anneau groupé	produit un graphique en anneau distinct pour chaque valeur unique de la variable attribuée à la fonction Grouper anneaux par .

Anneau : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Remarque : les fonctions disponibles dépendent du type de graphique sélectionné.

Nom du rôle	Description
Colonne à représenter	Les valeurs de la colonne que vous attribuez à cette fonction déterminent les différents secteurs d'un anneau. La colonne peut être soit alphanumérique soit numérique.
Grouper les anneaux par	Un anneau est créé pour chaque valeur unique de la colonne que vous attribuez à cette fonction. Tous les graphiques s'affichent sur la même page.
Somme de	La colonne que vous attribuez à cette fonction détermine la taille des sections d'un anneau. Si vous n'attribuez pas de colonne à cette fonction, la fréquence de chaque valeur de la Colonne à représenter est utilisée pour déterminer la taille des sections.
Grouper l'analyse par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de graphes créés. Un graphique distinct est créé pour chaque valeur unique. Chaque anneau s'affiche sur une nouvelle page.

Anneau : Définir les options d'apparence

Définir les options de l'anneau

Dans la liste, sous l'en-tête Apparence, cliquez sur **Anneau** pour accéder à ces options.

Nom d'option	Description
Indiquer les couleurs	indique la couleur des barres. Vous avez le choix entre ces options : ■ Jeu de couleurs par défaut utilise le jeu de couleurs par défaut. Ce dernier est déterminé par le style actuel. Vous pouvez indiquer le style à l'aide du Gestionnaire de styles

Nom d'option	Description
	ou de la boîte de dialogue Propriétés du résultat. Dans SAS Enterprise Guide, vous pouvez également spécifier le style de chaque type de résultat à partir de la boîte de dialogue Options. Dans SAS Add-In for Microsoft Office, vous pouvez appliquer des styles aux formats Rapport SAS et HTML à l'aide de la boîte de dialogue Options de SAS Add-In. ■ Couleurs existantes (Spécifiées dans une version précédente) permet d'utiliser les couleurs qui ont été spécifiées dans une version précédente de SAS Enterprise Guide ou SAS Add-In for Microsoft Office. Cette option n'est disponible que pour le contenu SAS ayant été migré depuis une version antérieure. ■ Couleurs personnalisées permet de spécifier des couleurs personnalisées. En vous servant des listes déroulantes, vous pouvez indiquer jusqu'à 12 couleurs pour un histogramme. Si votre graphique nécessite plus de 12 couleurs, sélectionnez les couleurs supplémentaires dans le jeu de couleurs par défaut actuel. Cliquez sur Réinitialiser pour remplacer les valeurs personnalisées par les valeurs par défaut. Remarque : selon le type de graphique en anneau, il est possible que toutes les couleurs personnalisées ne puissent pas être utilisées.
Spécifier un nombre de secteurs	spécifie le nombre de secteurs dans le graphique. Par défaut, le nombre de secteurs est déterminé automatiquement. Pour les colonnes continues, les valeurs sont divisées par intervalles, à raison d'un secteur par intervalle. Les secteurs qui représenteraient 4 pour cent ou moins du total sont groupés dans un secteur nommé "Autre". Pour indiquer le nombre de secteurs devant apparaître dans le graphique, activez la case à cocher Spécifier un nombre de secteurs.

Définir la présentation

Dans la liste, sous l'en-tête Apparence, cliquez sur **Présentation** pour accéder à ces options.

Indiquez si vous voulez utiliser le nom, le pourcentage, la valeur statistique, ou encore une combinaison des trois dans le libellé de chaque secteur. Pour chacun des composants du libellé, sélectionnez l'une des positions suivantes pour le libellé :

- à l'intérieur du secteur
- à l'extérieur du diagramme circulaire ou de l'anneau
- à l'extérieur du secteur avec une flèche pointant dessus.

Vous pouvez choisir d'utiliser la même couleur de police pour les libellés des secteurs que les secteurs eux-mêmes. Si vous voulez que les libellés apparaissent à l'intérieur des secteurs, n'activez pas cette option.

Vous pouvez indiquer le pourcentage et le libellé pour le secteur "Autre". Ce secteur est l'ensemble de tous les milieux ayant des valeurs statistiques inférieures ou égales au pourcentage du total. La valeur par défaut du pourcentage du total est 4. De ce fait, tout secteur représentant 4% ou moins du total est placé dans le secteur "Autre".

Ce libellé peut être une chaîne de texte comprenant jusqu'à 16 caractères. Le libellé par défaut est Autre. Le secteur "Autre" est le dernier secteur du diagramme, quel que soit l'ordre des secteurs.

Si un seul milieu entre dans la catégorie "Autre", le secteur de ce milieu reste à sa position d'origine sur le graphique et conserve son libellé d'origine. Supposons, par exemple, qu'un diagramme circulaire illustrant le pourcentage de ventes de boissons comporte les secteurs et valeurs suivants : Café 35%, thé 15%, soda 5% et lait 45%. Si le pourcentage du total est 5, le secteur correspondant au soda reste le troisième secteur du graphique plutôt que de devenir le dernier secteur, et il est intitulé "Soda" plutôt que "Autre."

Pour les graphiques groupés, utilisez la zone **Disposition groupée** afin d'organiser les graphiques. Indiquez le nombre de colonnes sur la largeur de la zone de sortie ou le nombre de lignes sur la longueur de la zone de sortie.

Définir les options avancées

Dans la liste, sous l'en-tête Apparence, cliquez sur **Options avancées** pour accéder à ces options.

Sélectionnez dans la liste déroulante la statistique à utiliser pour le calcul des secteurs du graphique. Si vous n'attribuez pas de variable à la fonction **Somme de**, vous pouvez calculer la fréquence ou le pourcentage de chaque secteur. Par défaut, la fréquence est calculée. Si vous n'attribuez pas de variable à la fonction **Somme de**, vous pouvez calculer la somme ou la moyenne de chaque secteur. Par défaut, la somme est calculée.

Options d'apparence supplémentaires

Certaines options sont communes aux tâches graphiques. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir les options de légende" page 733](#)

- “Personnaliser l'aire” page 734

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,

- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Télécharger des fichiers de données vers un PC

<i>À propos de la tâche Télécharger des fichiers de données vers un PC</i>	211
<i>Télécharger des fichiers de données vers un PC : Sélectionner les tables SAS à télécharger</i>	212
<i>Télécharger des fichiers de données vers un PC : Définir les options de téléchargement</i>	212

À propos de la tâche Télécharger des fichiers de données vers un PC

La tâche Télécharger des fichiers de données vers un PC permet de sélectionner une ou plusieurs tables SAS (*.sas7bdat) sur un ou plusieurs SAS Servers et de les copier sur le système de fichiers Microsoft Windows local.

Cette tâche génère un journal qui récapitule les fichiers copiés, leur taille en octets, la temporisation de chaque fichier et de tous les fichiers regroupés, ainsi que les éventuelles erreurs apparues au cours du transfert.

Télécharger des fichiers de données vers un PC : Sélectionner les tables SAS à télécharger

- 1 Cliquez sur **Ajouter** pour sélectionner les tables SAS que vous voulez copier. La boîte de dialogue **Ouvrir** apparaît.
- 2 Accédez au serveur et à la bibliothèque appropriés et sélectionnez le nom du fichier à copier. Pour sélectionner plusieurs fichiers, maintenez la touche CTRL pendant la sélection des fichiers à ajouter.
- 3 Cliquez sur **Ouvrir** pour ajouter les fichiers à la liste des tables à copier.

.....
Remarque : vous pouvez sélectionner des fichiers stockés sur différents SAS Servers en cliquant de nouveau sur **Ajouter** et en sélectionnant des fichiers sur un autre serveur.
.....

.....
Remarque : un nom en caractères minuscules est attribué à chaque fichier copié dans le dossier cible.
.....

- 4 Cliquez sur **Suivant** pour définir les options de téléchargement.

Télécharger des fichiers de données vers un PC : Définir les options de téléchargement

- 1 Avant de définir les options de téléchargement, vous devez sélectionner les tables à télécharger.
- 2 Saisissez le chemin du dossier dans lequel vous voulez copier les fichiers. L'emplacement par défaut est le dossier Mes documents.

- 3 Pour remplacer, sur l'ordinateur local, des fichiers portant le même nom que les fichiers à copier, activez la case à cocher **Remplacer les fichiers existants portant le même nom**. Par défaut, les fichiers existants ne sont pas remplacés.
- 4 Cliquez sur **Terminer** pour copier les tables.

Remarque : si la bibliothèque source est une bibliothèque de moteur SAS Base ne comportant qu'un seul chemin de fichier, les fichiers seront directement transférés depuis le chemin de cette bibliothèque vers le dossier cible. S'il ne s'agit pas d'une bibliothèque de moteur SAS Base ou s'il s'agit d'une bibliothèque concaténée, les fichiers seront dans un premier temps transférés vers une zone de transit de la bibliothèque WORK, puis copiés vers le dossier cible au moyen d'une instruction PROC SQL.

Analyse factorielle

<i>À propos de la tâche Analyse factorielle</i>	215
<i>Analyse factorielle : Attribuer des variables aux fonctions d'analyse</i>	216
<i>Analyse factorielle : Définir les options de la méthode de factorisation</i>	217
<i>Analyse factorielle : Définir les options des facteurs communs</i>	219
<i>Analyse factorielle : Définir les options de rotation et de graphe</i>	221
<i>Analyse factorielle : Définition des options des résultats</i>	223
<i>Indiquer des titres et des notes de bas de page</i>	225

À propos de la tâche Analyse factorielle

L'Analyse factorielle effectue toute une variété d'analyses de facteurs et de composants communs ainsi que de rotations. Les entrées peuvent être des données multivariées, une matrice de corrélation, une matrice de covariance, une représentation des facteurs ou la matrice des valeurs du vecteur du test du Score.

Vous pouvez utiliser cette tâche pour réaliser une analyse de facteur commun sur un entretien annuel des salariés. Supposons, par exemple, que 103 officiers de police ont été évalués par leurs supérieurs sur 14 échelles (variables). Vous pouvez mener une analyse de facteur commun sur ces variables pour voir quels facteurs latents se cachent derrière ces évaluations.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	FACTOR
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	Aucun

Analyse factorielle : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables d'analyse	Indique les variables faisant l'objet de l'analyse. Vous devez attribuer au moins une variable à cette fonction.
Variables partielles	spécifie la variable utilisée pour partialiser les variables à analyser. L'analyse est ainsi basée sur une matrice de corrélation ou de covariance partielle.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : Vous ne pouvez pas grouper la table en fonction d'une variable que vous avez déjà sélectionnée pour Variables d'analyse.
Nombre d'occurrences	indique une variable représentant la fréquence d'occurrence des autres valeurs dans l'observation. Ainsi, la tâche traite chaque ligne comme si elle se produisait n fois, n représentant le nombre d'occurrences de cette observation. Le nombre total d'observations est considéré comme équivalent à la somme de la variable lorsque les degrés de liberté pour les probabilités de signification sont déterminés. Si la valeur de la variable de fréquence est manquante ou inférieure à un, l'observation n'est pas utilisée dans l'analyse. Si la valeur n'est pas un nombre entier, seule la partie entière est utilisée.

Nom du rôle	Description
Poids relatif	spécifie une variable à utiliser comme pondération relative pour chaque observation dans la table d'entrée. Cela se produit souvent lorsque la variance associée à chaque observation est différente et lorsque les valeurs de la variable de poids relatif sont proportionnelles aux réciproques des variances. Si une valeur pondérée relative est négative ou manquante, elle est exclue de l'analyse.

Analyse factorielle : Définir les options de la méthode de factorisation

Dans la liste, cliquez sur **Méthode de factorisation** pour accéder à ces options.

Nom d'option	Description
Méthode de factorisation	spécifie la méthode de factorisation. Vous avez le choix entre ces options : ■ L'option Analyse en composantes principales effectue une analyse des composantes principales si tous les a priori sont égaux à un. Sinon, une analyse des facteurs principaux est effectuée. Il s'agit de la méthode par défaut, sauf pour les tables de type FACTOR. ■ L'option Analyse du composant d'Harris effectue une analyse du composant d'Harris de la matrice de S-IRS-1, une approximation non itérative d'analyse de composantes canonique. ■ Matrice de covariance de l'image procède à une analyse en composantes principales de la matrice de covariance de l'image. Une matrice de corrélation non singulière est requise. ■ Analyse factorielle Alpha effectue une analyse factorielle Alpha. ■ Analyse factorielle de probabilité maximale effectue une analyse factorielle du maximum de vraisemblance avec un algorithme de Fuller, excepté pour les détails mineurs. Une matrice de corrélation non singulière est requise. ■ Analyse factorielle principale itérée effectue une analyse factorielle principale itérée.

Nom d'option	Description
	■ Analyse factorielle des moindres carrés non pondérés effectue une analyse factorielle des moindres carrés non pondérés. ■ Représentation de facteur existante lit une représentation de facteur à partir d'une table de type FACTOR, CORR, UCORR, COV ou UCOV. Il s'agit de la méthode par défaut pour les tables de type FACTOR. ■ Coefficients des valeurs du vecteur du test de score lit les coefficients de score d'une table de type FACTOR, CORR, UCORR, COV ou UCOV. La table doit également contenir une matrice de corrélation ou une matrice de covariance. Une fois que vous avez sélectionné une méthode, vous pouvez définir les options Propriétés numériques et Nombre de facteurs dans la table. Si vous changez au moins deux des options Nombre de facteurs, le nombre de facteurs considérés est le nombre minimum répondant à l'un des critères suivants : ■ Valeur propre la plus petite pour laquelle un facteur est considéré ■ Nombre de facteurs à considérer ■ Pourcentage de variance commune représentée par le facteur considéré
Options de la matrice	spécifie les options de la matrice ■ Factoriser une matrice de covariance au lieu d'une matrice de corrélation utilise la matrice de covariance à la place de la matrice de corrélation pour la factorisation lorsque la méthode de factorisation est une analyse en composantes principales, une matrice de covariance de l'image, une analyse factorielle principale itérée ou une analyse factorielle des moindres carrés non pondérés. ■ Factoriser une matrice pondérée utilise une matrice de corrélation ou de covariance pondérée pour la factorisation lorsque la méthode de factorisation est une analyse en composantes principales, une matrice de covariance de l'image, une analyse factorielle principale itérée, ou une analyse des moindres carrés non pondérés. Remarque : cette option n'est disponible que si la méthode de factorisation est une Analyse en composantes principales, une Matrice de covariance de l'image, une Analyse factorielle principale itérée ou bien une Analyse factorielle des moindres carrés non pondérés, et si la table d'entrée est de type CORR, UCORR, COV, UCOV ou FACTOR.

Nom d'option	Description
	■ Corriger la matrice résultante de la moyenne utilise l'intercept issu de l'analyse ; les covariances ou corrélations sont corrigées pour la moyenne. Par défaut, cette option est sélectionnée.
Diviseur pour variance	spécifie le diviseur utilisé pour le calcul des variances, des écarts-types et des covariances. Par défaut, le diviseur utilisé pour le calcul de la variance est le degré de liberté.

Analyse factorielle : Définir les options des facteurs communs

Dans la liste, cliquez sur **Facteur commun** pour accéder à ces options.

Nom d'option	Description
Estimations des facteurs communs a priori	Sélectionnez l'une de ces options : ■ Proportionnelles aux corrélations multiples au carré définit les estimations des facteurs communs a priori proportionnelles aux corrélations multiples au carré, mais les ajuste de manière que leur somme soit égale aux corrélations absolues maximales. ■ Corrélation multiple carrée avec toutes les autres colonnes définit les estimations des facteurs communs a priori pour chaque variable à sa corrélation multiple au carré avec toutes les autres variables. ■ Corrélation absolue maximale avec d'autres colonnes définit les estimations des facteurs communs a priori pour chaque variable à sa corrélation absolue maximale avec toute autre variable. ■ Définir tous les a priori à un définit tous les facteurs communs a priori à 1,0. ■ Saisie à partir de PRIORS ou COMMUNAL dans la table d'entrée lit les estimations des facteurs communs a priori à partir de la première observation lorsque la variable <code>_TYPE_</code> contient PRIORS ou COMMUNAL et lorsque le type de table est FACTOR. Remarque : cette option n'est disponible que si la méthode de factorisation que vous avez sélectionnée

Nom d'option	Description
	est une analyse en composantes principales, une analyse factorielle Alpha, une analyse factorielle de probabilité maximale, une analyse factorielle principale itérée, ou bien une analyse factorielle des moindres carrés non pondérés, et si la table d'entrée est de type FACTOR. ■ Définir les valeurs a priori par un nombre pseudo-aléatoire entre 0 et 1 définit pour les estimations des facteurs communs a priori des nombres pseudo-aléatoires uniformément distribués entre 0 et 1. ■ Indiquer explicitement les valeurs a priori pour chaque colonne permet de sélectionner l'estimation des facteurs communs pour chaque variable à partir de la zone de saisie, puis de spécifier une nouvelle valeur. Par défaut, l'estimation des facteurs communs pour chaque variable est égale à 0. Remarque : les estimations par défaut des facteurs communs a priori dépendent de la méthode de factorisation que vous utilisez.
Options d'Heywood	Par défaut, l'analyse factorielle Alpha de la méthode de factorisation, l'analyse factorielle de probabilité maximale, l'analyse factorielle principale itérée, ainsi que l'analyse factorielle des moindres carrés non pondérés stoppent les itérations et définissent le nombre de facteurs à 0 si un facteur commun estimé dépasse 1. Vous avez aussi le choix entre ces options : ■ Définir à un tout facteur commun supérieur à un définit à 1 tous les facteurs communs supérieurs à 1, en autorisant les itérations. ■ Autoriser les facteurs communs à dépasser 1 autorise les itérations lorsque les facteurs communs sont supérieurs à 1. Les problèmes de convergence sont possibles parce que les facteurs communs peuvent devenir très élevés et des matrices hessiennes mal conditionnées peuvent se produire.

Analyse factorielle : Définir les options de rotation et de graphe

Dans la liste, cliquez sur **Rotation et graphes** pour accéder à ces options.

Nom d'option	Description
Méthode de rotation	spécifie la méthode de rotation. Vous avez le choix entre ces méthodes. En fonction de la méthode que vous sélectionnez, d'autres options peuvent être disponibles. ■ Aucune rotation indique qu'aucune rotation ne doit être effectuée et, par conséquent, indique de conserver la solution orthogonale de départ. Il s'agit de la méthode par défaut. ■ Equamax orthogonal spécifie une rotation Equamax orthogonal. Cette option revient à utiliser Orthomax général avec une pondération Orthomax égale à $n/2$, où n correspond au nombre de facteurs. ■ Harris-Kaiser case II orthoblique spécifie une rotation Harris-Kaiser case II orthoblique. Vous pouvez spécifier la puissance des racines carrées des valeurs propres utilisées pour modifier l'échelle des valeurs propres, sous l'hypothèse que les facteurs sont extraits par la méthode d'analyse factorielle principale. Pour les autres méthodes d'extraction, les vecteurs propres sont remplacés par les colonnes standardisées de la matrice de facteurs sans rotation, et les valeurs propres sont remplacées par les constantes de normalisation de colonne. La valeur propre par défaut égale à zéro produit la solution de classification indépendante, dans laquelle chaque variable tend à disposer d'un grand chargement sur un seul facteur. ■ Orthomax général spécifie la rotation Orthomax avec une pondération Orthomax par défaut égale à zéro. Il n'existe pas de restriction de valeurs correctes pour la pondération Orthomax, bien que les valeurs les plus courantes soient comprises entre zéro et le nombre de variables. Une pondération Orthomax égale à un se traduit par une rotation Varimax. ■ Parsimax orthogonal spécifie une rotation Parsimax orthogonal. Cette option revient à utiliser Orthomax général avec une pondération Orthomax égale à $[(nvar \times$

Nom d'option	Description
	$(n - 1)/(nvar + n - 2)$], où <i>nvar</i> correspond au nombre de variables et <i>n</i> au nombre de facteurs. ■ Analyse Procruste oblique spécifie la rotation Procrustes oblique avec le modèle cible fourni par la table TARGET. La méthode des moindres carrés non limités est utilisée et l'échelle des facteurs est adaptée à la variance unitaire après rotation. ■ Promax oblique spécifie une rotation Promax oblique. Vous pouvez définir la méthode de pré-rotation orthogonale ou oblique. Vous pouvez également changer la puissance utilisée pour calculer le modèle cible à partir de la valeur par défaut 3. ■ Quartilmax orthogonal spécifie une rotation Quartilmax orthogonal, ce qui revient à utiliser Orthomax général avec une pondération Orthomax égale à zéro. ■ Varimax orthogonal spécifie une rotation Varimax orthogonal, ce qui revient à utiliser Orthomax général avec une pondération Orthomax égale à un.
Méthode de normalisation des lignes du facteur	spécifie la méthode de normalisation des lignes. Vous avez le choix entre ces options : ■ Normalisation de Kaiser utilise la normalisation de Kaiser, qui est la méthode par défaut. ■ Pondération selon Cureton-Mulaik pondère les lignes à l'aide de la technique de Cureton-Mulaik. ■ Modifier l'échelle pour représenter les covar. adapte l'échelle des lignes de la matrice de modèle pour représenter les covariances plutôt que les corrélations. ■ Ne pas normaliser empêche la normalisation de se produire.
Graphes à afficher	spécifie les graphes à inclure dans les résultats. Vous avez le choix entre ces options : ■ Représentation des facteurs (après rotation) affiche une courbe de représentation des facteurs après rotation. Utilisez l'option Nombre de facteurs à tracer pour définir le nombre <i>n</i> de facteurs à représenter de manière à représenter toutes les paires de premiers facteurs <i>n</i> et à produire un total de $n(n - 1)/2$ courbes. Saisissez un nombre entier ou utilisez les flèches vers le haut et vers le bas pour ajuster la valeur par incrément de un. La valeur par défaut est 2. ■ Tracer les caractéristiques du facteur sans rotation affiche une courbe des caractéristiques du facteur sans rotation.

Nom d'option	Description
	■ Afficher un graphe des valeurs propres affiche un graphe des valeurs propres. Par défaut, aucun graphe n'est affiché. Après avoir sélectionné les graphes à afficher, vous pouvez également indiquer le nombre de facteurs à représenter. Vous avez le choix entre ces options : ■ Utiliser tous les facteurs représente tous les facteurs. Il s'agit de la présentation par défaut. ■ Nombre de facteurs à tracer permet d'indiquer le nombre de facteurs à représenter (n). La plus petite valeur possible est 2. Lorsque vous spécifiez une valeur, toutes les paires des n premiers facteurs sont tracées, produisant un total de $n(n - 1)/2$ graphes.

Analyse factorielle : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Enregistrer les données de sortie	Activez les cases à cocher Facteurs ou Statistiques pour les données de sortie que vous voulez enregistrer de façon permanente, dans la zone Enregistrer les données de sortie. Si vous sélectionnez Statistiques, vous pouvez choisir d'exclure la matrice de corrélation de la table de sortie. La matrice de corrélation n'est exclue que si la matrice de corrélation n'est pas requise pour une autre sortie demandée (par ex., si les scores ou les corrélations résiduelles sont affichées) et si vous sélectionnez Représentation de facteur existante comme Méthode de factorisation. Il s'agit de la méthode par défaut pour les tables de type FACTOR. Remarque : l'option Facteurs n'est pas disponible si la table d'entrée est de type CORR, UCORR, COV, UCOV ou FACTOR. Elle n'est pas disponible non plus si une variable a été attribuée à la fonction Variables partielles.

Nom d'option	Description
Afficher la sortie	Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir. Sélectionnez l'une des options suivantes dans la zone Afficher la sortie : ■ Matrice de corrélation résiduelle affiche la matrice de corrélation résiduelle et la matrice de corrélation partielle associée. Les éléments diagonaux de la matrice de corrélation résiduelle sont les uniques variances. ■ Vecteurs propres affiche les vecteurs propres de la matrice de corrélation réduite, dont les éléments diagonaux sont remplacés par les estimations de facteurs communs. Lorsque vous utilisez la méthode de factorisation Analyse factorielle de probabilité maximale, les vecteurs propres sont destinés à la matrice de corrélation réduite pondérée. ■ Coefficients des valeurs du vecteur du test de score affiche les coefficients des valeurs du vecteur du test de score. La corrélation multiple au carré de chaque facteur avec les variables est également affichée, sauf dans le cas de composantes principales sans rotation. ■ Reclasser lignes de la matrice par un chargement max. reclasse les lignes (variables) des différentes matrices de facteurs dans la sortie. Les variables avec leur chargement maximal (chargement de structure de référence pour les rotations obliques) sur le premier facteur sont affichées les premières, du chargement le plus grand au plus petit, suivies des variables avec leur chargement maximal sur le second facteur, etc. ■ Moyennes et écarts-types des colonnes d'entrée affiche les moyennes, les écarts-types et le nombre d'observations. ■ Matrice de corrélation des colonnes d'entrée affiche la matrice de corrélation ou la matrice de corrélation partielle. ■ Mesure d'adéquation de l'échantillonnage de Kaiser produit les corrélations partielles entre chaque paire de variables, en contrôlant toutes les autres variables (corrélations anti-image négatives) et la mesure d'adéquation de l'échantillonnage de Kaiser.

Nom d'option	Description
	■ Représentation des facteurs affiche la représentation des facteurs et les statistiques apparentées. Dans les cas de rotation oblique, les structures de référence et de facteur sont calculées et affichées. Remarque : cette option n'est disponible que si vous avez sélectionné la méthode de représentation de facteur existante. ■ Valeurs du vecteur du test de score affiche les valeurs du vecteur du test de score et les statistiques apparentées. Dans les cas de rotation oblique, les structures de référence et de facteur sont calculées et affichées. Remarque : cette option n'est disponible que si vous avez sélectionné la méthode des coefficients des valeurs du vecteur du test de score. Par défaut, aucun résultat de facteur ni calcul statistique n'est sélectionné.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Créer un projet Forecast Studio

<i>À propos de l'Assistant Créer un projet Forecast Studio</i>	227
<i>Créer un projet Forecast Studio : Vérifier les données</i>	228
<i>Créer un projet Forecast Studio : Sélectionner les variables de prévision</i>	228
<i>Créer un projet Forecast Studio : Sélectionner les variables de classification</i>	229
<i>Créer un projet Forecast Studio : Enregistrer le projet</i>	231

À propos de l'Assistant Créer un projet Forecast Studio

L'Assistant **Créer un projet Forecast Studio** aide à spécifier les variables de prévision, à indiquer si vous souhaitez prédire vos données de manière hiérarchique, et à indiquer l'horizon de prévision pour un nouveau projet SAS Forecast Studio.

SAS Forecast Studio est une application de prévision, conçue pour accélérer le processus de prévision grâce à l'automatisation. Le logiciel fournit la sélection automatique des modèles de série chronologique qui sont utilisés dans les données de prévision horodatées. Pour pouvoir enregistrer ce projet, vous devez avoir accès à SAS Forecast Server.

Remarque : pour créer un projet de prévision qui utilise les données SAS de votre PC, vous devez sélectionner la source de données via le système de fichiers plutôt que de sélectionner une source de données sur votre machine locale.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	Procédures de SAS Forecast Server

Nom du prérequis	Noms des procédures et des produits
Produits SAS requis	SAS Forecast Server
Autres produits SAS recommandés	Aucun

Créer un projet Forecast Studio : Vérifier les données

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Cliquez sur **Suivant** pour sélectionner les variables de prévision.

Créer un projet Forecast Studio : Sélectionner les variables de prévision

- 1 Avant de sélectionner les variables de prévision, vous devez vérifier les données.
- 2 Attribuez des variables à ces fonctions :

ID temps

Cette variable indique l'ID temps des données. L'ID temps est une variable de la table d'entrée qui contient la valeur date, date-heure ou heure SAS de chaque observation. Elle permet de déterminer la fréquence et l'ordre des données et d'extrapoler les valeurs ID temps des prévisions. Vous ne pouvez attribuer qu'une variable à cette fonction et il doit s'agir d'une variable date, d'une variable date-heure ou d'une variable numérique contenant des valeurs date ou date-heure.

Prévision (Variables dépendantes)

Les colonnes que vous attribuez à cette fonction spécifient les variables à modéliser et à appliquer pour la prévision. Les Variables de prévision sont également appelées Variables dépendantes. Vous devez attribuer au moins

une variable numérique à cette fonction. Utilisez les flèches vers le haut et vers le bas pour indiquer l'ordre des variables.

- 3 Vous pouvez éventuellement attribuer une ou plusieurs variables numériques à la fonction **Prédire à l'aide de (Variables indépendantes)**. La colonne que vous attribuez à cette fonction spécifie une variable explicative, d'entrée, indépendante, ou de facteur causal. Les variables Prédire à l'aide de sont également appelées Variables indépendantes. Spécifier l'ordre des variables.
- 4 Cliquez sur **Suivant** pour sélectionner les variables de classification.

Créer un projet Forecast Studio : Sélectionner les variables de classification

- 1 Avant de sélectionner les variables de classification, vous devez sélectionner la variable de prévision.
- 2 Attribuez des variables à la fonction **Pour chaque valeur de (Variable de classification)**. Les variables attribuées à cette fonction sont appelées variables de classification. Les observations qui ont la même valeur pour la variable de classification sont groupées. L'attribution d'une variable de classification permet d'obtenir une analyse distincte pour chaque groupe d'observations. Vous pouvez attribuer des variables alphanumériques et numériques à cette fonction.

L'ordre des variables de classification décrit la structure de la hiérarchie. Utilisez les flèches vers le haut et vers le bas pour indiquer l'ordre des variables.

- 3 Indiquez le mode d'affichage de la sortie. Vous avez le choix entre ces options :
 - **Vue sous forme de liste** ne crée pas une hiérarchie des données. La série de prévision est organisée dans une liste.
 - **Hiérarchie** utilise les valeurs des variables de classification pour organiser les données dans une hiérarchie. Cette option est sélectionnée par défaut.

Remarque : si vous attribuez plusieurs variables à la fonction **Prévision (variables dépendantes)** lorsque vous sélectionnez les variables de prévision, l'option **Hiérarchie** n'est pas disponible.

- 4 Si vous choisissez d'afficher la sortie sous forme de hiérarchie, vous pouvez spécifier les paramètres de hiérarchie en cliquant sur **Paramètres**. La fenêtre **Paramètres de la hiérarchie** apparaît.

Vous avez le choix entre ces options :

Méthode d'agrégation

Indiquez la statistique permettant d'agréger les données aux différents niveaux de la hiérarchie. Par défaut, toutes les variables dépendantes et indépendantes sont listées dans la table. Dans la colonne **Agrégation**, sélectionnez la statistique d'agrégation à utiliser pour chaque variable.

Par exemple, votre table contient les ventes pour un groupe de produits. Vous voulez connaître les ventes totales pour une catégorie. Dans ce cas, vous sélectionnez **Somme** comme statistique d'agrégation. Si vos données contiennent le prix de chaque produit, vous voulez éventuellement connaître le prix d'une ligne de produits. Dans ce cas, vous sélectionnez **Moyenne** comme statistique d'agrégation.

Niveau de rapprochement

Sélectionnez le niveau de la hiérarchie qui contient les données que vous voulez utiliser pour la prévision. Vous avez le choix entre ces options :

- **Ascendant** utilise les données au niveau le plus bas de la hiérarchie pour générer les prévisions. Ces prévisions sont ensuite utilisées pour générer les prévisions rapprochées pour les niveaux plus haut dans la hiérarchie.

La méthode Ascendant permet de voir tous les modèles (tels que la saisonnalité) dans les données. Néanmoins, du fait que vous vous trouvez au niveau le plus bas de la hiérarchie, vous pouvez également avoir trop de bruit ou de caractère aléatoire dans les données. De plus, ces prévisions peuvent échouer du fait que les données au plus bas niveau de la hiérarchie peuvent être sporadiques ou trop rares.

- **Transversal** ajoute les données des niveaux inférieurs, puis utilise ces valeurs pour générer les prévisions pour le niveau moyen. Du fait que certaines hiérarchies ont plusieurs niveaux moyens, vous devez spécifier lequel vous voulez utiliser. Les prévisions du niveau moyen s'utilisent pour générer les prévisions pour à la fois les niveaux supérieur et inférieur.
- **Descendant** ajoute les données des niveaux inférieurs dans la prévision, puis utilise cette valeur pour générer les prévisions. La prévision au niveau supérieur s'utilise pour générer les prévisions des niveaux inférieurs de la hiérarchie.

La méthode Descendant permet de supprimer le bruit excessif des données aux niveaux inférieurs de la hiérarchie. Cependant, vous risquez également de perdre le modèle (tel que la saisonnalité) dans les prévisions.

Méthode de désagrégation

Sélectionnez la méthode de désagrégation pour la hiérarchie. Vous avez le choix entre ces options :

- **Proportions de la prévision** constitue la prévision en se basant sur la moyenne historique. Il s'agit de la présentation par défaut.
- **Fractionnement égal de l'écart** constitue la prévision en se basant sur les écarts de prévision entre les niveaux.

- 5 Spécifiez dans le champ **Nombre de périodes à prédire (horizon)** le nombre de périodes dans le futur pour lesquelles les prévisions à étapes multiples seront réalisées. La valeur par défaut est 12.
- 6 Cliquez sur **Suivant** pour enregistrer le projet.

Créer un projet Forecast Studio : Enregistrer le projet

- 1 Avant de pouvoir enregistrer le projet, vous devez sélectionner la variable de classification.
- 2 Saisissez un nom pour le projet dans le champ **Enregistrer sous**. Le nom du projet doit être un nom SAS valide. Il peut contenir 32 caractères, et doit commencer par une lettre (A–Z). Les caractères suivants peuvent être des lettres ou des chiffres numériques (0–9). Les majuscules et les minuscules sont valides. Pour plus d’informations sur les conventions d’appellation SAS, voir [SAS Language Reference: Concepts](#)
- 3 Cliquez sur **Terminer** pour créer le projet SAS Forecast Studio. Une fois que ce projet est créé, vous pouvez l’ouvrir dans SAS Forecast Studio.

Ouvrir un projet Forecast Studio

<i>À propos de l'Assistant Ouvrir un projet Forecast Studio</i>	233
<i>Ouvrir un projet Forecast Studio : Sélectionner le projet de prévision</i>	234
<i>Ouvrir un projet Forecast Studio : Sélectionner la série de prévisions</i>	234
<i>Ouvrir un projet Forecast Studio : Sélectionner les types de résultat</i>	235
<i>Ouvrir un projet Forecast Studio : Indiquer les titres et les notes de bas de page</i>	235

À propos de l'Assistant Ouvrir un projet Forecast Studio

L'Assistant **Ouvrir un projet Forecast Studio** aide à ouvrir la série sélectionnée dans un projet SAS Forecast Studio existant et à indiquer comment afficher les résultats. SAS Forecast Studio est une application de prévision, conçue pour accélérer le processus de prévision grâce à l'automatisation. Le logiciel fournit la sélection automatique des modèles de série chronologique qui sont utilisés dans les données de prévision horodatées.

Avec cette tâche, vous pouvez ouvrir un projet qui a été créé dans SAS Forecast Studio et réaliser d'autres analyses. Pour pouvoir ouvrir un projet de prévision, vous devez avoir accès à SAS Forecast Server.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	Procédures de SAS Forecast Server
Produits SAS requis	SAS Forecast Server
Autres produits SAS recommandés	Aucun

Ouvrir un projet Forecast Studio : Sélectionner le projet de prévision

Dans le volet **Projets de prévision**, vous pouvez voir la liste de tous les projets disponibles ainsi que la date de la dernière modification. Sélectionnez le projet de prévision à ouvrir. Vous ne pouvez ouvrir qu'un projet à la fois.

Par défaut, les détails du projet ne s'affichent pas. Pour voir la description, les variables dépendantes, et les variables de classification (ou BY) qui sont associées au projet sélectionné, cliquez sur **Détails**. Du fait que ces informations sont stockées sur le Forecasting Server, les détails peuvent prendre plusieurs secondes à s'afficher. Pour masquer les détails, cliquez sur **Masquer**.

Cliquez sur **Suivant** pour sélectionner la série de prévision.

Ouvrir un projet Forecast Studio : Sélectionner la série de prévisions

- 1 Avant de sélectionner la série de prévision, vous devez sélectionner le projet de prévision.
- 2 Dans le volet **Sélectionnez la série à afficher**, activez la case à cocher de la série de prévision à ouvrir.

Les séries sont présentées dans une liste si les données du projet n'ont pas été prévues de manière hiérarchique. Les séries sont présentées dans une hiérarchie si les données ont été prévues de manière hiérarchique. Cliquez sur **Développer tout** pour développer tous les noeuds de la hiérarchie. Pour réduire les noeuds, cliquez sur **Réduire tout**.

- 3 Cliquez sur **Suivant** pour sélectionner les types de résultat.

Ouvrir un projet Forecast Studio : Sélectionner les types de résultat

- 1 Avant de sélectionner les types de résultat, vous devez sélectionner la série de prévision.
- 2 Sélectionnez les types de résultat à ouvrir. Vous avez le choix entre ces options :
 - **Courbe des prévisions** affiche les valeurs historiques, les valeurs ajustées, ainsi que la prévision réconciliée pour le modèle sélectionné.
 - **Table des prévisions** affiche le modèle pour la prévision statistique, l'ajustement de la réconciliation, la prévision réconciliée, les valeurs de remplacement manuel, et la prévision finale.
 - **Table de sélection du modèle** affiche tous les modèles qui ont été définis pour le projet en cours.
 - **Table des résultats estimés des paramètres** affiche les estimations, les erreurs types, et les tests de signification pour chacun des paramètres du modèle.
 - **Table des composants** affiche les tables pour chacun des composants utilisés par le modèle.
 - **Table des statistiques d'ajustement** affiche toutes les statistiques d'ajustement disponibles pour le modèle de prévision.
 - **Table de synthèse des prévisions** affiche les valeurs de prévision de la variable dépendante pour la plage d'heure de l'horizon de prévision.
- 3 Cliquez sur **Suivant** pour indiquer des titres et des notes de bas de page.

Ouvrir un projet Forecast Studio : Indiquer les titres et les notes de bas de page

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.
.....

Projet de valeur corrigée Forecast Studio

<i>À propos de l'Assistant Projet de valeur corrigée Forecast Studio</i>	237
<i>Projet de valeur corrigée Forecast Studio : Sélectionner le projet de prévisions</i>	238
<i>Projet de valeur corrigée Forecast Studio : Sélectionner la série de prévisions et soumettre les valeurs corrigées</i>	238
Sélectionner la série de prévisions	239
Créer et soumettre les valeurs corrigées	239
Calculer les valeurs corrigées pour les périodes temporelles sélectionnées	240
Résoudre les conflits de valeurs corrigées	241

À propos de l'Assistant Projet de valeur corrigée Forecast Studio

SAS Forecast Studio est une application de prévision, conçue pour accélérer le processus de prévision grâce à l'automatisation. Le logiciel fournit la sélection automatique des modèles de série chronologique qui sont utilisés dans les données de prévision horodatées. Pour pouvoir ouvrir un projet de prévision dans SAS Enterprise Guide ou dans SAS Add-In to Microsoft Office, vous devez avoir accès à SAS Forecast Server.

L'Assistant **Projet de valeur corrigée Forecast Studio** permet de soumettre les valeurs corrigées pour les données de prévision dans un projet existant de SAS Forecast Studio. Une valeur corrigée est une valeur de remplacement que vous indiquez à la place d'une valeur prévisionnelle. Vous ne pouvez soumettre des valeurs corrigées que pour les valeurs futures.

Une fois que vous avez soumis une valeur corrigée, SAS Forecast Server réconcilie le projet. La façon dont la valeur corrigée est appliquée dépend de si les données sont prédites de manière hiérarchique ou non.

- Si les données du projet sont prédites de manière hiérarchique, la valeur corrigée sera alors un ajustement qui est effectué en tenant compte de la prévision statistique réconciliée.
- Si les données du projet ne sont pas prédites de manière hiérarchique, la valeur corrigée sera alors un ajustement qui est effectué en tenant compte de la prévision statistique.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	Procédure de SAS Forecast Server
Produits SAS requis	SAS Forecast Server
Autres produits SAS recommandés	Aucun

Projet de valeur corrigée Forecast Studio : Sélectionner le projet de prévisions

Dans le volet **Projets de prévision**, vous pouvez voir la liste de tous les projets disponibles ainsi que la date de la dernière modification. Sélectionnez le projet de prévision à ouvrir. Vous ne pouvez ouvrir qu'un projet à la fois.

Par défaut, les détails du projet ne s'affichent pas. Pour voir la description, les variables dépendantes, et les variables de classification (ou BY) qui sont associées au projet sélectionné, cliquez sur **Détails**. Du fait que ces informations sont stockées sur le Forecasting Server, les détails peuvent prendre plusieurs secondes à s'afficher. Pour masquer les détails, cliquez sur **Masquer**.

Cliquez sur **Suivant** pour sélectionner la série de prévision.

Projet de valeur corrigée Forecast Studio : Sélectionner la série de

prévisions et soumettre les valeurs corrigées

Sélectionner la série de prévisions

- 1 Avant de sélectionner la série de prévision, vous devez sélectionner le projet de prévision.
- 2 Dans l'arborescence, sélectionnez les séries pour lesquelles vous voulez soumettre une valeur corrigée.

Si les données du projet ont été prédites de manière hiérarchique, les séries sont présentées dans une hiérarchie. Cliquez sur **Développer tout** pour développer tous les noeuds de la hiérarchie. Pour réduire les noeuds, cliquez sur **Réduire tout**.

Les séries sont présentées dans une liste si les données du projet n'ont pas été prévues de manière hiérarchique.

- 3 Une fois que vous avez sélectionné les séries, cliquez sur **Corriger les données de prévision** pour créer et soumettre les valeurs corrigées.
- 4 Répétez les étapes 2 et 3 pour toutes les séries pour lesquelles vous voulez soumettre une valeur corrigée.
- 5 Une fois que vous avez créé et soumis vos valeurs corrigées, cliquez sur **Terminer** pour fermer cet assistant.

Créer et soumettre les valeurs corrigées

Deux graphiques se trouvent en haut de la boîte de dialogue **Corriger les données de prévision**. Dans le graphique de gauche, le trait bleu montre la prévision et les cercles représentent les données historiques. Dans le graphique de droite, le trait bleu foncé montre la prévision, les traits en bleu plus clair correspondent aux intervalles de confiance de la prévision ; et les signes plus en noir représentent les valeurs corrigées spécifiées.

Pour créer et soumettre les valeurs corrigées :

- 1 Dans le tableau **Modifier les valeurs corrigées**, sélectionnez la période où vous voulez ajouter ou modifier une valeur corrigée. Vous pouvez sélectionner plusieurs périodes en appuyant sur Ctrl et en sélectionnant les lignes dans le tableau.

- 2 Pour créer une valeur corrigée pour une période individuelle, cliquez sur la cellule dans la colonne **Valeurs corrigées** pour cette période. Puis saisissez la valeur corrigée et appuyez sur Entrée.

Pour calculer les valeurs corrigées pour les périodes sélectionnées, cliquez sur **Calculer valeurs corrigées**. La boîte de dialogue **Calculateur de valeurs corrigées** apparaît.

Pour supprimer les valeurs corrigées que vous avez spécifiées durant cette session, cliquez sur **Réinitialiser**. Les valeurs corrigées indiquées auparavant ou spécifiées dans SAS Forecast Studio ne sont pas supprimées.

- 3 Une fois que vous avez créé des valeurs corrigées pour les séries en cours, cliquez sur **Soumettre**. SAS Forecast Server réconcilie le projet. Une fois que le projet est réconcilié, la fenêtre **Corriger les données de prévision** se ferme, et vous retournez à l'étape Sélectionner les séries de prévision de l'assistant. Dans l'arborescence, l'icône Valeur corrigée apparaît face aux séries pour lesquelles vous avez soumis les valeurs corrigées.

Remarque : Si SAS Forecast Server détecte un conflit de valeurs corrigées lors de la réconciliation du projet, un message d'erreur apparaît, indiquant que le projet ne peut pas être réconcilié et que vous devez utiliser SAS Forecast Studio pour résoudre les conflits.

Calculer les valeurs corrigées pour les périodes temporelles sélectionnées

Vous pouvez calculer les valeurs corrigées pour une ou plusieurs périodes temporelles en utilisant ces options :

Nom d'option	Description
Répartir proportionnellement à la prévision actuelle pour chaque période	Permet de baser les valeurs corrigées sur l'augmentation ou la diminution spécifiée de la prévision statistique réconciliée. Remarque : si les prévisions de vos données ne sont pas hiérarchiques, ou bien s'il n'y a aucune prévision de réconciliation disponible, vous pouvez sélectionner cette option pour baser les valeurs corrigées sur l'augmentation ou la diminution spécifiée de la prévision statistique. À l'aide des listes déroulantes et du champ de saisie, spécifiez le nombre d'unités ou le pourcentage pour augmenter ou réduire la valeur en cours. Sélectionnez l'option Augmenter ou Diminuer. Dans le champ de saisie de l'option sélectionnée, spécifiez la valeur et, dans la liste déroulante, indiquez si cette valeur est une unité ou un pourcentage.

Nom d'option	Description
	Par exemple, vous voulez créer une valeur corrigée qui est 10% supérieure à la prévision statistique réconciliée. Sélectionnez l'option Augmenter. Dans le champ de saisie, tapez 10 et dans la liste déroulante, sélectionnez % (pourcentage). Les valeurs corrigées sont 10% supérieures à la prévision statistique réconciliée.
Répartir une valeur entre les périodes sélectionnées	Répartie une seule valeur corrigée sur toutes les périodes sélectionnées. Spécifiez cette valeur dans le champ de saisie. Le mode d'application de la valeur corrigée dépend de l'option que vous sélectionnez. Vous avez le choix entre ces options : ■ Répartir proportionnellement à la prévision actuelle pour chaque période - la valeur indiquée est répartie proportionnellement entre les périodes sélectionnées. Par exemple, vous sélectionnez deux périodes avec des prévisions de 550 et 430. Vous spécifiez une valeur de 100, et vous choisissez de diviser cette valeur proportionnellement entre les deux périodes. À présent, les prévisions sont 606,12 et 473,88 ■ Répartir équitablement dans les périodes la valeur indiquée est répartie équitablement parmi les périodes sélectionnées. Par exemple, vous sélectionnez deux périodes avec des prévisions de 550 et 430. Vous spécifiez une valeur de 100, et vous choisissez de diviser cette valeur de façon égale entre les deux périodes. À présent, les prévisions sont 600 et 480.
Affecter une valeur à chaque période sélectionnée	Applique une seule valeur corrigée sur toutes les périodes sélectionnées. Spécifiez cette valeur dans le champ de saisie.

Une fois que vous avez calculé les valeurs corrigées pour les périodes sélectionnées, cliquez sur **OK** pour soumettre les valeurs corrigées.

Résoudre les conflits de valeurs corrigées

Si SAS Forecast Server détecte un conflit de valeurs corrigées lors de la réconciliation du projet, un message d'erreur apparaît, indiquant que le projet ne peut pas être réconcilié et que vous devez utiliser SAS Forecast Studio pour résoudre les conflits. Dans SAS Forecast Studio, vous pouvez voir toutes les valeurs corrigées ayant été soumises, et vous pouvez choisir quelles valeurs corrigées appliquer aux données de prévision.

Pour résoudre les conflits de valeurs corrigées :

- 1 Cliquez sur **Annuler** pour quitter l'Assistant **Soumettre les valeurs corrigées de Forecast Studio** et fermez Microsoft Excel.
- 2 Ouvrez le projet dans SAS Forecast Studio. Lorsque vous ouvrez le projet dans SAS Forecast Studio, l'avertissement "Reconciliation is out of date" apparaît en haut de l'espace de travail.
- 3 Face à cet avertissement, cliquez sur **Mise à jour**. La boîte de dialogue **Conflit de valeurs corrigées** apparaît.
- 4 Pour résoudre les conflits de valeurs corrigées, vous pouvez modifier les valeurs corrigées en utilisant la boîte de dialogue **Conflits de valeurs corrigées**, ou bien vous pouvez fermer cette boîte de dialogue et modifier les valeurs corrigées en utilisant la table de données.

Pour plus d'informations sur la résolution des conflits de valeurs corrigées, voir l'aide de SAS Forecast Studio.

Modèle linéaire généralisé

<i>À propos de la tâche Modèles linéaires généralisés</i>	243
<i>Modèles linéaires généralisés : Attribuer des variables aux fonctions d'analyse</i>	244
<i>Modèles linéaires généralisés : Créer un modèle</i>	245
<i>Modèles linéaires généralisés : Définir les options du modèle</i>	246
<i>Modèles linéaires généralisés : Sélectionner les tests post hoc</i>	248
<i>Modèles linéaires généralisés : Générer des graphes</i>	248
<i>Modèles linéaires généralisés : Définir les options de prédiction</i>	249
<i>Indiquer des titres et des notes de bas de page</i>	250
<i>Afficher les propriétés</i>	251

À propos de la tâche Modèles linéaires généralisés

La tâche Modèles linéaires généralisés est une extension du modèle linéaire traditionnel qui vous permet de modéliser les données qui ne sont pas normalement distribuées, telles que des effectifs ou des proportions mesurées qui sont considérés comme discrets. Vous pouvez également utiliser cette tâche pour modéliser les données pour lesquelles la moyenne a été limitée à une plage de valeurs, ou les données pour lesquelles la variance n'est pas constante pour toutes les observations.

Vous pouvez utiliser cette tâche pour :

- comparer la durée de vie de produits de différents fabricants. Par exemple, supposons que vous souhaitiez analyser les temps de défaillance de pièces de machine, dont certaines sont fabriquées par le fabricant A et d'autres par le fabricant B. Vous pouvez utiliser cette tâche pour tester les différences entre les durées de vie des pièces des deux fabricants.

- ajuster un modèle à des données mesurées sur une échelle ordinale. Par exemple, supposons que vous disposiez d'une table contenant les résultats des tests gustatifs de trois marques de crème glacée. Les trois marques sont évaluées en termes de goût sur une échelle en cinq points, allant de très bon (tb) à très mauvais (tm). L'analyse consiste à déterminer les différences d'évaluation des trois marques.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GENMOD
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Modèles linéaires généralisés : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables dépendantes	Indique les variables qui doivent être utilisées comme variables dépendantes (de réponse). Vous devez attribuer au moins une variable à cette fonction.
Variables quantitatives	spécifie les variables à utiliser comme variables indépendantes (explicatives) continues (également appelées effets).
Variables de classification	indique les variables à utiliser en tant qu'effets indépendants discrets. Les variables que vous attribuez à cette fonction peuvent être numériques ou alphanumériques, mais doivent avoir un nombre limité de valeurs discrètes.

Nom du rôle	Description
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper en fonction d'une variable que vous avez déjà sélectionnée comme variable dépendante ou variable indépendante.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Poids relatif	spécifie une variable à utiliser comme pondération relative pour les valeurs correspondantes de la variable sélectionnée comme variable de réponse. Vous ne pouvez attribuer qu'une seule variable à cette fonction.
Décalage relatif	spécifie une variable à utiliser en tant que variable de décalage. Remarque : vous ne pouvez attribuer qu'une seule variable à cette fonction. Vous ne pouvez pas faire un décalage en fonction d'une variable que vous avez déjà sélectionnée comme variable dépendante, comme variable indépendante ou comme variable de classification.

Modèles linéaires généralisés : Créer un modèle

Vous pouvez créer des modèles dans plusieurs tâches SAS. Pour plus d'informations, voir [Annexe 3, "Créer un modèle" page 749](#).

Modèles linéaires généralisés : Définir les options du modèle

Dans la liste, cliquez sur **Options du modèle** pour accéder à ces options.

Nom d'option	Description
Distribution	affiche la probabilité d'une variable de réponse dans votre modèle. Remarque : les distributions binomiales ne sont disponibles que si la variable dépendante comporte deux niveaux. Remarque : si une variable alphanumérique a été attribuée à la fonction Variables dépendantes, seules les distributions multinomiales et binomiales sont disponibles.
Fonctions de liaison	décrit la manière dont la valeur attendue est liée à l'estimateur linéaire. Plusieurs fonctions de liaison sont disponibles, selon le modèle que vous créez. Dans la liste déroulante Fonction de liaison, vous avez le choix entre les fonctions de liaison suivantes : ■ Aucun ■ Identité ■ Logit ■ Probit ■ Log ■ Log-log complémentaire ■ Puissance avec nombre Remarque : les fonctions de liaison Logit cumulé, Log-log complémentaire cumulé et Probit cumulé ne sont disponibles que lorsque vous sélectionnez une distribution multinomiale. Si vous ne sélectionnez aucune fonction de liaison dans la liste déroulante Fonction de liaison alors que vous avez spécifié une fonction de liaison personnalisée, cette dernière est utilisée. Si vous n'avez pas défini de fonction de liaison personnalisée, la fonction de liaison par défaut de la distribution du modèle est utilisée. ■ Pour une distribution binomiale, la fonction de liaison par défaut est logit. ■ Pour une distribution gamma, la fonction de liaison par défaut est inverse (puissance (-1)).

Nom d'option	Description
	■ Pour une distribution gaussienne inverse, la fonction de liaison par défaut est le carré inverse (puissance (-2)). ■ Pour une distribution multinomiale, la fonction de liaison par défaut est logit cumulé. ■ Pour une distribution binomiale négative, la fonction de liaison par défaut est log. ■ Pour une distribution normale, la fonction de liaison par défaut est identité. ■ Pour une distribution de Poisson, la fonction de liaison par défaut est log. Remarque : si vous sélectionnez Néant pour la distribution et la fonction de liaison, la distribution normale avec la fonction de liaison d'identité est utilisée dans l'analyse.
Tests des effets	spécifie les statistiques à inclure dans l'analyse. ■ Activez la case à cocher Calculer une analyse de Type 1 ou séquentielle pour afficher la statistique du test du rapport de vraisemblance entre des paires de modèles successives. Les modèles ajustés de manière séquentielle commencent par le modèle nul (terme d'intercept uniquement) et vont jusqu'aux modèles que vous avez définis. ■ Activez la case à cocher Calculer les statistiques pour les contrastes de Type 3 afin d'afficher les statistiques de contraste pour chaque effet dans les modèles que vous avez définis. L'analyse par défaut est de calculer les statistiques du rapport de vraisemblance pour les contrastes. Les statistiques de Wald sont également calculées si la case à cocher Calculer statistiques de Wald est activée.
Intervalle de confiance	spécifie le niveau de confiance pour les valeurs prédites. La valeur par défaut est 95 pour cent.
Détails des estimations	indique s'il faut afficher la matrice de corrélation d'estimation de paramètre, la matrice de covariance d'estimation de paramètre, ou les deux.
Intervalle de confiance	indique si un intervalle de confiance bilatéral doit être utilisé. ■ Activez la case à cocher Bilatéral pour calculer les intervalles de confiance bilatéraux pour tous les paramètres du modèle basés sur la fonction de vraisemblance du profil (également connue sous le nom de fonction de vraisemblance partiellement maximisée). ■ Activez la case à cocher Wald bilatéral pour calculer les intervalles de confiance bilatéraux de Wald pour tous les

Nom d'option	Description
	paramètres du modèle basés sur la normalité asymptotique des résultats estimés des paramètres.

Modèles linéaires généralisés : Sélectionner les tests post hoc

Dans la liste, cliquez sur **Tests post hoc** pour accéder à ces options.

Pour chaque effet de la zone **Effets à estimer**, spécifiez des options dans la zone **Options pour les tests de la moyenne**.

Modèles linéaires généralisés : Générer des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Par défaut, la sortie comprend tous les graphes appropriés pour la sélection de données actuelle. Vous pouvez néanmoins choisir les graphes à inclure dans la sortie en sélectionnant l'option **Liste personnalisée de graphes**. Vous avez le choix entre les options suivantes :

Nom d'option	Description
Graphe D de Cook	crée un graphique de la statistique D de Cook par rapport au nombre d'observations.
Graphes DFBETA	crée des graphiques de la différence dans l'estimation des paramètres en raison de la suppression de l'observation.
Graphes DFBETAS	crée des graphiques de la différence normalisée dans l'estimation des paramètres liée à la suppression de l'observation. Est utile pour évaluer l'effet d'une observation individuelle sur chaque paramètre estimé du modèle ajusté.

Nom d'option	Description
Graphe d'effet de levier	crée un graphe de la valeur de l'élément diagonal correspondant de la matrice H. Les éléments diagonaux de la matrice sont utiles pour détecter les points extrêmes dans l'espace où ils ont tendance à avoir des valeurs élevées.
Graphe des valeurs prédites	crée un graphique des valeurs prédites.
Graphique des résidus de Pearson	crée un graphique du résidu de Pearson (Khi) utilisé pour identifier les observations peu représentées pour le modèle par rapport aux valeurs prédites.
Graphique du résidu de l'écart	crée un graphique du résidu de l'écart pour identifier les observations dont l'ajustement est insatisfaisant par rapport aux valeurs prédites.
Graphe du résidu de vraisemblance	crée un graphique du résidu de vraisemblance utilisé pour identifier les valeurs dont l'ajustement est insatisfaisant par rapport aux valeurs prédites.
Graphique des résidus bruts	crée un graphique des résidus bruts par rapport aux valeurs prédites.
Résidu de Pearson standardisé	crée un graphique du résidu de Pearson (Khi) normalisé permettant d'identifier les observations peu représentées par le modèle par rapport aux valeurs prédites.
Résidu de l'écart normalisé	crée un graphique du résidu de l'écart normalisé utilisé pour identifier les valeurs dont l'ajustement est insatisfaisant par rapport aux valeurs prédites.

Modèles linéaires généralisés : Définir les options de prédiction

Dans la liste, cliquez sur **Prédictions** pour accéder à ces options.

Nom d'option	Description
Données à prédire	identifie les sources de données à utiliser pour prédire les données de sortie. Activez la case à cocher Données supplémentaires pour inclure ces données. Pour désigner l'emplacement des données supplémentaires, cliquez sur Parcourir.
Enregistrer les données de sortie	spécifie s'il faut enregistrer les prévisions et les statistiques de diagnostic dans une table de sortie. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Statistiques supplémentaires	spécifie s'il faut enregistrer des statistiques supplémentaires. ■ Résidus enregistre le résidu de probabilité utilisé pour identifier les observations peu ajustées, le résidu de Pearson (Chi) normalisé utilisé pour identifier les observations peu représentées pour le modèle, et le résidu d'écart normalisé utilisé pour identifier les observations dont l'ajustement est insatisfaisant. ■ Limites de prédiction enregistre l'intervalle de confiance supérieure et inférieure pour la valeur prédite.
Afficher les prédictions	affiche les prédictions dans les résultats.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

.....

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

.....

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

.....

Régression linéaire High-Performance

<i>À propos de la tâche Régression linéaire High-Performance</i>	253
<i>Régression linéaire High-Performance : Attribuer des variables aux fonctions d'analyse</i>	254
<i>Régression linéaire High-Performance : Créer un modèle</i>	256
Créer un modèle	256
Spécifier une méthode de sélection du modèle	256
<i>Régression linéaire High-Performance : Définir les options du modèle</i>	259
<i>Régression linéaire High-Performance : Définir les options de la partition</i>	260
<i>Régression High-Performance : Définir les options de performances</i>	261
<i>Régression High-Performance : Définir les options de prédiction</i>	261
<i>Indiquer des titres et des notes de bas de page</i>	262

À propos de la tâche Régression linéaire High-Performance

La tâche Régression linéaire High-Performance ajuste et réalise une sélection de modèles pour les modèles de moindres carrés ordinaires. Contrairement avec la tâche Régression linéaire, vous pouvez spécifier des modèles linéaires généraux qui comprennent des variables de classification. La tâche Régression linéaire High-Performance prend également en charge les méthodes LAR (Least Angle Regression) et de sélection Lasso.

La tâche Régression linéaire High-Performance autorise deux paramétrages (GLM et référence) pour les variables de classification. Contrairement à d'autres tâches qui autorisent plusieurs paramétrages, les variables de classification de cette tâche ont le même paramétrage. Le paramétrage GLM est celui par défaut.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	HPREG
Produits SAS requis	Base SAS, SAS High-Performance Analytics Server, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Régression linéaire High-Performance : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables dépendantes	Indique les variables qui doivent être utilisées comme variables dépendantes (de réponse). Vous devez attribuer au moins une variable à cette fonction.
Effets explicatifs	indique les variables à utiliser en tant que variables de régression (explicatives). Si vous ne spécifiez aucune variable de régression, chaque variable dépendante est ajustée à un modèle à moyenne uniquement. c'est-à-dire un modèle comportant seulement un terme constant.
Variables de classification	indique les variables de classification à utiliser dans l'analyse. Ce peut être des variables caractères ou numériques. Si elles sont numériques, les valeurs doivent être discrètes plutôt que continues. Ces variables saisissent l'analyse au travers de niveaux sur lesquels des valeurs uniques sont mappées. Lorsque les valeurs sont mises à niveau, les observations qui partagent la même valeur sont affectées au même niveau. Vous pouvez spécifier le style de codification pour toutes les variables de classification ou pour des variables de classification individuelles.

Nom du rôle	Description
	Pour spécifier le même style de codification pour toutes les variables de classification, utilisez les options dans la zone Options générales de la variable de classification. ■ Utilisez l'option GLM pour le codage d'une cellule de référence less-than-full-rank. Ce paramétrage s'utilise dans des procédures telles que GLM, MIXED, et GLIMMIX dans SAS/STAT. ■ Utilisez l'option Référence pour le codage de la cellule de référence. Indiquez si vous voulez utiliser la première ou la dernière valeur classée comme référence. Si vous sélectionnez Référence, vous pouvez remplacer le codage de cellule de référence pour une variable de classification individuelle. La case à cocher Fractionner spécifie quelles colonnes de la matrice d'expérience correspondent à un effet qui contient une variable de classification fractionnée. Chaque colonne peut être sélectionnée pour entrer dans le modèle ou le quitter indépendamment des autres colonnes de cet effet. Vous pouvez spécifier ceci en tant qu'option globale ou en tant qu'option sur une variable de classification individuelle. Supposons que la variable Temp ait trois niveaux avec les valeurs hot, warm, et cold et que la variable gender ait deux niveaux avec les valeurs M et F. ■ Si Fractionner est sélectionné comme option générale, les deux effets sont divisés en huit effets indépendants. L'effet gender est divisé en deux effets : gender_M et gender_F. L'effet gender*temp est divisé en six effets : gender_M*temp_hot, gender_F*temp_hot, gender_M*temp_warm, gender_F*temp_warm, gender_M*temp_cold, et gender_F*temp_cold ■ Si Fractionner est sélectionné comme option pour la variable temp, l'effet gender*temp est divisé en trois effets : gender*temp_hot, gender*temp_warm, et gender*temp_cold. Chacun de ces trois effets fractionnés a maintenant deux paramètres qui correspondent aux deux niveaux de gender.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Poids relatif	indique la variable contenant les valeurs qui sont des pondérations relatives pour un ajustement pondéré des moindres carrés. Les valeurs de la variable de poids relatif ne doivent pas être négatives. Vous ne pouvez attribuer qu'une seule variable à cette fonction.

Régression linéaire High-Performance : Créer un modèle

Créer un modèle

Vous pouvez créer des modèles dans plusieurs tâches SAS. Pour plus d'informations, voir [Annexe 3, "Créer un modèle" page 749](#).

Spécifier une méthode de sélection du modèle

Dans la liste, sous l'en-tête Modèle, cliquez sur **Sélection** pour accéder à ces options.

Nom d'option	Description
Méthode de sélection de modèle	Sélectionnez une méthode dans la liste déroulante des méthodes de sélection des modèles. Si vous ne choisissez pas de méthode, le modèle entier est utilisé. En fonction de la méthode de sélection du modèle, des options supplémentaires deviennent disponibles. Voici les méthodes disponibles : ■ Modèle complet ajusté (aucune sélection) est la valeur par défaut et ne fournit aucune sélection d'effet. Le modèle est celui créé lors de l'attribution des variables dépendantes et quantitatives. ■ La sélection ascendante commence sans effets dans le modèle. Pour chacune des variables explicatives, cette méthode calcule les statistiques F reflétant la contribution de la variable dans le modèle. Les p-values de ces statistiques F sont comparées au niveau de significativité spécifié pour inclure une variable dans le modèle. La valeur par défaut est 0,05. Pour changer le niveau de significativité, saisissez une valeur dans le champ Pour entrer dans le modèle. Si aucune statistique F n'a un niveau de significativité supérieur à cette valeur, la sélection ascendante s'arrête. Sinon, la méthode de sélection ascendante ajoute la variable dont la statistique F est la plus élevée dans le modèle. La méthode de sélection ascendante recalcule ensuite les statistiques F pour les variables

Nom d'option	Description
	restant en dehors du modèle, et le processus d'évaluation est répété. Ainsi, les variables sont ajoutées une à une au modèle jusqu'à ce qu'aucune variable restante ne produise une statistique F significative. Une fois qu'une variable a été ajoutée au modèle, elle y reste. ■ La sélection descendante commence par calculer les statistiques F pour un modèle, en incluant toutes les variables explicatives. Les variables sont ensuite supprimées du modèle une à une jusqu'à ce que toutes les variables restant dans le modèle produisent des statistiques F significatives. Le niveau de significativité est spécifié dans la zone de texte Pour rester dans le modèle. La valeur par défaut est 0,05. À chaque étape, la variable montrant la plus faible contribution au modèle est supprimée. ■ La sélection pas à pas est une variante de la méthode ascendante. Dans la méthode pas à pas, les variables qui sont déjà dans le modèle n'y restent pas nécessairement. De même que dans la méthode de sélection ascendante, les variables sont ajoutées une à une au modèle si la statistique F est significative. Le niveau de significativité est spécifié dans la zone de texte Pour rester dans le modèle. Cependant, après l'ajout d'une variable, la méthode pas à pas vérifie toutes les variables qui sont déjà incluses dans le modèle et supprime celles qui ne produisent pas une statistique F significative. Le niveau de significativité est spécifié dans la zone de texte Pour rester dans le modèle. Ce n'est qu'après cette vérification et après les suppressions requises qu'une autre variable peut être ajoutée au modèle. Le processus pas à pas prend fin lorsque l'une de ces conditions est remplie : ☐ aucune variable en dehors du modèle n'a une statistique F significative, et chaque variable du modèle est significative au niveau de significativité qui est spécifié pour rester dans le modèle. ☐ la variable à ajouter dans le modèle est la variable qui vient d'en être supprimée. ■ La méthode de Sélection Forward-swap est une extension de la méthode de sélection ascendante. Cette méthode utilise les permutations d'effets par paires à l'intérieur et à l'extérieur du modèle pour améliorer le critère de sélection. ■ La méthode LAR (Least Angle Regression), à l'instar de la sélection ascendante, démarre avec zéro effet dans le modèle et ajoute des modèles. Les résultats estimés des paramètres de n'importe quelle étape sont plus petits en comparaison avec les estimations des moindres carrés correspondants. Si le modèle contient des variables de classification, ces dernières sont fractionnées.

Nom d'option	Description
	■ La méthode Lasso ajoute et supprime les paramètres en se basant sur une version des moindres carrés ordinaires où la somme des coefficients de régression absolus est restreinte. Si le modèle contient des variables de classification, ces dernières sont fractionnées.
Niveaux de significativité	Pour les méthodes de sélection ascendante et LAR (Least Angle Regression), spécifiez un niveau de significativité pour saisir une variable quantitative (explicative) dans le modèle. Pour la méthode de sélection descendante, spécifiez un niveau de significativité de sorte qu'une variable quantitative reste dans le modèle. Pour les méthodes de sélection pas à pas et Lasso, spécifiez les deux niveaux de significativité.
Sélectionner le modèle avec la meilleure valeur	Utilise le critère à sélectionner (à chaque étape du processus de sélection) le modèle qui produit la meilleure valeur. Si la valeur optimale du critère spécifié survient pour les modèles à plus d'une étape, le modèle avec le plus petit nombre de paramètres est celui qui est choisi. Si vous ne spécifiez pas de critère, le modèle sélectionné est le modèle à l'étape finale du processus de sélection. Voici les critères disponibles : ■ AIC - Critère d'information d'Akaike ■ AICC - un "small-sample bias" qui est la version corrigée du critère d'information d'Akaike ■ BIC - Critère de Schwarz ■ ADJRSQ - critère R-carré ajusté ■ CP - Statistique C_p de Mallows ■ PRESS - somme résiduelle prédictive des carrés (Predictive Residual Sum of Squares) ■ VALIDATE - erreur quadratique moyenne sur les données de validation
Sélectionner les critères pour l'arrêt du processus de sélection	Utilise le critère pour stopper le processus de sélection. Si vous sélectionnez Néant, le processus de sélection s'arrête si aucun candidat d'ajout ou de suppression approprié ne peut être trouvé ou si une limite basée sur la taille est atteinte. Voici les critères disponibles : ■ SL - la significativité du test de score ■ AIC - Critère d'information d'Akaike ■ AICC - un "small-sample bias" qui est la version corrigée du critère d'information d'Akaike ■ BIC - Critère de Schwarz ■ ADJRSQ - critère R-carré ajusté ■ CP - Statistique C_p de Mallows

Nom d'option	Description
	■ PRESS - somme résiduelle prédictive des carrés (Predictive Residual Sum of Squares) ■ VALIDATE - erreur quadratique moyenne sur les données de validation
Effets à forcer dans le modèle	Sélectionnez l'effet que vous voulez ajouter au modèle. La variable est ajoutée dans la zone Effets à forcer dans le modèle. Vous pouvez réorganiser ces variables en les sélectionnant et en utilisant les boutons fléchés. Remarque : cette option n'est pas disponible si vous sélectionnez les méthodes de sélection des modèles Modèle complet ajusté (aucune sélection), LAR (Least angle regression), ou Lasso.

Régression linéaire High-Performance : Définir les options du modèle

Dans la liste, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Détails sur les estimations	Vous pouvez sélectionner ces statistiques dans la zone Détails sur les estimations : ■ Coefficients de régression standardisés produit des coefficients de régression standardisés. Un coefficient de régression standardisé est calculé en divisant l'estimation du paramètre par le rapport entre l'écart-type de l'échantillon de la variable dépendante et l'écart-type du régresseur. ■ IC pour les résultats estimés des paramètres demande les bornes supérieure et inférieure de l'intervalle de confiance pour les résultats estimés des paramètres. Par défaut, les bornes à 95% sont calculées.
Diagnostics de colinéarité	Vous pouvez sélectionner ces diagnostics : ■ Valeurs de tolérance pour estimation produit les valeurs de tolérance pour les estimations. La tolérance d'un paramètre est définie par $1 - R^2$, où R^2 est obtenu à partir de la régression du paramètre sur tous les autres paramètres du modèle. ■ Valeurs d'inflation de la variance produit des facteurs d'inflation de la variance avec les résultats estimés des

Nom d'option	Description
	paramètres. L'inflation de variance est la réciproque de la tolérance.
	Remarque : ces options ne sont pas disponibles si vous sélectionnez les méthodes de sélection de modèles LAR (Least angle regression) ou Lasso.

Régression linéaire High-Performance : Définir les options de la partition

Dans la liste, cliquez sur **Partition** pour accéder à ces options.

Nom d'option	Description
Ne pas partitionner les données pour la formation, la validation et le test du modèle	spécifie que les observations ne sont pas partitionnées.
Spécifier des proportions à utiliser pour les affectations aléatoires des observations	spécifie comment les observations dans la table d'entrée sont partitionnées dans les sous-ensembles disjoints pour l'apprentissage du modèle, la validation, et les tests. Vous pouvez spécifier que les proportions des observations peuvent être attribuées aléatoirement aux fonctions d'apprentissage et de validation. ■ Pourcentage pour l'apprentissage affiche la proportion des observations dans la table d'entrée qui sont attribuées de façon aléatoire à la fonction d'apprentissage. Les observations qui ne sont pas attribuées aux fonctions de validation ou de test sont attribuées à la fonction d'apprentissage. ■ Pourcentage de validation spécifie la proportion des observations dans la table d'entrée qui sont attribuées de façon aléatoire à la fonction de validation. ■ Pourcentage de test spécifie la proportion des observations dans la table d'entrée qui sont aléatoirement attribuées à la fonction de test.

Régression High-Performance : Définir les options de performances

Dans la liste, cliquez sur **Performances** pour accéder à ces options.

Nom d'option	Description
Utiliser les paramètres par défaut pour cette tâche	utilise les paramètres par défaut.
Utiliser les paramètres personnalisés pour cette tâche	permet de spécifier le noeud, l'hôte du Grid, l'emplacement de l'installation, et le Data Server pour la tâche. ■ Noeuds spécifie le nombre de noeuds dans l'environnement informatique distribué, à condition que les données ne soient pas traitées avec la base de données. ■ Hôte du Grid spécifie le nom de l'hôte de l'équipement. ■ Emplacement de l'installation spécifie le répertoire dans lequel les bibliothèques partagées de High-Performance Analytics sont installées sur le matériel. ■ Data Server spécifie le nom du serveur sur les systèmes Teradata tels que définis via le fichier d'hôtes et tels qu'utilisés dans l'instruction LIBNAME pour Teradata.
Afficher les étapes détaillées de la procédure	crée une table qui affiche une répartition des étapes de la procédure.

Régression High-Performance : Définir les options de prédiction

Dans la liste, cliquez sur **Prédictions** pour accéder à ces options.

Nom du graphe	Description
Enregistrer les données de sortie	indique si les prédictions et statistiques de diagnostic doivent être enregistrées de façon permanente dans une table de sortie. ■ Prédictions enregistre la probabilité prédite pour chaque niveau de réponse. Pour une variable de réponse Y à trois niveaux, 1, 2 et 3, les probabilités individuelles sont $\Pr(Y=1)$, $\Pr(Y=2)$, et $\Pr(Y=3)$. ■ Statistiques de diagnostic enregistre les statistiques de diagnostic DFBETAS, les diagnostics de décalage de l'intervalle de confiance (C et CBAR), la matrice H, et les estimations du prédicteur linéaire avec les estimations de l'erreur type. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Statistiques supplémentaires	■ Sélectionnez Résidus pour enregistrer ces statistiques : ☐ la modification de l'écart ☐ le résidu d'écart ☐ le résidu de Pearson (Khi-2) ☐ modification de l'écart imputable à la suppression d'une observation donnée ; ☐ modification de la statistique critère qualificatif de lissage du Khi-2 imputable à la suppression d'une observation donnée ; ■ Activez Limites de prédiction pour inclure bornes supérieure et inférieure d'un intervalle de confiance pour chaque prédiction et la moyenne attendue de la variable dépendante dans la sortie.
Sélectionner des variables de sortie	spécifie les colonnes à inclure dans la table de sortie.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Régression logistique High-Performance

<i>À propos de la tâche Régression logistique High-Performance</i>	265
<i>Régression logistique High-Performance : Attribuer des variables aux fonctions d'analyse</i>	266
<i>Régression logistique High-Performance : Créer un modèle</i>	267
Spécifier la réponse	267
Créer un modèle	269
Spécifier une méthode de sélection du modèle	269
<i>Régression logistique High-Performance : Définir les options du modèle</i>	271
<i>Régression High-Performance : Définir les options de performances</i>	273
<i>Régression High-Performance : Définir les options de prédiction</i>	273
<i>Indiquer des titres et des notes de bas de page</i>	274

À propos de la tâche Régression logistique High-Performance

La tâche Régression logistique High-Performance ajuste les modèles de régression logistique pour les données binaires, binomiales et multinomiales sur équipement SAS. La tâche Régression logistique High-Performance permet deux paramétrages (GLM et référence) pour les variables de classification. Contrairement à d'autres tâches qui autorisent plusieurs paramétrages, les variables de classification de cette tâche ont le même paramétrage. Le paramétrage GLM est celui par défaut.

La tâche Régression logistique High-Performance est spécifiquement conçue pour opérer dans un environnement distribué à hautes performances. Par défaut, les calculs sont réalisés dans plusieurs threads. La tâche Régression logistique utilise un thread unique.

Remarque : lorsque vous ouvrez la tâche Régression logistique High-Performance, vous remarquerez peut-être un léger décalage du fait que SAS Enterprise Guide détermine quelles variables ont des valeurs binaires. Une fois qu'elle est terminée, la tâche fournit des informations sur la variable attribuée à la fonction **Variable dépendante**.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	HPLOGISTIC
Produits SAS requis	Base SAS, SAS High-Performance Analytics Server, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Régression logistique High-Performance : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variable dépendante	indique la variable à utiliser en tant que variable dépendante (de réponse). Vous ne devez attribuer qu'une seule variable à cette fonction.
Variables quantitatives	indique quelles variables utiliser comme variables indépendantes (explicatives).
Variables de classification	indique les variables de classification à utiliser dans l'analyse. Ce peut être des variables caractères ou numériques. Si elles sont numériques, les valeurs doivent être discrètes plutôt que continues.

Nom du rôle	Description
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Poids relatif	spécifie une variable à utiliser comme pondération relative pour les valeurs correspondantes de la variable sélectionnée comme variable de réponse. Si une valeur de variable est négative ou manquante, elle est exclue de l'analyse.

Régression logistique High-Performance : Créer un modèle

Spécifier la réponse

Dans la liste, sous l'en-tête Modèle, cliquez sur **Réponse** pour accéder à ces options.

Nom du rôle	Description
Type de réponse	Les types de réponse disponibles varient selon le nombre de niveaux de la variable de réponse. ■ Si la variable de réponse ne comporte que deux niveaux, le type de réponse est Binaire. ■ Si la variable de réponse comporte plus de deux niveaux, vous avez le choix entre les types de réponse Classée et Non classée.
Type de modèle	Vous pouvez choisir parmi ces types de modèles : ■ logit affiche la sortie de la fonction log-odds qui est la sélection par défaut. Le modèle logit binaire est utilisé lorsqu'il existe deux catégories de réponses, tandis que le modèle logit cumulé est utilisé lorsqu'il en existe plus de deux. ■ probit - affiche la sortie de l'inverse de la fonction de répartition de la loi normale. Le modèle probit binaire est

Nom du rôle	Description
	utilisé lorsqu'il existe deux catégories de réponses, et le modèle probit cumulé, lorsqu'il en existe plus de deux. ■ log-log complémentaire - affiche la sortie de la fonction logarithmique complémentaire. Le modèle logarithmique complémentaire binaire s'utilise lorsqu'il y a deux catégories de réponse. Le modèle log-log complémentaire cumulé est utilisé lorsqu'il y a plus de deux catégories de réponse. ■ log-log affiche la sortie de la fonction log-log. Le modèle log-log binaire est utilisée lorsqu'il y a deux catégories de réponse. Le modèle log-log cumulé est utilisé lorsqu'il y a plus de deux catégories de réponse. ■ glogit - affiche la sortie de la fonction de logit généralisé. Dans le modèle logit généralisé, chaque catégorie de non-référence est comparée à la catégorie de référence. Remarque : cette option n'est disponible que si vous avez sélectionné le type de réponse Non classée.
Niveaux de réponse	La case Niveaux de réponse répertorie par ordre croissant les niveaux formatés pour la variable de réponse.
Ajuster le modèle au niveau	Vous pouvez spécifier la manière d'ajuster le modèle à un niveau. ■ Si vous avez choisi le type de réponse Binaire, sélectionnez la catégorie d'événement pour le modèle de réponse binaire dans la liste déroulante Ajuster le modèle au niveau. ■ Si vous avez sélectionné Non classé comme type de réponse, sélectionnez la catégorie de référence dans la liste pour le modèle logit généralisé dans la liste déroulante Ajuster le modèle au niveau. Chaque logit contraste une catégorie de non référence avec la catégorie de référence. ■ Si vous avez choisi le type de réponse Classée, vous pouvez utiliser les valeurs de la liste déroulante Ajuster le modèle au niveau pour préciser l'ordre des catégories de réponse. Si vous sélectionnez la première valeur triée, les catégories de réponse sont classées telles qu'elles apparaissent dans la zone Niveaux de réponse. Si vous sélectionnez la dernière valeur triée, les catégories de réponse sont classées dans l'ordre inverse.

Créer un modèle

Vous pouvez créer des modèles dans plusieurs tâches SAS. Pour plus d'informations, voir [Annexe 3, "Créer un modèle" page 749](#).

Spécifier une méthode de sélection du modèle

Dans la liste, sous l'en-tête Modèle, cliquez sur **Sélection** pour accéder à ces options.

Nom d'option	Description
Méthode de sélection de modèle	Sélectionnez une méthode dans la liste déroulante des méthodes de sélection des modèles. Si vous ne choisissez pas de méthode, le modèle entier est utilisé. En fonction de la méthode de sélection du modèle, des options supplémentaires deviennent disponibles. Voici les méthodes disponibles : ■ Modèle complet ajusté (aucune sélection) est la valeur par défaut et ne fournit aucune sélection d'effet. Le modèle est celui créé lors de l'attribution des variables dépendantes et quantitatives. ■ La sélection ascendante commence sans effets dans le modèle. Pour chacune des variables explicatives, cette méthode calcule les statistiques F reflétant la contribution de la variable dans le modèle. Les p-values de ces statistiques F sont comparées au niveau de significativité spécifié pour inclure une variable dans le modèle. La valeur par défaut est 0,05. Pour changer le niveau de significativité, saisissez une valeur dans le champ Pour entrer dans le modèle. Si aucune statistique F n'a un niveau de significativité supérieur à cette valeur, la sélection ascendante s'arrête. Sinon, la méthode de sélection ascendante ajoute la variable dont la statistique F est la plus élevée dans le modèle. La méthode de sélection ascendante recalcule ensuite les statistiques F pour les variables restant en dehors du modèle, et le processus d'évaluation est répété. Ainsi, les variables sont ajoutées une à une au modèle jusqu'à ce qu'aucune variable restante ne produise une statistique F significative. Une fois qu'une variable a été ajoutée au modèle, elle y reste. ■ La sélection descendante commence par calculer les statistiques F pour un modèle, en incluant toutes les variables

Nom d'option	Description
	explicatives. Les variables sont ensuite supprimées du modèle une à une jusqu'à ce que toutes les variables restant dans le modèle produisent des statistiques F significatives. Le niveau de significativité est spécifié dans la zone de texte Pour rester dans le modèle. La valeur par défaut est 0,05. À chaque étape, la variable montrant la plus faible contribution au modèle est supprimée. ■ La sélection pas à pas est une variante de la méthode ascendante. Dans la méthode pas à pas, les variables qui sont déjà dans le modèle n'y restent pas nécessairement. De même que dans la méthode de sélection ascendante, les variables sont ajoutées une à une au modèle si la statistique F est significative. Le niveau de significativité est spécifié dans la zone de texte Pour entrer dans le modèle. Néanmoins, après l'ajout d'une variable, la méthode pas à pas vérifie toutes les variables qui sont déjà incluses dans le modèle et supprime celles qui ne produisent pas une statistique F significative. Le niveau de significativité est spécifié dans la zone de texte Pour rester dans le modèle. Ce n'est qu'après cette vérification et après les suppressions requises qu'une autre variable peut être ajoutée au modèle. Le processus pas à pas prend fin lorsque l'une de ces conditions est remplie : ☐ aucune variable en dehors du modèle n'a une statistique F significative, et chaque variable du modèle est significative au niveau de significativité qui est spécifié pour rester dans le modèle. ☐ la variable à ajouter dans le modèle est la variable qui vient d'en être supprimée. ■ La sélection descendante (rapide sans réajustement du modèle) réalise une élimination descendante rapide. Cette méthode commence avec tous les effets dans le modèle et supprime les effets sans réajuster le modèle.
Niveaux de significativité	■ Pour la méthode de sélection ascendante, indiquez le niveau de significativité à utiliser pour saisir une variable quantitative (explicative) dans le modèle. ■ Pour les méthodes Sélection descendante et Sélection descendante (rapide sans réajustement du modèle), spécifiez un niveau de significativité de sorte qu'une variable quantitative reste dans le modèle. ■ Pour la méthode de sélection pas à pas, indiquez les deux niveaux de significativité.

Nom d'option	Description
Sélectionner le modèle avec la meilleure valeur	Utilise le critère à sélectionner (à chaque étape du processus de sélection) le modèle qui produit la meilleure valeur. Si la valeur optimale du critère spécifié survient pour les modèles à plus d'une étape, le modèle avec le plus petit nombre de paramètres est celui qui est choisi. Si vous ne spécifiez pas de critère, le modèle sélectionné est le modèle à l'étape finale du processus de sélection. Voici les critères disponibles : ■ AIC - Critère d'information d'Akaike ■ AICC - un "small-sample bias" qui est la version corrigée du critère d'information d'Akaike ■ BIC - Critère de Schwarz
Sélectionner les critères pour l'arrêt du processus de sélection	Utilise le critère pour stopper le processus de sélection. Si vous sélectionnez Néant, le processus de sélection s'arrête si aucun candidat d'ajout ou de suppression approprié ne peut être trouvé ou si une limite basée sur la taille est atteinte. Voici les critères disponibles : ■ SL - la significativité du test de score ■ AIC - Critère d'information d'Akaike ■ AICC - un "small-sample bias" qui est la version corrigée du critère d'information d'Akaike ■ BIC - Critère de Schwarz
Effets à forcer dans le modèle	Sélectionnez l'effet que vous voulez ajouter au modèle. La variable est ajoutée dans la zone Effets à forcer dans le modèle. Vous pouvez réorganiser ces variables en les sélectionnant et en utilisant les boutons fléchés. Remarque : cette option n'est pas disponible si vous sélectionnez Ajustement complet du modèle (aucune sélection) comme méthode de sélection de modèle.

Régression logistique High-Performance : Définir les options du modèle

Dans la liste, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Évaluation de l'ajustement du modèle	spécifie ces tests d'ajustement du modèle : ■ Test d'adéquation de Hosmer et Lemeshow affiche le test d'adéquation de Hosmer et Lemeshow dans le cas d'un modèle de réponse binaire. Remarque : cette option n'est disponible que si vous attribuez une variable binaire à la fonction Variable dépendante. ■ R carré généralisé affiche un coefficient généralisé de détermination pour le modèle ajusté.
Niveau de confiance	indique si les intervalles de confiance doivent être calculés sur la base des tests de Wald individuels. Vous pouvez également sélectionner une valeur dans la liste déroulante Niveau de conf. pour définir le niveau de significativité pour les intervalles de confiance des paramètres de régression ou des rapports de cotes .
Supprimer l'affichage des informations de niveau de classe	supprime l'affichage du tableau Informations sur le niveau de classe. Sélectionnez Supprimer le calcul des coefficients de régression logistique pour supprimer le calcul de la matrice de covariance et les erreurs standard des coefficients de régression logistique. Lorsque le modèle contient des milliers de variables, l'inversion de la matrice hessienne pour déduire la matrice de covariance et les erreurs standard des coefficients de régression peut être longue.
Style de codage des variables de classification	spécifie la méthode de paramétrage des variables de classification. ■ GLM s'utilise pour le codage d'une cellule de référence less-than-full-rank. ■ Référence s'utilise pour le codage d'une cellule de référence.
Convertir les variables CLASS en majuscules	indique s'il faut convertir les valeurs de variables de classification alphanumériques en majuscules avant la définition de leurs niveaux. Par exemple, si vous sélectionnez cette option et que les valeurs d'une variable de classification sont 'a', 'A', et 'b', alors 'a' et 'A' représentent le même niveau, et la variable de classification est traitée comme si elle n'avait que deux valeurs : 'A' et 'B'.

Régression High-Performance : Définir les options de performances

Dans la liste, cliquez sur **Performances** pour accéder à ces options.

Nom d'option	Description
Utiliser les paramètres par défaut pour cette tâche	utilise les paramètres par défaut.
Utiliser les paramètres personnalisés pour cette tâche	permet de spécifier le noeud, l'hôte du Grid, l'emplacement de l'installation, et le Data Server pour la tâche. ■ Noeuds spécifie le nombre de noeuds dans l'environnement informatique distribué, à condition que les données ne soient pas traitées avec la base de données. ■ Hôte du Grid spécifie le nom de l'hôte de l'équipement. ■ Emplacement de l'installation spécifie le répertoire dans lequel les bibliothèques partagées de High-Performance Analytics sont installées sur le matériel. ■ Data Server spécifie le nom du serveur sur les systèmes Teradata tels que définis via le fichier d'hôtes et tels qu'utilisés dans l'instruction LIBNAME pour Teradata.
Afficher les étapes détaillées de la procédure	crée une table qui affiche une répartition des étapes de la procédure.

Régression High-Performance : Définir les options de prédiction

Dans la liste, cliquez sur **Prédictions** pour accéder à ces options.

Nom du graphe	Description
Enregistrer les données de sortie	indique si les prédictions et statistiques de diagnostic doivent être enregistrées de façon permanente dans une table de sortie. ■ Prédictions enregistre la probabilité prédite pour chaque niveau de réponse. Pour une variable de réponse Y à trois niveaux, 1, 2 et 3, les probabilités individuelles sont $\Pr(Y=1)$, $\Pr(Y=2)$, et $\Pr(Y=3)$. ■ Statistiques de diagnostic enregistre les statistiques de diagnostic DFBETAS, les diagnostics de décalage de l'intervalle de confiance (C et CBAR), la matrice H, et les estimations du prédicteur linéaire avec les estimations de l'erreur type. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Statistiques supplémentaires	■ Sélectionnez Résidus pour enregistrer ces statistiques : ☐ la modification de l'écart ☐ le résidu d'écart ☐ le résidu de Pearson (Khi-2) ☐ modification de l'écart imputable à la suppression d'une observation donnée ; ☐ modification de la statistique critère qualificatif de lissage du Khi-2 imputable à la suppression d'une observation donnée ; ■ Activez Limites de prédiction pour inclure bornes supérieure et inférieure d'un intervalle de confiance pour chaque prédiction et la moyenne attendue de la variable dépendante dans la sortie.
Sélectionner des variables de sortie	spécifie les colonnes à inclure dans la table de sortie.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.
.....

Histogrammes

<i>À propos de la tâche Histogramme</i>	277
<i>Analyse de capacité : Attribuer des variables aux fonctions d'analyse</i>	278
<i>Histogrammes : Sélectionner la distribution</i>	279
<i>Histogrammes : Personnaliser la distribution</i>	281
Comment accéder aux options de distribution	281
Spécifier les paramètres de la distribution	281
Indiquer les estimations de densité du noyau	281
<i>Analyse de capacité : Définir les options d'apparence du graphe</i>	282
Définir les options d'un axe	282
Ajouter des lignes de référence	283
Ajouter un encart	283
Définir d'autres options de tracé	283
Définir les options des limites de spécification	286
<i>Analyse de capacité : Spécifier les types d'analyse</i>	286
<i>Indiquer des titres et des notes de bas de page</i>	287
<i>Afficher les propriétés</i>	287

À propos de la tâche Histogramme

Une analyse de capacité de processus compare la distribution des sorties d'un processus sous contrôle statistique à ses limites de spécification afin de déterminer la cohérence avec laquelle ces spécifications sont respectées. Les histogrammes sont utilisés dans les analyses de capacité de processus pour comparer la distribution des mesures d'un processus sous contrôle statistique à ses limites de spécification.

Vous pouvez utiliser la tâche Histogramme pour analyser un processus de fabrication. Par exemple, supposons qu'un fabricant de semi-conducteurs produise des circuits imprimés échantillonnés pour déterminer si l'épaisseur du revêtement en cuivre est comprise entre la limite de spécification inférieure de 3,45 mm et la limite supérieure de spécification de 3,55 mm. Chaque barre de l'histogramme

indique le pourcentage d'observations et l'épaisseur mesurée. Les limites de spécification inférieure et supérieure sont incluses dans l'histogramme, de sorte que vous pouvez visualiser le pourcentage d'observations comprises dans ces limites de spécification.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	CAPABILITY
Produits SAS requis	SAS Base, SAS/QC, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Analyse de capacité : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables d'analyse	indique les variables numériques faisant l'objet de l'analyse. Vous devez attribuer au moins une variable à cette fonction. Pour toutes les courbes (sauf la courbe P-P, de probabilités théoriques par rapport aux probabilités calculées) et pour chaque variable d'analyse, vous pouvez indiquer une limite supérieure, une valeur cible et une limite inférieure dans la zone Limites de spécification pour <i>nom_variable</i>. Lorsque vous indiquez ces limites, la sortie inclut automatiquement une table contenant les tests de la normalité, une table contenant les limites de spécification et une table contenant les indices de capacité du processus.

Nom du rôle	Description
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Poids relatif	spécifie la variable à utiliser comme variable de poids relatif. Lorsque vous attribuez une variable à cette fonction, la valeur de la variable de chaque observation est utilisée pour calculer les statistiques pondérées. Vous ne pouvez attribuer qu'une seule variable à cette fonction. Remarque : lorsque vous créez un histogramme, une courbe de probabilité, une courbe P-P ou une courbe CDF, il est déconseillé d'attribuer une variable à la fonction Poids relatif. Si vous attribuez malgré tout une variable à cette fonction, le résultat doit être interprété avec précaution.
Variables de classification (disponible pour la tâche d'histogramme uniquement)	indique une ou deux variables à utiliser en tant que variables de classification. Lorsque vous attribuez une variable à cette fonction, la tâche crée un histogramme comparatif vous permettant de comparer la distribution d'une variable d'analyse sur des niveaux de classification. Remarque : pour créer des courbes de probabilité comparatives ou des courbes Q-Q comparatifs, utilisez la tâche Analyse de distribution.

Histogrammes : Sélectionner la distribution

Dans la liste, sous l'en-tête Distributions, cliquez sur **Récapitulatif** pour accéder à ces options.

Vous pouvez ajuster plusieurs distributions théoriques sur un seul histogramme. Après avoir sélectionné la distribution pour le calcul, sélectionnez le nom de la distribution dans la liste pour spécifier les valeurs de paramètres et les propriétés de courbe pour cette distribution.

Nom d'option	Description
Distributions	Utilisez les cases à cocher pour sélectionner une ou plusieurs de ces courbes de densité : ■ Normal ajuste une densité normale avec une moyenne (μ) et un écart-type (σ). ■ Lognormal ajuste une densité lognormale avec un paramètre d'échelle (ζ), un paramètre de seuil (θ) et un paramètre de forme (σ). ■ Exponentiel ajuste une densité exponentielle avec un paramètre d'échelle (σ) et un paramètre de seuil (θ). ■ Weibull ajuste une densité Weibull à trois paramètres avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et un paramètre de forme (c). ■ Bêta ajuste une densité bêta avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et des paramètres de forme (α et β). ■ Gamma ajuste une densité gamma avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et un paramètre de forme (α). ■ Kernel ajuste l'estimation non-paramétrique de densité de noyau. Remarque : pour les histogrammes comparatifs, seules la distribution Normale et les estimations de densité de Noyau sont disponibles.
Style des graphiques	Sélectionnez le style des graphiques à utiliser dans la sortie. Vous avez le choix entre les graphiques SAS traditionnels et les graphiques statistiques ODS. Pour plus d'informations sur les graphiques ODS, voir SAS Output Delivery System: User's Guide.

Histogrammes : Personnaliser la distribution

Comment accéder aux options de distribution

Dans la liste, sous l'en-tête Distributions, cliquez sur le nom de la distribution que vous voulez personnaliser pour pouvoir accéder à ces options.

Remarque : ces options de distribution ne sont disponibles que si vous avez sélectionné la distribution au préalable sur la page **Récapitulatif**, ou bien si vous activez la case à cocher **Nom de la distribution** sur la page de cette distribution.

Spécifier les paramètres de la distribution

Pour les distributions normales, lognormales, exponentielles, Weibull, bêta et gamma, vous pouvez spécifier les paramètres pour toutes les variables d'analyse ou pour chacune d'entre elles. Par défaut, la tâche estime les valeurs de paramètre de toutes les variables d'analyse.

Pour définir les paramètres de chaque variable d'analyse :

- 1 Désactivez la case à cocher **Appliquer la distribution à toutes les variables**.
- 2 Activez la case à cocher de la variable dans la zone **Variables d'analyse**.
- 3 Spécifiez les valeurs des paramètres.

Indiquer les estimations de densité du noyau

Pour les estimations non paramétriques de densité du noyau, vous pouvez indiquer la largeur de bande (c) et la fonction du noyau (k). Par défaut, la tâche estime une largeur de bande qui minimise l'erreur au carré intégrée de la moyenne approximative (MISE). Si vous ne souhaitez pas estimer la largeur de bande, désactivez la case à cocher **Minimiser MISE (Mean Integrated Squared Error)**. Pour spécifier le paramètre de largeur de bande normalisée de la densité du noyau, saisissez la largeur de bande dans le champ de saisie. Vous pouvez saisir jusqu'à

cinq valeurs séparées par un espace. Par exemple, entrez ces valeurs dans la zone de texte : 0,5 1,0 1,5.

Remarque : si vous ne désactivez pas la case à cocher **Minimiser MISE (Mean Integrated Squared Error)**, la valeur MISE est considérée comme l'une des cinq valeurs de largeur de bande possibles.

Vous pouvez également définir la fonction du noyau (normale, quadratique ou triangulaire) utilisée pour calculer l'estimation de la densité du noyau. La fonction du noyau par défaut est **Normale**.

Vous pouvez définir des fonctions de noyau pour un maximum de cinq estimations. Si vous définissez plusieurs fonctions, leur ordre doit toujours être le suivant : normale, quadratique, triangulaire. Pour ajouter des fonctions ou en modifier l'ordre, vous devez modifier le code de la tâche.

Si vous définissez plus de fonctions de noyau que de largeurs de bande, la dernière largeur de bande est reprise pour les estimations restantes. Par exemple, si vous définissez les largeurs de bande 0,5 et 1 et les fonctions de noyau normale, quadratique et triangulaire, la largeur de bande 1 est utilisée pour la fonction triangulaire. Si vous définissez plus de largeurs de bande que de fonctions de noyau, la dernière fonction de noyau est reprise pour les largeurs de bande restantes.

Par exemple, si vous indiquez 0,5, 1 et 1,5 pour définir les largeurs de bande, et normale, quadratique et triangulaire pour définir les fonctions de noyau, la fonction quadratique est utilisée pour la largeur de bande 1,5.

Analyse de capabilité : Définir les options d'apparence du graphe

Définir les options d'un axe

Dans la liste, sous l'en-tête **Apparence** ⇒ **Axes**, cliquez sur **Axes** pour accéder à ces options.

Vous pouvez définir ces options :

- la couleur des axes
- la couleur des graduations
- un libellé pour les axes horizontaux et verticaux

Ajouter des lignes de référence

Dans la liste, sous l'en-tête **Apparence** ⇒ **Axes**, cliquez sur **Horizontal** ou **Vertical** pour accéder à ces options.

Pour utiliser les lignes de référence, sélectionnez **Utiliser les lignes de référence**. Vous pouvez formater la ligne de référence et spécifier l'emplacement de chaque ligne de référence.

Ajouter un encart

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Encart** pour accéder à ces options.

Vous pouvez améliorer un graphe en ajoutant une case ou une table (appelée encart) contenant les statistiques descriptives directement sur le graphe. Si vous créez un histogramme comparatif, un encart apparaît dans chaque histogramme.

Sélectionnez **Inclure l'encart** pour ajouter des statistiques descriptives au graphe. Sélectionnez ensuite les statistiques que vous voulez inclure dans l'encart.

Remarque : les statistiques permettant de calculer les indices de capacité ne sont disponibles que si vous avez créé des limites de spécification pour la variable attribuée à la fonction **Variables d'analyse**. Vous pouvez créer des limites de spécification pour les histogrammes, les courbes CDF, les courbes de probabilité et les diagrammes Q-Q. Pour plus d'informations sur le calcul d'indices de capacité, consultez l'aide relative à la procédure CAPABILITY.

Définir d'autres options de tracé

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Options** pour accéder à ces options.

Les options d'apparence disponibles dépendent du type de graphe créé par la tâche.

Nom du rôle	Description
Arrière-plan	spécifie la couleur d'arrière-plan du graphe.
Contour de la barre et remplissage de la barre	spécifie les couleurs pour les barres d'un histogramme.

Nom du rôle	Description
	Remarque : ces options ne sont disponibles que pour les histogrammes.
Inverser les deux axes en plaçant les centiles sur l'axe opposé	inverse les axes du graphe. Si vous sélectionnez cette option, les options et les lignes de référence que vous avez sélectionnées pour les axes risquent d'être incorrectes en raison du changement d'orientation du graphe. Remarque : cette option n'est disponible que pour les courbes de probabilité et les courbes Q-Q.
Supprimer légendes par défaut	empêche l'affichage de la légende par défaut avec le graphe.
Utiliser des milieux	spécifie les milieux dans l'histogramme. Les options disponibles dans la liste déroulante varient selon que vous créez un histogramme ou un histogramme comparatif. Si vous attribuez une variable à la fonction Variables de classification, vous créez un histogramme comparatif. Pour les histogrammes comparatifs, ces options sont disponibles : ■ Spécifier les milieux permet d'indiquer les milieux de début et de fin. Vous pouvez également spécifier l'intervalle d'escalier à utiliser pour calculer les autres milieux. Le milieu correspond à la valeur située au milieu de la plage de valeurs que représente chaque barre. ■ Clé détermine les milieux des données de la cellule clé. Par défaut, les niveaux de la variable de classification s'affichent de haut en bas (de gauche à droite) dans l'ordre croissant des valeurs internes (non formatées) de la première variable de classification. Si vous n'avez spécifié qu'une seule de ces variables, la cellule clé correspond au niveau qui apparaît le premier dans cet ordre. Si vous en avez spécifié deux, la cellule clé correspond à une combinaison des niveaux des variables 1 et 2 qui apparaissent en premier dans cet ordre. Ainsi, le choix de la cellule clé détermine l'axe horizontal uniforme utilisé pour les autres cellules. La liste des milieux de la cellule clé est donc étendue dans l'une ou l'autre direction si besoin jusqu'à ce qu'elle couvre les données dans les cellules restantes. ■ Uniforme détermine le nombre de milieux en fonction de la taille de l'échantillon complet. Si vous n'attribuez pas de variable à la fonction Variables de classification, vous créez un histogramme. Les options disponibles sont alors les suivantes : ■ Spécifier les milieux permet d'indiquer les milieux de début et de fin. Vous pouvez également spécifier l'intervalle d'escalier à utiliser pour calculer les autres milieux. Le

Nom du rôle	Description
	milieu correspond à la valeur située au milieu de la plage de valeurs que représente chaque barre. ■ Indiquer le nombre de classes permet d'indiquer le nombre de classes (également appelées intervalles d'histogrammes) des données. Par défaut, le nombre de classes affichées dépend de l'étendue des données. Cependant, vous pouvez indiquer un maximum. Remarque : vous ne pouvez indiquer des milieux que si vous utilisez la tâche Histogramme.
Indiquer taille du graphique personnalisé (en pixels)	spécifie la taille de l'image. À titre d'indication, la taille totale de l'écran est affichée dans les zones Largeur et Hauteur .

Remarque : des options de comparaison sont disponibles pour les histogrammes comparatifs. Vous pouvez créer un histogramme comparatif en attribuant une variable à la fonction **Variables de classification** dans la tâche Histogramme.

Nom du rôle	Description
Lignes et Colonnes	Si vous attribuez au moins une variable de classification à la fonction d'analyse, vous pouvez utiliser les cases Lignes par page et Colonnes par page pour modifier la disposition des colonnes et des lignes dans l'histogramme comparatif. Si vous attribuez une variable de classification, la disposition par défaut est de deux lignes et d'une colonne par page. Si vous attribuez deux variables de classification, la disposition par défaut est de deux lignes et de deux colonnes par page. Vous pouvez également utiliser les listes déroulantes Côté du cadre et Haut du cadre afin d'indiquer les couleurs utilisées pour remplir les zones de libellés de ligne et libellés de colonne.
Nombre de classes	Par défaut, le nombre de classes affichées dépend de l'étendue des données. Cependant, vous pouvez indiquer un maximum. Vous pouvez également spécifier un écart-type (par exemple, le nombre de classes affichées), sur la base d'une valeur de l'écart-type comprise au-dessus et en dessous de la moyenne des données.

Définir les options des limites de spécification

Remarque : ces options sont disponibles pour les histogrammes, les courbes de probabilité, les graphiques Q-Q et les courbes CDF.

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Limites de spécification** pour indiquer la limite de spécification inférieure, la cible, et la limite de spécification supérieure.

Remarque : les options **Limites de spécification** ne sont disponibles que si vous indiquez une ou plusieurs limites pour au moins une variable d'analyse. Pour indiquer une limite de spécification pour une variable d'analyse, cliquez sur **Données** dans la liste, puis sélectionnez la variable d'analyse.

Analyse de capabilité : Spécifier les types d'analyse

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Par défaut, le rapport Analyse de capabilité se compose des tables suivantes : une table d'intervalles de confiance de base, une table de mesures de base, une table de tests de position et une table de moments.

Pour ajouter une table au rapport, sélectionnez la table dans la zone **Tables**.

Pour spécifier les options d'analyse, sélectionnez la table dans la zone **Tables**. Si les options sont disponibles, une case s'affiche sur la droite avec les options à spécifier.

Pour supprimer toutes les tables de statistiques descriptives du rapport, activez la case à cocher **Supprimer stat. descriptives et indices de capabilité**. La sélection de cette option ne supprime pas ces tables :

- tables créées par la distribution sélectionnée. Pour supprimer ces tables, activez la case à cocher **Supprimer les tables de distribution** dans le panneau correspondant à cette distribution.
- Tables créées par l'instruction INTERVALS. Pour supprimer ces tables, vous pouvez spécifier l'option NOPRINT. Pour plus d'informations sur l'instruction INTERVALS, consultez l'aide relative à la procédure CAPABILITY.

Pour enregistrer vos résultats dans une table de sortie, activez la case à cocher **Enregistrer les statistiques de sortie dans une table**. La table inclura les statistiques à une variable et les indices de capabilité.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,

- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Carte aux mesures individuelles

<i>À propos de la tâche Carte aux mesures individuelles</i>	289
<i>Carte aux mesures individuelles : Attribuer des variables aux fonctions d'analyse</i>	290
<i>Cartes de contrôle : Indiquer des limites de contrôle</i>	291
<i>Carte de contrôle : Sélectionner les tests à effectuer</i>	293
<i>Cartes de contrôle : Définir les options d'apparence du graphe</i>	294
<i>Cartes de contrôle : Enregistrer les résultats d'analyse</i>	294
<i>Indiquer des titres et des notes de bas de page</i>	295
<i>Afficher les propriétés</i>	296

À propos de la tâche Carte aux mesures individuelles

La carte de contrôle de Shewhart est un outil graphique et analytique qui permet de déterminer si un processus est sous contrôle statistique. La tâche Carte aux mesures individuelles crée des cartes de contrôle pour chaque mesure et pour les étendues mobiles. Ces cartes sont appropriées si une seule mesure est disponible pour chaque échantillon de sous-groupe et si les mesures sont indépendantes et distribuées normalement.

Vous pouvez utiliser cette tâche pour analyser un processus de fabrication. Par exemple, supposons qu'un constructeur aéronautique qui produit des moteurs à réaction mesure le diamètre intérieur de la face avant de chaque moteur (en centimètres). Les mesures de diamètre de 20 moteurs sont enregistrées dans une table SAS. Chaque point des cartes aux mesures individuelles indique le diamètre interne d'un moteur. Chaque point des cartes d'étendue mobile indique l'intervalle des deux mesures les plus récentes. Si toutes les mesures individuelles et si les étendues mobiles sont comprises dans les limites de contrôle, vous pouvez en conclure que le processus est sous contrôle statistique.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	SHEWHART
Produits SAS requis	SAS Base, SAS/QC, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Carte aux mesures individuelles : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Mesure	indique les variables de traitement à analyser. Ces variables contiennent les mesures individuelles. Vous devez attribuer au moins une variable à cette fonction.
Identificateur de mesure	indique la variable qui identifie les sous-groupes dans les données. Les valeurs de la variable indiquent de quelle manière les observations de la table d'entrée sont organisées en sous-groupes rationnels. En général, les valeurs sont les suivantes : ■ les indices qui donnent l'ordre dans lequel les sous-échantillons sont collectés ■ les dates ou heures auxquelles les sous-échantillons ont été collectés ■ les libellés qui identifient de façon unique les échantillons de sous-groupe Un identifiant de la mesure est requis et vous ne pouvez attribuer qu'une variable à cette fonction.

Nom du rôle	Description
	Pour trier les données selon les sous-groupes identifiés, activez la case à cocher Trier par sous-groupe .
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.
Variables de bloc	indique deux variables maximum, qui groupent les données par blocs de sous-groupes consécutifs. Ces blocs sont libellés dans une légende, et chaque variable de bloc fournit un niveau pour les libellés de la légende.

Cartes de contrôle : Indiquer des limites de contrôle

Dans la liste, cliquez sur **Limites de contrôle** pour accéder à ces options.

Nom d'option	Description
Limites Sigma	spécifie la largeur des limites de contrôle commun multiple de l'erreur type (sigma) des statistiques récapitulatives du sous-groupe tracé dans le graphique. La largeur doit être positive. Le multiple par défaut est 3.
Sélectionner une méthode de calcul	spécifie la méthode de calcul pour les limites de contrôle. Vous avez le choix entre ces options : ■ Calculer les limites de contrôle à partir des données actives ■ Calculer les limites de contrôle à partir de la table sélectionnée Pour désigner la table contenant les limites de contrôle ou les paramètres à partir desquels les limites de contrôle peuvent être calculées, cliquez sur

Nom d'option	Description
	Parcourir. Sélectionnez la table à utiliser et cliquez sur Ouvrir. Remarque : la table que vous sélectionnez doit être une table LIMITS=. Pour plus d'informations sur les variables requises dans une table LIMITS=, consultez la documentation des procédures du graphique que vous êtes en train de créer. Pour représenter les statistiques descriptives de tous les sous-groupes, activez la case à cocher Trace les statistiques descriptives de tous les sous-groupes, que la taille de l'échantillon du sous-groupe soit égale à la taille nominale des limites de contrôle ou non. Vous pouvez représenter ces statistiques lorsque la taille de l'échantillon est la même pour pratiquement tous les sous-groupes et lorsque vous voulez afficher les limites de contrôle fixes (non soumises à variation) correspondant à la taille d'échantillon nominale n. Si vous sélectionnez cette option alors que les tailles des échantillons de vos sous-groupes sont très différentes, l'interprétation des limites de contrôle n'a de sens que pour les sous-groupes dont la taille de l'échantillon est égale à n. Si vous souhaitez ajouter des symboles spéciaux pour les points de données où la taille de l'échantillon est différente de n, activez la case à cocher Ajouter des symboles spéciaux pour les points correspondant aux tailles d'échantillon différentes de la taille nominale pour les limites de contrôle fixes. ■ Spécifier les limites de contrôle Dans les champs Supérieure, Centrale, et Inférieure, indiquez respectivement la limite de contrôle supérieure, la ligne centrale et la limite de contrôle inférieure. Si l'effectif du sous-groupe varie, vous devez indiquer un effectif nominal. Remarque : si vous attribuez des variables à la fonction Grouper l'analyse par, vous ne pouvez pas saisir des valeurs pour la limite de contrôle.

Carte de contrôle : Sélectionner les tests à effectuer

Dans la liste, cliquez sur **Tests** pour accéder à ces options.

Remarque : ces tests ne s'appliquent pas aux limites de contrôle qui ne sont pas trois limites sigma ou qui varient en fonction de la taille de l'échantillon du sous-groupe.

Nom d'option	Description
Sélectionner un test	Utilisez les cases à cocher pour demander un ou plusieurs tests pour des sujets spéciaux, également connus sous les noms de test de dérive : Run test, Pattern et Western Electric tests. Ces tests détectent les dérives particulières non aléatoires dans les points représentés sur le graphique de contrôle principal. L'occurrence du modèle non aléatoire est appelé signal. Les tests peuvent indiquer la présence d'une cause spéciale de variation. Lorsque vous sélectionnez l'un des tests, une description de ce test apparaît dans la zone Description.
Libellé	Utilisez les champs de saisie afin de saisir des libellés pour les points où les tests pour causes spéciales sont positifs. Les libellés ne peuvent contenir plus de 16 caractères. Chaque libellé apparaît à tous les points du test où un signal est rencontré.
Texte identifiant les signaux de test	■ Utilisez la liste déroulante afin de sélectionner une couleur pour les libellés que vous avez indiqués. ■ Utilisez la case à cocher Afficher lignes de zone pour inclure des lignes qui représentent les zones A, B et C sur un graphique principal. Ces zones se construisent en divisant l'intervalle entre les limites de contrôle en six sous-intervalles espacés de façon égale. ■ Utilisez la case Surpasser la limite 3 sigma pour tester des sujets spéciaux si vous spécifiez une limite de contrôle sigma différente du multiple par défaut de 3. Pour changer la limite sigma, cliquez sur Limites de contrôle dans le volet de sélection.

Nom d'option	Description
	■ Cochez la case Appliquer le test aux dérives des points se chevauchant afin d'appliquer des tests pour les causes spéciales aux dérives de points se chevauchant.

Cartes de contrôle : Définir les options d'apparence du graphe

Certaines options d'apparence sont communes à plusieurs tâches de la carte de contrôle. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir les options d'un axe" page 743](#)
- ["Ajouter des lignes de référence" page 744](#)
- ["Définir d'autres options de tracé" page 744](#)
- ["Définir les options de variable de bloc" page 746](#)

Cartes de contrôle : Enregistrer les résultats d'analyse

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Pour enregistrer les résultats de l'analyse, vous pouvez créer des tables de sortie contenant plusieurs statistiques. Activez les cases à cocher pour les données de sortie que vous voulez enregistrer de façon permanente.

Nom du rôle	Description
Statistiques de sous-groupe et table des limites de contrôle	crée une table de sortie contenant les informations représentées sur le graphe. Cela inclut les valeurs de la variable de sous-groupe ainsi que les statistiques descriptives correspondantes et les limites de contrôle. Remarque : cette option crée une table OUTTABLE=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option TABLE= décrite par la procédure SHEWHART. Pour plus

Nom du rôle	Description
	d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des statistiques de sous-groupe	crée une table de sortie contenant les statistiques descriptives du sous-groupe. Remarque : cette option crée une table OUTHISTORY=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option HISTORY= décrite par la procédure SHEWHART. Pour plus d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des limites de contrôle	crée une table de sortie qui enregistre les limites de contrôle. Vous pouvez utiliser cette table lorsque vous sélectionnez la méthode de calcul pour les limites de contrôle lors d'une utilisation ultérieure de cette tâche.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Tables de survie

<i>À propos de la tâche Tables de survie</i>	297
<i>Tables de survie : Attribuer des variables aux fonctions d'analyse</i>	298
<i>Tables de survie : Définir les options de la méthode</i>	300
<i>Tables de survie : Générer des graphes</i>	301
<i>Tables de survie : Définition des options des résultats</i>	302
<i>Indiquer des titres et des notes de bas de page</i>	303
<i>Afficher les propriétés</i>	304

À propos de la tâche Tables de survie

La tâche Tables de survie calcule les estimations non paramétriques de la distribution de survie des données qui pourraient être censurées de droit du fait de l'abandon ou de la fin de l'étude. Pour ces données, vous savez seulement que la durée de vie dépasse la valeur indiquée. Généralement, vous voulez comparer les courbes de survie pour déterminer si deux échantillons proviennent de la même distribution de survie.

Cette tâche calcule les tests de rang et un test de rapport des vraisemblances pour tester l'homogénéité des fonctions de survie à travers les niveaux. La tâche génère également des graphes qui permettent d'examiner les courbes de survie.

Vous pouvez utiliser cette tâche pour étudier la distribution de survie d'une population spécifique. Par exemple, supposons que vous vouliez étudier le délai de survie de 40 rats exposés à un agent carcinogène au cours d'une petite expérience aléatoire, répartis dans un ou deux groupes de traitements. Le délai de survie est le délai qui s'écoule entre la sélection aléatoire et la mort. L'événement présentant un intérêt est la mort provoquée par le cancer induit par l'agent carcinogène. Vous pouvez utiliser la tâche Tables de survie pour déterminer si les distributions de survie diffèrent entre les deux groupes de traitements.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	LIFETEST
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Tables de survie : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Temps de survie	indique la variable à utiliser en tant que variable de temps de survie. Les temps de survie sont souvent appelés temps de défaillance et les temps d'événement sont des temps de survie non tronqués. Le temps de survie est requis et vous pouvez attribuer une seule variable à cette fonction.
Variable de coupure	indique quelle variable utiliser comme variable de coupure. Dans la zone Valeurs de coupure à droite , sélectionnez les valeurs de cette variable qui indiquent si un temps de survie est tronqué. Les valeurs de la variable de coupure doivent être numériques, non manquantes. La variable de coupure n'est pas obligatoire. Cependant, si vous l'utilisez, vous ne pouvez indiquer qu'une variable et vous devez indiquer les valeurs de coupure.
Variables de niveau (de discrétisation)	indique les variables dont les valeurs déterminent l'amplitude des niveaux de discrétisation. Si vous n'indiquez pas de points extrêmes pour les variables numériques, les niveaux sont déterminés par la combinaison des niveaux

Nom du rôle	Description
	des valeurs uniques des variables de niveau de discrétisation. Vous pouvez activer la case à cocher pour permettre aux valeurs manquantes de former des amplitudes de niveau de discrétisation possibles. Pour les variables de niveau numérique, vous pouvez indiquer les intervalles de niveaux de discrétisation en cliquant sur Indiquer des intervalles dans la zone des Amplitude de niveau. Définissez ensuite les intervalles en ajoutant des points extrêmes. Par défaut, les valeurs uniques forment les amplitudes de niveau de discrétisation. Indiquez les points extrêmes déterminant la stratification. L'intervalle au-dessous d'un point extrême ne contient pas le point extrême et l'intervalle au-dessus du point extrême contient le point extrême. Pour indiquer les points extrêmes d'une variable, saisissez des valeurs et cliquez sur Ajouter. Les points extrêmes ne se limitent pas aux valeurs entières. Vous pouvez saisir soit des valeurs simples séparées par des virgules ou des espaces, soit des intervalles. Par exemple, saisissez 5 10 15 pour produire des niveaux uniques de discrétisation pour les valeurs qui tombent dans les intervalles $\{[-\infty, 5), [5, 10), [10, 15), [15, \infty)\}$. L'intervalle 5 ... 10 produit des niveaux uniques pour les valeurs tombant dans les intervalles suivants : $\{[-\infty, 5), [5, 6), [6, 7), [7, 8), [8, 9), [9, 10), [10, \infty)\}$. Si vous attribuez plusieurs variables à cette fonction, définissez les intervalles séparément pour chaque variable.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Covariables du test	indique une liste de covariables numériques (continues) utilisées pour tester l'association avec le temps de défaillance. Deux groupes de statistiques de rang sont calculés. Les statistiques et leurs variances sont groupées. Les statistiques du test (marginales) sont affichées pour chacune des covariables. Les statistiques du test sont également affichées pour les effets joints des covariances.
Variables ID	indique une ou plusieurs variables à utiliser comme variables ID. Les valeurs des variables ID sont utilisées pour attribuer des libellés aux observations des estimations de survie.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les

Nom du rôle	Description
	données sont triées par ordre croissant des valeurs de ces variables de groupe.
	Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables .
	Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une fonction d'analyse.

Tables de survie : Définir les options de la méthode

Dans la liste, cliquez sur **Méthodes** pour accéder à ces options.

Nom d'option	Description
Méthode d'estimation	Pour indiquer une méthode d'estimation, cliquez sur le bouton Estimation de Kaplan-Meier ou Table de survie. Les deux méthodes génèrent des résultats similaires. Cependant la méthode de Kaplan-Meier est basée sur les temps de survie réels et la table de survie groupe les temps de survie par intervalles. Lorsque la taille de l'échantillon est très élevée, il peut être utile de grouper les temps de survie par intervalles et d'effectuer l'analyse de table de survie. La méthode par défaut est la méthode de Kaplan-Meier.
Intervalles	Si vous sélectionnez la méthode d'estimation de Kaplan-Meier, vous pouvez réduire le nombre d'estimations de survie affichées en définissant des points extrêmes pour chaque intervalle. La tâche affiche uniquement les estimations pour la durée la plus courte dans chaque intervalle spécifié. Si vous sélectionnez Table de survie comme méthode d'estimation, vous pouvez spécifier des informations sur la création des intervalles utilisés dans le calcul de la fonction de survie. Par défaut, 10 intervalles sont utilisés. Vous pouvez spécifier une autre valeur en sélectionnant Nombre et en saisissant la valeur de votre choix. Vous pouvez également créer des intervalles en indiquant la largeur de l'intervalle. Sélectionnez Largeur et entrez votre valeur.

Nom d'option	Description
	Si vous voulez définir des points extrêmes exacts à utiliser pour la création des intervalles, sélectionnez Points extrêmes .
Niveau de confiance pour la fonction de survie	Vous pouvez soit saisir une valeur pour un niveau de confiance soit en sélectionner une dans la liste déroulante. La valeur par défaut est 95 pour cent.

Tables de survie : Générer des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Vous pouvez produire des courbes sur les fonctions de survie estimées, les fonctions de hasard et les fonctions de densité de probabilité par rapport au temps.

Nom d'option	Description
Afficher la courbe de fonction de survie	Cette option génère un graphe de la fonction de distribution de survie, qui s'utilise pour décrire la durée de vie d'une population d'intérêt. ■ Pour inclure l'intervalle de confiance, activez la case à cocher Inclure des intervalles de confiance instantanés. L'intervalle de confiance pour les courbes est contrôlé par le paramètre Méthodes pour les intervalles de confiance. La valeur par défaut est 95 pour cent. ■ Pour afficher toutes les courbes de différents niveaux sur un seul graphe assorti d'une légende, activez la case à cocher Recouvrir niveau (de discrétisation) sur graphe simple. Si vous avez attribué une variable à la fonction Variables de niveau (de discrétisation) et activé la case à cocher Afficher la courbe de fonction de survie, toutes les courbes des différents niveaux sont par défaut affichées sur un seul graphe. Remarque : la case à cocher Recouvrir niveau (de discrétisation) sur graphe simple n'est activée que si une variable est attribuée à la fonction Variables de niveau (de discrétisation). Si plus d'une variable a été attribuée, vous ne pouvez pas sélectionner cette option.
Probit(1 - fonction de survie)	Cette option génère une courbe de probit de 1 moins la fonction de survie estimée par rapport au logarithme de temps. Ce type de graphe s'utilise pour évaluer les

Nom d'option	Description
	distributions lognormales. Si le modèle lognormal est approprié, la courbe doit être approximativement linéaire.
-Log(Fonction de survie)	Cette option génère une courbe du logarithme négatif de la fonction de survie estimée par rapport au temps. Ce type de graphe s'utilise pour évaluer les distributions exponentielles. Si le modèle exponentiel est approprié, la courbe doit être approximativement linéaire.
Log(-log(Fonction de survie))	Cette option génère une courbe du logarithme négatif de la fonction de survie estimée par rapport au logarithme de temps. Ce type de graphe s'utilise pour évaluer une distribution de Weibull. Si le modèle de Weibull est approprié, la courbe doit être approximativement linéaire.
Estimation des tables de survie	Pour générer des courbes sur la fonction de hasard et la densité de probabilité par rapport au temps, vous devez d'abord sélectionner la méthode de table de survie pour calculer les estimations de la fonction de survie. Vous pouvez ensuite activer les cases à cocher Fonction de hasard ou Densité de probabilité .

Tables de survie : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Enregistrer les données de sortie	crée des tables de sortie contenant les estimations de survie ainsi que diverses statistiques. ■ Activez la case à cocher Estimations de survie pour créer une table de sortie contenant les estimations de la fonction de survie et l'intervalle de confiance correspondant pour tous les niveaux de discrétisation. ■ Activez la case à cocher Statistiques pour créer une table de sortie contenant les statistiques globales du test du Chi-deux à associer avec le temps de défaillance pour les variables contenues dans les fonctions Covariables du test que vous avez sélectionnées. Cette table inclut également

Nom d'option	Description
	les valeurs des statistiques de rang à une variable et leur matrice de covariance estimée. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Supprimer l'affichage des résultats	N'affiche pas les résultats. Si vous activez cette case à cocher, les options du volet Graphes ne sont pas disponibles. Par défaut, les résultats sont affichés.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

.....

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Courbe

<i>À propos de la tâche Courbe</i>	305
<i>Courbe : Sélectionner un type de graphe</i>	306
<i>Courbe : Attribuer des variables aux fonctions d'analyse</i>	307
<i>Courbe : Définir les options d'apparence</i>	309
Définir d'autres options de tracé	309
Options d'apparence supplémentaires	310
<i>Indiquer des titres et des notes de bas de page</i>	310
<i>Afficher les propriétés</i>	311

À propos de la tâche Courbe

La tâche Courbe crée une courbe, un nuage de points, une courbe spline, un diagramme en bâtons ou en escalier, un graphique d'analyse de régression, une courbe lissée, STD, d'interpolation de Lagrange ou de superposition qui représente la relation mathématique entre des variables, en mettant en évidence les tendances ou les motifs des points de données.

Par exemple, une courbe permet de montrer s'il existe une relation entre les frais de publicité et les ventes d'un produit donné.

Pour plus d'informations, voir ["Sélectionner un format de sortie pour les graphiques" page 6](#).

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GPLOT
Produits SAS requis	SAS Base, SAS/STAT

Nom du prérequis	Noms des procédures et des produits
Autres produits SAS recommandés	SAS/GRAPH

Courbe : Sélectionner un type de graphe

Dans la liste, cliquez sur **Courbe** pour accéder à ces options.

Nom du rôle	Description
Courbe	crée une courbe où les points de données sont reliés par des droites. Les points sont reliés dans l'ordre dans lequel ils apparaissent dans la table d'entrée.
Courbe spline	spécifie qu'une routine de lissage sera utilisée pour l'interpolation de la courbe.
Diagramme en bâtons	trace un trait vertical à partir de chaque point de données jusqu'à un trait horizontal à la valeur 0 sur l'axe vertical ou à la valeur minimale sur l'axe vertical si celle-ci est supérieure à 0. La ligne horizontale est tracée automatiquement.
Diagramme en escaliers	indique que les données sont tracées à l'aide d'une fonction par paliers. Dans la courbe, les escaliers sont joints par une ligne verticale. Avant la représentation, les données sont triées par la variable indépendante.
Nuage de points avec courbe de régression	crée un nuage de points et ajuste la meilleure droite de régression linéaire pour les données. Par défaut, les droites de régression ne sont pas forcées par les origines de courbe et l'intervalle de confiance n'est pas affiché.
Courbe lissée	spécifie un ajustement d'une courbe lissée par rapport aux données à l'aide d'une routine de lissage. Cette méthode permet de lisser les données bruyantes. Les points du tracé ne se trouvent pas systématiquement sur la ligne. Avant la représentation, les données sont triées par la variable indépendante.
Courbe avec écart-type	spécifie qu'une ligne pleine relie la valeur moyenne de la variable Y avec deux écarts-types pour chaque valeur de la variable X.

Nom du rôle	Description
Courbe d'interpolation de Lagrange	spécifie qu'un polynôme de Lagrange du premier degré soit utilisé pour lisser la courbe. Avant la représentation, les données sont triées par la variable indépendante.
Courbes à colonne verticale avec superposition	crée une courbe distincte pour chaque variable Y par rapport à une même variable X. Toutes les courbes apparaissent dans le même graphique.
Plusieurs courbes par colonne de groupe	crée une courbe séparée pour chaque valeur distincte de la variable de groupe. Toutes les courbes apparaissent dans le même graphique.
Nuage de points	crée un nuage de points qui présente les relations entre deux ou trois variables en mettant en évidence des motifs ou des concentrations de points de données.

Courbe : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Horizontal	La colonne que vous attribuez à cette fonction est la variable d'axe horizontal ou X du graphique.
Vertical	La colonne que vous attribuez à cette fonction est la variable d'axe vertical ou Y du graphique. Vous pouvez créer un récapitulatif statistique des valeurs uniques de la variable d'axe horizontal. Sélectionnez la variable sous la fonction Vertical . Sélectionnez Agréger pour

Nom du rôle	Description
	chaque valeur horizontale distincte et sélectionnez la fonction d'agrégation. Par exemple, vos données d'entrée répertorient les ventes et les dépenses pour chaque pays et par région. Vous attribuez la variable Country à la fonction Horizontal et la variable Sales à la fonction Vertical. Si vous voulez représenter la somme des ventes pour chaque pays, activez la case à cocher, puis sélectionnez Somme dans la liste déroulante. La tâche ajoute les ventes régionales pour un pays et présente une seule valeur par pays dans le résultat final.
Vertical (Droite)	La colonne que vous attribuez à cette fonction est la variable d'axe vertical ou Y du côté droit du graphique. Cette variable est représentée par rapport à la variable d'axe X de sorte qu'un autre graphe soit généré dans le diagramme. Vous pouvez créer un récapitulatif statistique des valeurs uniques de la variable d'axe horizontal. Sélectionnez la variable sous la fonction Vertical. Sélectionnez Agréger pour chaque valeur horizontale distincte et sélectionnez la fonction d'agrégation. Par exemple, vos données d'entrée répertorient les ventes et les dépenses pour chaque pays et par région. Vous attribuez la variable Country à la fonction Horizontal et la variable Expenses à la fonction Vertical (Droite). Si vous voulez représenter la somme des dépenses pour chaque pays, activez la case à cocher, puis sélectionnez Somme dans la liste déroulante. La tâche ajoute les dépenses régionales pour un pays et présente une seule valeur par pays dans le résultat final. Remarque : la fonction Vertical (Droite) n'est pas disponible si vous sélectionnez Plusieurs courbes par colonne de groupe comme type de courbe.
Groupe	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de courbes tracées sur le graphique. Une courbe distincte est en effet créée pour chaque valeur unique de cette colonne. Toutes les courbes utilisent le même ensemble d'axes. Remarque : la fonction Vertical (Droite) n'est pas disponible si vous sélectionnez Plusieurs courbes par colonne de groupe comme type de courbe.
Grouper les graphiques par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de graphes créés. Un graphique distinct est créé pour chaque valeur unique. Activez la case à cocher Produire la même mise à l'échelle pour tous les graphiques si vous voulez que l'échelle de

Nom du rôle	Description
	chaque axe soit la même pour tous les graphiques créés. Par défaut, la plage de valeurs de chaque axe est basée sur les valeurs minimales et maximales des données représentées, et peuvent donc varier d'un graphique à l'autre. Si vous activez cette option, la plage de chaque axe est la même pour tous les graphiques.

Courbe : Définir les options d'apparence

Définir d'autres options de tracé

Dans la liste, sous l'en-tête Apparence, cliquez sur **Graphes** pour accéder à ces options.

Remarque : ces options ne sont disponibles que si vous sélectionnez une courbe ou un nuage de points en 2D.

Pour contrôler l'apparence d'un graphe pour une variable spécifique :

- 1 Sélectionnez une variable dans la liste. Les variables répertoriées dans la liste sont celles que vous avez attribuées aux fonctions **Vertical** et **Vertical (Droite)**. Si vous avez sélectionné le type de courbe **Plusieurs courbes par colonne de groupe**, cette liste contient les valeurs uniques de la variable que vous avez attribuée à la fonction **Groupe**.
- 2 Dans la zone **Trait**, indiquez le style, l'épaisseur et la couleur du trait.
- 3 Sélectionnez une couleur de contour. Cette option spécifie la couleur pour ces éléments :
 - Contour des zones délimitées
 - Intervalle de confiance d'une analyse de régression
 - Traits ou contours des courbes avec écart-type
- 4 Dans la zone **Symbole pour un point de données**, spécifiez le type du symbole, le symbole à utiliser, la hauteur de chaque symbole, ainsi que la couleur des symboles des points de données.

Remarque : certains symboles ne sont pas pris en charge par tous les formats de sortie graphique. Si vous sélectionnez un symbole non pris en charge par le format de sortie, un autre symbole (pris en charge) s'affiche.

- 5 Indiquez si vous souhaitez interrompre les lignes de la courbe aux valeurs manquantes.
-

Remarque : la case à cocher **Interrompre la seconde ligne de la courbe aux valeurs manquantes** n'est disponible que si vous attribuez une variable à la fonction **Vertical (Droite)**.

Options d'apparence supplémentaires

Certaines options sont communes aux tâches graphiques. Pour plus d'informations sur ces options, voir ces rubriques :

- "Sélectionner la méthode d'interpolation" page 738
- "Définir les options d'un axe" page 729
- "Ajouter des lignes de référence" page 732
- "Définir les options de légende" page 733
- "Personnaliser l'aire" page 734

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Assistant Courbe

<i>À propos de l'assistant Courbe</i>	313
<i>Assistant Courbe : Sélectionner les données</i>	313
<i>Assistant Courbe : Attribuer des variables aux fonctions</i>	314
<i>Assistant Courbe : Définir les options de la courbe</i>	314
<i>Assistant Courbe : Définir les options d'apparence</i>	316
<i>Assistant Courbe : Indiquer les titres et les notes de bas de page</i>	316

À propos de l'assistant Courbe

L'Assistant **Courbe** vous aide à créer des courbes bidimensionnelles, avec des options pour les symboles et les interpolations.

Assistant Courbe : Sélectionner les données

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Cliquez sur **Suivant** pour attribuer des variables aux fonctions.

Assistant Courbe : Attribuer des variables aux fonctions

- 1 Avant d'attribuer des variables aux fonctions, vous devez sélectionner les données.
- 2 Attribuez des variables aux fonctions suivantes :
 - **Horizontal** - la colonne que vous attribuez à cette fonction est la variable d'axe horizontal ou X du graphique.
 - **Vertical** la colonne que vous attribuez à cette fonction est la variable d'axe vertical ou Y du graphique.
- 3 (Facultatif) Attribuez une variable à la fonction **Représenter par**. Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de courbes à produire. Une courbe distincte est créée pour chaque valeur unique. Vous pouvez attribuer plusieurs variables à cette fonction.
- 4 (Facultatif) Activez la case à cocher **Agréger pour chaque valeur horizontale distincte** pour créer un récapitulatif statistique des valeurs uniques de la variable d'axe horizontale. Sélectionnez une fonction statistique dans la liste déroulante **Fonction**.

Par exemple, vos données d'entrée répertorient les ventes et les dépenses pour chaque pays et par région. Vous attribuez la variable Country à la fonction **Horizontal** et la variable Sales à la fonction **Vertical**. Si vous voulez représenter la somme des ventes pour chaque pays, activez la case à cocher, puis sélectionnez **Somme** dans la liste déroulante. La tâche ajoute les ventes régionales pour un pays et présente une seule valeur par pays dans le résultat final.
- 5 Cliquez sur **Suivant** pour spécifier les options de la courbe.

Assistant Courbe : Définir les options de la courbe

- 1 Avant de définir options d'une courbe, vous devez attribuer des variables aux fonctions.

- 2 Sélectionnez une interpolation de courbe dans la liste déroulante **Type**. Cliquez sur **Détail** pour spécifier les options de l'interpolation sélectionnée. Ces interpolations sont disponibles :
 - **Lagrange** trace une courbe lissée reliant les points de données. Cette méthode est principalement utilisée pour les données correspondant à des valeurs précises mises en tableau. Un polynomial du degré indiqué (1, 3 ou 5) est ajusté sur les 2, 4 ou 6 points les plus proches. En général, la première valeur dérivée n'est pas continue. Si les valeurs de la variable X ne sont pas strictement croissantes, la méthode paramétrique correspondante est utilisée.
 - **Courbe** relie les points de données par des lignes droites. Les points sont reliés dans l'ordre dans lequel ils apparaissent dans les données d'entrée.
 - La méthode **Bâton** trace un trait vertical à partir de chaque point de données jusqu'à un trait horizontal à la valeur 0 sur l'axe vertical ou à la valeur minimale sur l'axe vertical si celle-ci est supérieure à 0. La ligne horizontale est tracée automatiquement.
 - **Régression** crée un graphique d'analyse de régression. Cette méthode d'interpolation n'est disponible que si vous attribuez des variables numériques aux fonctions Horizontal et Vertical.
 - **Nuage de points** supprime toute interpolation. Une courbe des points de données est créée.
 - **Courbe lissée** ajuste une courbe lissée par rapport aux données à l'aide d'une routine de lissage. Cette méthode permet de lisser les données bruyantes. Les points du tracé ne se trouvent pas systématiquement sur la ligne.
 - **Spline** trace le trait en utilisant une routine spline.
 - **Écart-type** utilise une ligne pleine pour relier la valeur Y moyenne à des écarts-types de ± 1 , 2 ou 3 pour chaque X. Par défaut, deux écarts-types sont utilisés.
 - **Escalier** représente les données sous forme d'un escalier en utilisant la fonction par paliers.
- 3 Indiquez le style, l'épaisseur et la couleur du trait.
- 4 Spécifiez le symbole à utiliser, la hauteur de chaque symbole, ainsi que la couleur des symboles des points de données.

Remarque : certains symboles ne sont pas pris en charge par tous les formats de sortie graphique. Si vous sélectionnez un symbole non pris en charge par le format de sortie, un autre symbole (pris en charge) s'affiche.

- 5 Pour définir les options d'apparence, cliquez sur **Suivant**.

Assistant Courbe : Définir les options d'apparence

- 1 Avant de définir l'apparence d'une courbe, vous devez spécifier les options de la courbe.
- 2 Indiquez si vous souhaitez inclure des lignes de référence et des graduations sur le graphique.
- 3 (Facultatif) Cliquez sur les libellés de l'axe vertical et horizontal.
- 4 Cliquez sur **Suivant** pour indiquer des titres et des notes de bas de page.

Assistant Courbe : Indiquer les titres et les notes de bas de page

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

- 1 Avant d'indiquer les titres et les notes de bas de page, vous devez spécifier les options d'apparence.
- 2 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie. Vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.
- 3 Cliquez sur **Terminer** pour exécuter la tâche.

Modèles linéaires

<i>À propos de la tâche Modèles linéaires</i>	317
<i>Modèles linéaires : Attribuer des variables aux fonctions d'analyse</i>	318
<i>Modèles linéaires : Créer un modèle</i>	319
<i>Modèles linéaires : Définir les options du modèle</i>	319
<i>Modèles linéaires : Définir les options avancées</i>	320
<i>Modèles linéaires : Sélectionner les tests post hoc</i>	322
À propos des tests post-hoc	322
Indiquer les options des tests des moyennes pour les moindres carrés	322
Indiquer les options des tests des moyennes pour les moyennes arithmétiques ..	324
<i>Modèles linéaires : Générer des graphes</i>	326
<i>Modèles linéaires : Définir les options de prédiction</i>	327
<i>Indiquer des titres et des notes de bas de page</i>	328
<i>Afficher les propriétés</i>	328

À propos de la tâche Modèles linéaires

La tâche Modèles linéaires utilise la méthode des moindres carrés pour ajuster les modèles linéaires généraux. La tâche analyse les données dans la structure des modèles linéaires généraux. La tâche traite les modèles en reliant une ou plusieurs variables continues dépendantes à une ou plusieurs variables indépendantes. Les variables indépendantes peuvent être des variables de classification, qui divisent les observations en groupes discrets ou en variables continues.

Vous pouvez utiliser cette tâche pour :

- effectuer une analyse de régression. Par exemple, un test peut porter sur la consommation d'un véhicule à différentes vitesses pour déterminer à laquelle de ces vitesses le véhicule peut parcourir la plus longue distance. Un modèle quadratique est ajusté aux données expérimentales.

- analyser des mesures répétitives multivariées. Par exemple, supposons que deux réponses, Y1 et Y2, aient été mesurées chacune trois fois par sujet (en prétraitement, en post-traitement et lors d'une visite de suivi). Chaque sujet reçoit l'un des trois traitements suivants : A, B ou le contrôle. Dans cette tâche, vous pouvez identifier les différentes réponses ainsi qu'un autre facteur répétitif pour distinguer les différents instants de mesure. L'analyse des mesures répétitives inclut les tests multivariés pour le temps et les effets principaux du traitement, ainsi que pour leurs interactions, à travers les réponses.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GLM
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Modèles linéaires : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables dépendantes	Indique les variables qui doivent être utilisées comme variables dépendantes (de réponse). Vous devez attribuer au moins une variable à cette fonction.
Variables quantitatives	spécifie les variables à utiliser comme variables indépendantes (explicatives) continues (également appelées effets).

Nom du rôle	Description
Variables de classification	indique les variables à utiliser en tant qu'effets indépendants discrets. Les variables que vous attribuez à cette fonction peuvent être numériques ou alphanumériques, mais doivent avoir un nombre limité de valeurs discrètes.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper en fonction d'une variable que vous avez déjà sélectionnée comme variable dépendante ou variable indépendante.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Poids relatif	spécifie une variable à utiliser comme pondération relative pour les valeurs correspondantes de la variable sélectionnée comme variable de réponse. Vous ne pouvez attribuer qu'une seule variable à cette fonction.

Modèles linéaires : Créer un modèle

Vous pouvez créer des modèles dans plusieurs tâches SAS. Pour plus d'informations, voir [Annexe 3, "Créer un modèle" page 749](#).

Modèles linéaires : Définir les options du modèle

Dans la liste, cliquez sur **Options du modèle** pour accéder à ces options.

Nom d'option	Description
Afficher les tests associés à l'intercept	affiche les résultats à partir de tests d'hypothèse réalisés sur la constante.
Somme des carrés à afficher	affiche les sommes des carrés qui sont associés aux fonctions estimables pour chaque effet. Les fonctions estimables sont des combinaisons linéaires des paramètres du modèle. Les sommes des carrés sont calculées pour les tests d'hypothèse. Les sommes des carrés de Type I et Type III sont les options par défaut. Vous pouvez ajouter ou supprimer un type de somme des carrés en activant la case à cocher correspondante. Remarque : au moins une case à cocher de Somme des carrés à afficher doit être activée.
Afficher les paramètres estimés	affiche les paramètres estimés. Un paramètre est une quantité de population généralement fixe mais inconnue, de sorte qu'elle est prédite à partir des données de l'échantillon selon un modèle. Par défaut, les résultats estimés des paramètres sont affichés. Vous ne pouvez désactiver la case à cocher Afficher les résultats estimés des paramètres que si une variable a été attribuée à la fonction Variables de classification. Activez la case à cocher IC pour résultats estimés des paramètres afin d'afficher les intervalles de confiance pour les résultats estimés des paramètres. Un intervalle de confiance se compose de bornes inférieure et supérieure qui constituent un intervalle de valeurs raisonnables pour le paramètre estimé avec un niveau de confiance prédéfini. Vous pouvez soit saisir une valeur pour un niveau de confiance soit en sélectionner une dans la liste déroulante. La valeur par défaut est 95 pour cent.

Modèles linéaires : Définir les options avancées

Dans la liste, cliquez sur **Options avancées** pour accéder à ces options.

Nom d'option	Description
Fonctions estimables pour chaque effet du modèle	■ Activez la case à cocher Afficher forme générale de toutes les fonctions estimables pour afficher les coefficients de la matrice des fonctions estimables de Type I - Type IV pour chaque effet du modèle. ■ Activez la case à cocher face au type pour afficher les coefficients de la matrice pour une fonction estimable spécifique. ■ Activez la case à cocher Afficher les fonctions estimables en tant que structure de repliement pour afficher les fonctions estimables en tant que structure de repliement. Dans une structure de repliement, chaque ligne affiche la combinaison linéaire des paramètres qui est estimée pour chaque fonction estimable, et une colonne contenant les mêmes informations est ajoutée dans la table des résultats estimés des paramètres. Remarque : vous pouvez aussi afficher les fonctions estimables pour l'intercept en activant la case Afficher les tests associés à l'intercept.
Matrices à afficher	■ Activez la case à cocher X'X pour afficher la matrice X'X augmentée des produits croisés. ■ Activez la case à cocher X'X inverse pour afficher la matrice X'X inverse (ou inverse généralisée).
Paramètres de mise au point	Remarque : ces paramètres de mise au point ne sont utiles que lorsque des problèmes sont mal conditionnés. ■ Pour changer la sensibilité dans les fonctions de Entrée III ou IV, saisissez la nouvelle valeur dans la zone Sensibilité pour les fonctions de Type III ou IV. Tout élément contenu dans la base de la fonction estimable ayant une valeur absolue inférieure à la valeur spécifiée est défini à 0. La valeur par défaut est $1\text{E}-08$. ■ Pour changer la sensibilité utilisée lors de l'analyse d'une matrice, saisissez la valeur voulue dans la zone Valeur du test de la singularité. Cette option ajuste la sensibilité de la routine de régression aux dépendances linéaires. Si un élément pivot diagonal est inférieur à la somme des carrés corrigés pour la variable multipliée par la valeur que vous avez spécifiée, la colonne associée est déclarée linéairement dépendante des colonnes précédentes, et le paramètre associé est défini à 0. La valeur par défaut est $1\text{E}-07$. ■ Pour afficher les tolérances qui sont utilisées lors de la mise au point de la sensibilité de la routine de régression basée sur la valeur spécifiée dans la zone Valeur du test

Nom d'option	Description
	de la singularité, activez la case à cocher Afficher les tolérances .

Modèles linéaires : Sélectionner les tests post hoc

À propos des tests post-hoc

Sous l'en-tête Tests post hoc de la liste, cliquez sur **Moindres carrés** ou bien sur **Arithmétique** pour accéder à ces options.

- 1 Cliquez sur **Ajouter** pour ajouter un effet dans la zone **Effets à estimer**.
- 2 Pour chaque effet de la liste, indiquez des options dans la zone **Options pour les tests de la moyenne**.

Remarque : les options de la zone **Options pour les tests de la moyenne** ne sont pas les mêmes pour les moyennes des moindres carrés et les moyennes arithmétiques.

Indiquer les options des tests des moyennes pour les moindres carrés

Les options suivantes sont disponibles dans la zone **Options pour les tests de la moyenne**.

Nom d'option	Description
Effets de classe à utiliser	Sélectionnez les effets de classe à utiliser dans l'analyse. Par défaut, aucun effet de classe n'est utilisé.
Comparaisons	Indiquez les options de comparaison à utiliser dans l'analyse.

Nom d'option	Description
	■ Afficher les p-values des différences calcule les p-values des différences des moyennes des moindres carrés. Les options suivantes sont disponibles : ☐ Néant ne calcule pas les p-values. ☐ Par défaut calcule les différences appariées sans ajustement. ☐ Toutes les différences appariées calcule toutes les différences appariées. ☐ Contrôle utilisant le premier niveau calcule les différences avec un contrôle qui est par défaut le premier niveau de chacun des effets spécifiés de la moyenne des moindres carrés. ☐ Test pour inférieur au contrôle test déterminant si les autres niveaux sont inférieurs à celui de contrôle. Vous déclarez un niveau comme étant sensiblement inférieur au contrôle si la borne supérieure de l'intervalle de confiance associée de ce niveau moins le contrôle est inférieure à zéro, et si vous ignorez les bornes inférieures de l'intervalle de confiance associées (définies à moins l'infini). ☐ Test pour supérieur au contrôle test déterminant si les autres niveaux sont supérieurs à celui de contrôle. Vous déclarez un niveau comme étant sensiblement supérieur au contrôle si la borne inférieure de l'intervalle de confiance associée de ce niveau moins le contrôle est supérieure à zéro, et si vous ignorez les bornes supérieures de l'intervalle de confiance associées (définies à l'infini). ■ Méthode d'ajustement pour la comparaison demande un ajustement de comparaison multiple pour les p-values et l'intervalle de confiance, pour les différences des moyennes des moindres carrés.
Intervalle de confiance	Sélectionnez Vrai pour l'option Afficher l'intervalle de confiance afin d'afficher l'intervalle de confiance pour les moyennes des moindres carrés. Si vous spécifiez une valeur pour l'option Afficher les p-values des différences, l'intervalle de confiance pour les différences entre les moyennes sont également affichées.

Indiquer les options des tests des moyennes pour les moyennes arithmétiques

Les options suivantes sont disponibles dans la zone **Options pour les tests de la moyenne**.

Nom d'option	Description
Effets de classe à utiliser	Sélectionnez les effets de classe à utiliser dans l'analyse. Par défaut, aucun effet de classe n'est utilisé.
Comparaisons	Sélectionnez la méthode de comparaison à utiliser dans l'analyse : ■ Défaut indique qu'aucune méthode de comparaison n'est réalisée. ■ Tests t de Bonferroni réalise des tests t de Bonferroni des différences entre les moyennes pour toutes les moyennes des effets principaux. ■ Tests multivariés de Duncan réalise un test multivarié de Duncan sur toutes les moyennes des effets spéciaux ■ Tests de comparaison multiples de Gabriel réalise une procédure de comparaison multiple de Gabriel sur toutes les moyennes des effets spéciaux ■ Tests t appariés réalise des tests t appariés, équivalents du test de la différence la moins significative de Fisher dans le cas de cellules de taille égale, pour toutes les moyennes des effets principaux ■ Test de Ryan-Einot-Gabriel-Welsch réalise le test multivarié de Ryan-Einot-Gabriel-Welsch sur toutes les moyennes des effets principaux ■ Procédure de comparaison multiple de Scheffe réalise la procédure de comparaison multiple de Scheffe sur toutes les moyennes des effets principaux ■ Test t apparié et corrigé de Sidak réalise des tests t appariés sur les différences entre les moyennes avec des niveaux ajustés selon l'inégalité de Sidak pour toutes les moyennes des effets principaux ■ Tests multivariés de Student-Newman-Keuls réalise le test multivarié de Student-Newman-Keuls sur toutes les moyennes des effets principaux ■ Test de l'étendue "studentisée" de Tukey réalise le test de l'étendue "studentisée" de Tukey sur toutes les moyennes des effets principaux

Nom d'option	Description
	■ Test t du rapport K de Waller-Duncan réalise le test <i>t</i> du rapport <i>K</i> de Waller-Duncan sur toutes les moyennes des effets principaux.
Carré moyen de l'erreur	Sélectionnez l'effet de l'erreur à utiliser dans les comparaisons multiples. Vous pouvez également spécifier le type du carré moyen pour l'effet de l'erreur.
Options des moyennes	Spécifiez ces options pour les moyennes : ■ Afficher moyennes de indique si les moyennes doivent être affichées pour toutes les variables du modèle ou seulement pour les variables dépendantes. Par défaut, la tâche affiche les moyennes pour toutes les variables continues. ■ Joindre sous-groupes non significatifs répertorie les moyennes ayant été calculées avec la méthode de comparaison par ordre décroissant et indique les sous-groupes non significatifs par segment de ligne, face aux moyennes correspondantes. ■ Trier moyennes par ordre décroissant trie les moyennes par ordre décroissant. Si vous sélectionnez Vrai à la fois pour l'option Trier moyennes par ordre décroissant et pour l'une des options de la zone Intervalles de confiance, les moyennes ne sont pas triées.
Intervalles de confiance	Indiquez si les intervalles de confiance doivent être inclus dans votre analyse. Ces options sont disponibles : ■ Afficher pour les moyennes présente les résultats de la méthode de comparaison comme des intervalles de la moyenne pour chaque niveau de variables spécifié. ■ Afficher toutes les différences appariées présente les résultats de la méthode de comparaison comme des intervalles de confiance pour toutes les différences appariées entre les moyennes.
Homogénéité de variance	Indiquez si un test d'homogénéité de variance doit être demandé pour les groupes définis par l'effet des moyennes. Ces options sont disponibles : ■ Néant indique qu'aucun test n'est réalisé. Il s'agit de la présentation par défaut. ■ Bartlett spécifie un test de Bartlett, variante du test de rapport des vraisemblances de la théorie normale. ■ Brown et Forsythe spécifie une variante de Brown et Forsythe du test de Levene. ■ Levene (carré) utilise les résidus carrés du test de Levene. ■ Levene (absolu) utilise les résidus absolus du test de Levene.

Nom d'option	Description
	■ O'Brien spécifie le test de O'Brien, qui est en fait une variante du test de Levene utilisant les résidus carrés.

Modèles linéaires : Générer des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Par défaut, la sortie comprend tous les graphes appropriés pour la sélection de données actuelle. Vous pouvez néanmoins choisir les graphes à inclure dans la sortie en sélectionnant l'option **Liste personnalisée de graphes**. Vous avez le choix entre les options suivantes :

Nom d'option	Description
Courbes des diagnostics	crée un panel de diagnostics récapitulatifs pour l'ajustement à afficher. Le panel comprend des nuages de points des résidus, résidus absolus, résidus "studentisés" et réponses observées par valeurs prédites, des résidus "studentisés" par effet de levier et du D de Cook par observation. Il comprend aussi un graphique Q-Q des résidus, un histogramme des résidus et un graphique de variabilité Ajustement-Résidu.
Courbes résiduelles	crée des nuages de points des résidus par rapport à chaque covariable continue à afficher.
Boîte à moustaches	crée un graphe pour l'effet du modèle dans une analyse à une dimension du modèle de la variance. Pour créer ce graphe, votre modèle ne peut contenir qu'une seule variable de classification. Si vous n'attribuez pas de variable de classification, la boîte à moustaches n'est pas créée. Si vous avez attribué une variable de poids relatif, ces poids sont ignorés lors de la création de la boîte à moustaches.
Graphe interactif	crée un graphe pour une analyse à deux facteurs du modèle de la variance. Pour pouvoir créer ce graphe, il faut que votre modèle ne contienne que deux variables de classification. Si vous n'attribuez pas deux variables de classification, le graphe des interactions n'est pas créé.

Modèles linéaires : Définir les options de prédiction

Dans la liste, cliquez sur **Prédictions** pour accéder à ces options.

Nom d'option	Description
Données à prédire	identifie les sources de données à utiliser pour prédire les données de sortie. Activez la case à cocher Données supplémentaires pour inclure ces données. Pour désigner l'emplacement des données supplémentaires, cliquez sur Parcourir.
Enregistrer les données de sortie	spécifie s'il faut enregistrer les prévisions et les statistiques de diagnostic dans une table de sortie : Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Statistiques supplémentaires	spécifie s'il faut enregistrer des statistiques supplémentaires. ■ Résidus enregistre le résidu de probabilité utilisé pour identifier les observations peu ajustées, le résidu de Pearson (Chi) normalisé utilisé pour identifier les observations peu représentées pour le modèle, et le résidu d'écart normalisé utilisé pour identifier les observations dont l'ajustement est insatisfaisant. ■ Limites de prédiction enregistre l'intervalle de confiance supérieure et inférieure pour la valeur prédite.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Régression linéaire

<i>À propos de la tâche Régression linéaire</i>	331
<i>Régression linéaire : Attribuer des variables aux fonctions d'analyse</i>	332
<i>Régression linéaire : Définir les options du modèle</i>	333
<i>Régression linéaire : Définir les options statistiques</i>	336
<i>Régression linéaire : Générer des graphes</i>	337
<i>Régression linéaire : Définir les options de prédiction</i>	339
<i>Indiquer des titres et des notes de bas de page</i>	340
<i>Afficher les propriétés</i>	341

À propos de la tâche Régression linéaire

L'analyse de régression linéaire tente d'attribuer une fonction linéaire à vos données en utilisant la méthode des moindres carrés. En utilisant cette tâche, vous pouvez effectuer une analyse de la régression linéaire sur plusieurs variables dépendantes et indépendantes.

Par exemple, une analyse de régression permet de déterminer avec quelle précision il est possible de prédire le poids d'un enfant à partir de sa taille. Supposons qu'une table SAS contienne le poids et la taille de 19 enfants. En utilisant le poids comme variable dépendante et la taille comme variable indépendante, vous pouvez réaliser une analyse de régression linéaire sur ces données.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	REG
Produits SAS requis	SAS Base, SAS/STAT

Nom du prérequis	Noms des procédures et des produits
Autres produits SAS recommandés	SAS/GRAPH

Régression linéaire : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables dépendantes	indique les variables qui doivent être utilisées comme variables dépendantes (de réponse). Vous devez attribuer au moins une variable à cette fonction.
Variables explicatives	indique les variables à utiliser en tant que variables de régression (explicatives). Si vous ne spécifiez aucune variable de régression, chaque variable dépendante est ajustée à un modèle à moyenne uniquement. c'est-à-dire un modèle comportant seulement un terme constant.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : vous ne pouvez pas regrouper l'analyse avec une variable que vous avez sélectionnée en tant que variable de réponse.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Poids relatif	indique la variable contenant les valeurs qui sont des pondérations relatives pour un ajustement pondéré des moindres carrés. Les

Nom du rôle	Description
	valeurs de la variable de poids relatif ne doivent pas être négatives. Vous ne pouvez attribuer qu'une seule variable à cette fonction.

Régression linéaire : Définir les options du modèle

Dans la liste, cliquez sur **Modèle** pour accéder à ces options.

Nom d'option	Description
Méthode de sélection de modèle	Sélectionnez une méthode dans la liste déroulante des méthodes de sélection des modèles. Si vous ne choisissez pas de méthode, le modèle entier est utilisé. En fonction de la méthode de sélection du modèle, des options supplémentaires deviennent disponibles. Modèle complet ajusté (aucune sélection) Il s'agit de la méthode par défaut qui ne propose aucune option de sélection de modèle. Le modèle est celui créé lors de l'attribution des variables dépendantes et quantitatives (explicatives). Sélection ascendante La méthode de sélection ascendante commence sans variable dans le modèle. Pour chacune des variables explicatives, cette méthode calcule les statistiques F reflétant la contribution de la variable dans le modèle. Les p-values de ces statistiques F sont comparées au niveau de significativité spécifié pour inclure une variable dans le modèle. La valeur par défaut est 0,5. Pour changer le niveau de significativité, saisissez une valeur dans le champ Pour entrer dans le modèle. Si aucune statistique F n'a un niveau de significativité supérieur à cette valeur, la sélection ascendante s'arrête. Sinon, la méthode de sélection ascendante ajoute la variable dont la statistique F est la plus élevée dans le modèle. La méthode de sélection ascendante recalcule ensuite les statistiques F pour les variables restant en dehors du modèle, et le processus d'évaluation est répété. Ainsi, les variables sont ajoutées une à une au modèle jusqu'à ce qu'aucune variable restante ne produise une statistique F significative. Une fois qu'une variable a été ajoutée au modèle, elle y reste. Élimination descendante La méthode de sélection descendante commence par calculer les statistiques F pour un modèle, en incluant toutes les variables explicatives. Les variables sont ensuite supprimées du modèle une à une jusqu'à ce que toutes les variables restant dans le modèle produisent des statistiques F significatives. Le niveau de significativité est spécifié dans la zone de texte Pour rester dans le

Nom d'option	Description
	modèle. La valeur par défaut est 0, 10. À chaque étape, la variable montrant la plus faible contribution au modèle est supprimée. Sélection pas à pas La méthode pas à pas est une modification de la méthode de sélection ascendante. Dans la méthode pas à pas, les variables qui sont déjà dans le modèle n'y restent pas nécessairement. À l'instar de la méthode de sélection ascendante, les variables sont ajoutées une à une dans le modèle, et la statistique F d'une variable à ajouter doit être significative. Le niveau de significativité est spécifié dans la zone de texte Pour entrer dans le modèle. Néanmoins, après l'ajout d'une variable, la méthode pas à pas vérifie toutes les variables qui sont déjà incluses dans le modèle et supprime celles qui ne produisent pas une statistique F significative. Le niveau de significativité est spécifié dans la zone de texte Pour rester dans le modèle. Ce n'est qu'après cette vérification et après les suppressions requises qu'une autre variable peut être ajoutée au modèle. Le processus pas à pas prend fin lorsque l'une de ces conditions est remplie : ■ aucune variable en dehors du modèle n'a une statistique F significative, et chaque variable du modèle est significative au niveau de significativité qui est spécifié pour rester dans le modèle ■ la variable à ajouter dans le modèle est la variable qui vient d'en être supprimée
Méthode de sélection de modèle (suite)	Gain maximal sur le R-carré La méthode du Gain maximal sur le R-carré ne s'arrête pas à un seul modèle. Elle essaie de trouver le meilleur modèle à une variable, le meilleur modèle à deux variables, etc. Il n'est pas garanti de trouver le modèle avec le plus grand R^2 pour chaque taille. Cette méthode commence par rechercher le modèle à une variable qui produit le R^2 le plus grand. Puis une autre variable, celle qui réalise la plus grande augmentation du R^2, est ajoutée. Une fois que le modèle à deux variables est obtenu, chaque variable du modèle est comparée à chaque variable ne se trouvant pas dans le modèle. Pour chaque comparaison, cette méthode détermine si la suppression d'une variable et son remplacement par une autre variable augmente le R^2. Une fois qu'elle a comparé toutes les permutations possibles, la méthode réalise la permutation produisant la plus grande augmentation du R^2. Les comparaisons reprennent et le processus se poursuit jusqu'à ce qu'aucune autre permutation ne puisse augmenter le R^2. Ainsi, le modèle à deux variables résultant est considéré comme le meilleur modèle à deux variables que la méthode puisse trouver. Une autre variable Ajouter à est ensuite ajoutée au modèle et le processus de comparaison-permutation est réitéré de façon à trouver le meilleur modèle à trois variables, etc. La différence entre la méthode de sélection pas à pas et celle du R^2 maximal réside dans le fait que dans la méthode du R^2 maximal, toutes les permutations sont évaluées avant que l'une d'elles soit réalisée. Dans la méthode de sélection pas à pas, la "plus mauvaise" variable peut être supprimée sans tenir compte de ce que l'ajout de la "meilleure" variable restante pourrait apporter.

Nom d'option	Description
	Gain minimal sur le R-carré La méthode du Gain minimal sur le R-carré ressemble beaucoup à la méthode du Gain maximal sur le R-carré, mais la permutation choisie est celle qui produit la plus faible augmentation du R^2. Pour un nombre donné de variables du modèle, les méthodes du R-carré maximal et du R-carré minimal produisent généralement le meilleur modèle, mais la méthode du R-carré minimal considère un plus grand nombre de modèles pour chaque taille. La méthode de sélection du R-carré recherche les sous-ensembles de variables explicatives qui prédisent le mieux une variable dépendante par régression linéaire dans l'échantillon donné. Vous pouvez indiquer le nombre de variables explicatives le plus grand et le plus petit devant apparaître dans un sous-ensemble, et sélectionner le nombre de sous-ensembles de chaque taille. La méthode du R-carré peut réaliser efficacement toutes les régressions possibles sur les sous-ensembles et afficher les modèles par ordre décroissant d'importance du R^2 pour chaque taille de sous-ensemble. D'autres statistiques sont disponibles pour comparer les sous-ensembles de différentes tailles. Ces statistiques, de même que les coefficients de régression estimés, peuvent être affichées ou envoyées dans une table SAS. La méthode de sélection du R-carré diffère des autres méthodes de sélection dans la mesure où la méthode de sélection du R-carré identifie toujours le modèle avec R^2 le plus élevé pour chaque nombre de variables considéré. Sélection avec comparaison des valeurs du R-carré ajusté Cette méthode est similaire à la méthode de sélection R-carré, sauf que la statistique de R^2 ajusté est utilisée comme critère pour la sélection des modèles, et la méthode recherche les modèles ayant le R^2 ajusté le plus élevé dans l'étendue de tailles. Sélection de Cp de Mallows Cette méthode est similaire à la méthode du R-carré ajusté, excepté que la statistique Cp de Mallows est utilisée comme critère de sélection du modèle. Les modèles sont présentés par ordre de C_p croissant.
Niveaux de significativité	Pour la méthode de sélection ascendante, indiquez un niveau de significativité pour saisir une variable explicative (indépendante) dans le modèle. Pour la méthode de sélection descendante, indiquez un niveau de significativité pour qu'une variable explicative reste dans le modèle. Pour la méthode de sélection pas à pas, indiquez les deux niveaux de significativité.
Critères statistiques de lissage	Pour la sélection du R carré, la sélection avec comparaison des valeurs du R-carré ajusté ou la sélection du Cp de Mallow, utilisez les cases à cocher pour sélectionner les critères statistiques de lissage à inclure dans les résultats.
Effets à forcer dans le modèle	Pour toutes les méthodes hormis pour la méthode d'ajustement complet du modèle, activez les cases à cocher des effets à imposer au modèle. Utilisez les flèches pour changer l'ordre des effets. Pour supprimer un effet, désactivez sa case à cocher.

Nom d'option	Description
Avec intercept	Pour toutes les méthodes, désactivez cette case à cocher si vous voulez supprimer les effets de l'intercept dans l'analyse. Par défaut, l'intercept est inclus.

Régression linéaire : Définir les options statistiques

Dans la liste, cliquez sur **Statistiques** pour accéder à ces options.

Nom d'option	Description
Détails sur les estimations	définit les options d'estimation. Vous avez le choix entre ces options : ■ les coefficients de régression standardisés. Un coefficient de régression standardisé est calculé en divisant l'estimation du paramètre par le rapport entre l'écart-type de l'échantillon de la variable dépendante et l'écart-type du régresseur. ■ la somme séquentielle des carrés (Type 1) et la somme partielle des carrés (Type 2) ainsi que les paramètres pour chaque terme du modèle. ■ la matrice de corrélation des estimations. ■ la matrice de covariance des estimations. ■ le niveau de confiance pour les paramètres estimés. La valeur par défaut est 95 pour cent.
Corrélations	indique s'il faut inclure les coefficients de corrélation partielle au carré et semi-partielle au carré. ■ Les coefficients de corrélation partielle sont calculés en tant que $SS/(SS+SSE)$, où SSE correspond à la somme des carrés des erreurs. ■ Les coefficients de corrélation semi-partielle sont calculés en tant que SS/SST, où SST est le SS total corrigé. Si vous choisissez de ne pas inclure l'intercept dans le modèle, le SS total non corrigé est utilisé dans le dénominateur. Chaque type de corrélation utilise les sommes des carrés de Type 1 et Type 2 pour calculer les coefficients de corrélation.

Nom d'option	Description
Diagnostics	spécifie les analyses de diagnostic à effectuer. ■ analyse de colinéarité. Cette option demande une analyse détaillée de la colinéarité entre les régresseurs. Cette analyse inclut les valeurs propres, les indices de condition et la décomposition des variances des estimations pour chaque valeur propre. ■ analyse de colinéarité sans intercept. Cette option demande la même analyse que l'option d'analyse de colinéarité, excepté que l'intercept n'est pas compris dans l'analyse. ■ valeurs de tolérance pour les estimations. La tolérance d'une variable est définie par $1-R^2$, où R^2 est obtenu à partir de la régression de la variable sur tous les autres régresseurs du modèle. valeurs d'inflation de la variance. L'inflation de variance est la réciproque de la tolérance. ■ test d'hétéroscédasticité. Cette option teste si le premier et le second moments du modèle sont correctement spécifiés. ■ matrice de covariance asymptotique. Cette option affiche la matrice de covariance asymptotique estimée sous l'hypothèse d'hétéroscédasticité. ■ statistique de Durbin-Watson. La statistique de Durbin-Watson indique si les erreurs ont une autocorrélation de premier ordre. (Ce test est approprié uniquement pour les données de série chronologique). L'autocorrélation des résidus des échantillons est également produite.

Régression linéaire : Générer des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Par défaut, la sortie comprend tous les graphes appropriés pour la sélection de données actuelle. Vous pouvez néanmoins choisir les graphes à inclure dans la sortie en sélectionnant l'option **Liste personnalisée de graphes**. Vous avez le choix entre ces options :

Nom du graphe	Description
Histogramme des résidus	crée un histogramme des résidus d'ajustement.

Nom du graphe	Description
Graphe des résidus sur les valeurs prédites	crée un graphique des résidus comparés aux valeurs prédites. Un résidu correspond à la différence entre la valeur observée et la valeur prédite.
Graphe des résidus 'studentisés' sur les valeurs prédites	crée un graphique des résidus comparés aux valeurs prédites.
Graphe des valeurs observées sur les prédites	crée un graphique des valeurs observées par rapport aux valeurs prédites.
Tracer la statistique D de Cook	crée un graphique de la statistique D de Cook par rapport au nombre d'observations.
Graphe des résidus 'studentisés' sur l'effet de levier	crée un graphique des résidus comparés aux valeurs prédites. Un résidu 'studentisé' représente la valeur du résidu divisée par l'erreur type, avec suppression de l'observation courante. L'effet levier pour une observation correspond à la valeur d'un élément diagonal correspondant dans la matrice H .
Courbe de quantile normal des résidus	crée un tracé QQ.
Graphe Ajustement-Résidu	crée deux graphiques côte à côte des quantiles de l'ajustement centré et des résidus.
Boîte à moustaches des résidus	crée une boîte à moustaches des résidus.
Courbes des diagnostics	crée des graphiques de diagnostics d'ajustement.
Graphiques DFFITS	crée des graphiques de la différence standardisée dans l'ajustement (DFFITS). La statistique DFFITS représente l'influence standard de l'observation sur la valeur prédite. Une valeur élevée indique que l'observation est très influente.

Nom du graphe	Description
Graphes DFBETAS	crée des graphiques de la différence normalisée dans l'estimation des paramètres liée à la suppression de l'observation. Est utile pour évaluer l'effet d'une observation individuelle sur chaque paramètre estimé du modèle ajusté.
Courbes résiduelles	crée des graphiques des résidus par rapport aux régresseurs.
Nuage de points avec courbe de régression	crée une courbe de régression superposée à un nuage de points des données.

Régression linéaire : Définir les options de prédiction

Dans la liste, cliquez sur **Prédictions** pour accéder à ces options.

Nom d'option	Description
Données à prédire	identifie les sources de données à utiliser pour prédire les données de sortie. Activez la case à cocher Données supplémentaires pour inclure ces données.
Statistiques supplémentaires	spécifie les statistiques supplémentaires à inclure. ■ La sélection de Résidus a pour effet d'inclure les résidus ordinaires, standardisés et "studentisés" dans la sortie. Un résidu ordinaire représente la différence entre la valeur observée et la valeur prédite. Un résidu standardisé représente la valeur du résidu divisée par l'erreur type. Un résidu 'studentisé' représente la valeur du résidu divisée par l'erreur type, avec suppression de l'observation courante. ■ La sélection de Limites de prédiction a pour effet d'inclure les bornes supérieure et inférieure d'un intervalle de confiance de $100 \cdot (1 - \alpha)\%$ pour chaque prédiction et la moyenne attendue de la variable dépendante dans la sortie.

Nom d'option	Description
Enregistrer les données de sortie	indique si les prédictions et statistiques de diagnostic doivent être enregistrées de façon permanente dans une table de sortie. Voici les statistiques de diagnostic : ■ La statistique d'influence D de Cook ■ l'influence standardisée de l'observation sur la valeur prédite (statistique DFFITS) ■ le levier, ■ l'erreur type de la valeur prédite individuelle ■ l'erreur type de la valeur prédite de la moyenne ■ l'erreur type du résidu Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur Parcourir.
Afficher la sortie et les graphes	Affiche les résultats. Pour supprimer les résultats, désactivez la case à cocher. Si vous désactivez cette case à cocher, les options du volet Graphes ne sont pas disponibles. Par défaut, les résultats sont affichés. Utilisez la case à cocher Afficher les prédictions pour inclure les prédictions dans les résultats. Cette option n'est disponible que si une source de données est sélectionnée dans la zone Données à prédire et si la case à cocher Afficher la sortie et les graphes est activée.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.

- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.
.....

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

.....
Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.
.....

Liste des données

<i>À propos de la tâche Liste des données</i>	343
<i>Liste des données : Attribuer des variables aux fonctions d'analyse</i>	344
<i>Liste des données : Définir les options de liste</i>	345
<i>Indiquer des titres et des notes de bas de page</i>	347
<i>Afficher les propriétés</i>	347

À propos de la tâche Liste des données

La tâche Liste des données imprime les observations d'une table SAS, en utilisant tout ou partie des variables. Vous pouvez créer différents types de rapport, allant d'une simple liste à des rapports extrêmement personnalisés qui regroupent les données et calculent des totaux et des sous-totaux pour des variables numériques.

Par exemple, vous pouvez utiliser la tâche Liste des données pour créer un rapport qui additionne les dépenses et les recettes de chaque zone de ventes, compare ces valeurs aux dépenses et aux recettes totales de la société, affiche le nombre d'observations dans chaque sous-groupe et dans l'ensemble du rapport, et contient un titre personnalisé indiquant le nom de la zone de ventes.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	PRINT
Produits SAS requis	Base SAS
Autres produits SAS recommandés	Aucun

Liste des données : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables de liste	imprime les variables dans l'ordre où elles sont indiquées. Vous devez attribuer au moins une variable à cette fonction.
Grouper l'analyse par	spécifie comment trier la table. Une liste est générée pour chaque valeur distincte (ou groupe BY) dans la variable ou combinaison de variables.
Page par	imprime une nouvelle page chaque fois que la valeur de la variable spécifiée change ou bien lorsque le groupe BY suivant commence. La variable que vous attribuez à cette fonction doit également être variable de la fonction Grouper l'analyse par . Vous ne pouvez attribuer qu'une seule variable à cette fonction.
Total de	imprime les sommes de la variable sélectionnée au bas du rapport de listes.
Sous-total de	imprime un sous-total chaque fois que la valeur de la variable indiquée change ou bien lorsque le groupe BY suivant commence. La variable que vous attribuez à cette fonction doit également être variable de la fonction Grouper l'analyse par . Vous ne pouvez attribuer qu'une seule variable à cette fonction.
Libellé d'identification	spécifie comment identifier les lignes plutôt que d'utiliser des numéros d'observation.

Liste des données : Définir les options de liste

Dans la liste, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Lignes à afficher	spécifie les lignes à lister. Vous avez le choix entre ces options : ■ Toutes les lignes : imprime toutes les lignes. ■ Les n premières lignes : imprime le nombre de lignes indiqué dans la zone de saisie Quantité (n). Par exemple, si vous définissez <i>n</i> à 50, les données de sortie contiennent les lignes 1 à 50 de la table d'entrée. ■ Les n % de premières lignes imprime le pourcentage de lignes indiqué dans la zone de saisie Quantité (n). Par exemple, si votre table d'entrée contient 100 lignes et si vous définissez <i>n</i> à 25 %, les données de sortie contiennent les lignes 1 à 25 de la table d'entrée. ■ Une ligne sur n imprime chaque <i>n</i>ème ligne de la table. Vous indiquez la valeur de <i>n</i> dans le champ de saisie Quantité (n). Par exemple, si vous définissez <i>n</i> à 10, les données de sortie contiennent chaque 10ème ligne de la table d'entrée.
Imprimer le numéro de ligne	inclut dans la sortie une colonne répertoriant le numéro de ligne de chaque observation. Vous pouvez indiquer un libellé pour cette colonne dans le champ de saisie En-tête de colonne . Le nom par défaut est Numéro de ligne .
Utiliser libellés de variable comme en-têtes de col.	utilise le libellé de la variable plutôt que son nom comme en-tête de colonne.
Imprimer le nombre de lignes	affiche le nombre de lignes de la table à la fin de la sortie ou le nombre de lignes contenues dans chaque groupe BY à la fin de chaque sortie de groupe BY.
Arrondir les valeurs avant d'additionner	arrondit chaque valeur numérique au nombre de décimales de son format ou bien à deux décimales si aucun format n'a été indiqué. Si cette option est sélectionnée, la tâche Lister

Nom d'option	Description
	les données effectue l'opération avant de totaliser la variable.
Diviser la page en sections	divise chaque page en sections s'il y a plus de colonnes que la page peut en contenir. Si vous désactivez l'option Diviser la page en sections, la tâche Lister les données place autant d'observations que possible sur chaque page. S'il y a plus de colonnes qu'il n'en peut sur une page, les colonnes en trop sont placées sur les pages suivantes. En fonction du nombre de lignes et de colonnes contenues dans la table, cette option peut résulter en un nombre réduit de pages. Remarque : cette option n'a d'effet que sur une sortie texte. Elle n'affecte pas les résultats aux formats Rapport SAS, HTML, RTF, ou PDF.
Direction de l'en-tête	imprime les en-têtes de colonne horizontalement ou verticalement. Sélectionnez Par défaut et laissez SAS déterminer la disposition optimale pour chaque colonne.
Largeur de colonne	indique comment la tâche Liste de données détermine la largeur des colonnes : ■ Par défaut : La tâche Liste des données détermine la largeur des colonnes en fonction de chaque page. ■ Plein: La tâche Liste des données utilise la largeur du format (ou la largeur par défaut si aucune largeur n'est spécifiée pour le format) pour toutes les pages. ■ Minimum: La tâche Liste des données utilise la plus petite largeur de colonne possible en fonction de chaque page. ■ Uniforme : Lit entièrement la table pour déterminer la largeur de colonne appropriée avant de générer la sortie. Si cette option n'est pas sélectionnée, la largeur d'une même colonne peut être différente d'une page à l'autre. ■ Uniforme par: La tâche Liste des données imprime uniformément toutes les pages au sein de chaque groupe BY. Remarque : cette option n'a d'effet que sur une sortie texte. Elle n'affecte pas les résultats aux formats Rapport SAS, HTML, RTF, ou PDF.
Fractionner libellés à	indique que si le libellé d'une variable contient un caractère de fractionnement (*, !, @, #, \$, %, ^, &, ou +), il doit être fractionné. Par exemple, pour une variable contenant "Ceci est*un libellé", si le caractère * est sélectionné comme

Nom d'option	Description
	caractère de fractionnement, l'en-tête de colonne sera le suivant :
	Ceci est un libellé
	Il n'est pas utile de sélectionner à la fois Utiliser le libellé de la variable comme en-tête de col , et Fractionner libellés à . Fractionner libellés à indique que vous voulez utiliser les libellés des variables.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Assistant Rapport sous forme de liste

<i>À propos de l'Assistant Rapport sous forme de liste</i>	349
<i>Assistant Rapport sous forme de liste : Sélectionner les données</i>	350
<i>Assistant Rapport sous forme de liste : Définir la liste</i>	350
<i>Assistant Rapport sous forme de liste : Afficher les totaux</i>	351
<i>Assistant Rapport sous forme de liste : Indiquer les titres et les notes de bas de page</i>	352

À propos de l'Assistant Rapport sous forme de liste

L'Assistant **Rapport sous forme de liste** permet de créer des rapports détaillés ou de synthèse. Un rapport détaillé contient une ligne pour chaque observation sélectionnée pour le rapport. Un rapport de synthèse consolide les données de sorte que chaque ligne représente plusieurs observations. Ces deux types de rapport peuvent contenir des totaux qui récapitulent les données numériques d'un ensemble de lignes ou de la totalité des lignes. Vous pouvez également créer des tables distinctes ou des sauts de page en fonction des valeurs uniques d'une variable.

Par exemple, vous pouvez utiliser l'Assistant **Rapport sous forme de liste** pour créer un rapport répertoriant les ventes de soda et de café dans les épiceries de chaque région. Les responsables de chaque région sont groupés dans une table distincte. Dans chaque table, les ventes de chacun de ces produits et le total des ventes sont indiqués pour chaque responsable.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	REPORT
Produits SAS requis	Base SAS
Autres produits SAS recommandés	Aucun

Assistant Rapport sous forme de liste : Sélectionner les données

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Cliquez sur **Suivant** pour définir la liste.

Assistant Rapport sous forme de liste : Définir la liste

Si la source de données contient moins de 20 colonnes, l'Assistant **Rapport sous forme de liste** inclut toutes les colonnes de la source de données. Ces colonnes apparaissent dans la zone **Aperçu**. Si en revanche les données contiennent plus de 20 colonnes, aucune colonne n'est incluse dans le rapport sous forme de liste et la zone **Aperçu** est vide.

Vous devez alors désigner les colonnes à ajouter au rapport. Vous pouvez ajouter toutes les colonnes de la source de données, une colonne spécifique ou la fréquence.

Cliquez sur **Suivant** pour afficher les totaux.

Assistant Rapport sous forme de liste : Afficher les totaux

Vous pouvez afficher les sous-totaux ou les totaux généraux de la totalité des variables ou des variables numériques qui ne sont pas des variables de classification.

Pour spécifier les totaux :

- 1 Définissez tout d'abord la liste.
- 2 Dans le volet **Sélectionner des totaux**, activez les cases à cocher des variables dont vous voulez calculer le total. Pour afficher les totaux de toutes les variables, activez la case à cocher **Sélectionner tout**.
- 3 Pour spécifier le type de total, la position de la valeur totale et le libellé de la ligne contenant le total, cliquez sur **Modifier**. La boîte de dialogue **Modifier les totaux** apparaît.

Vous avez le choix entre ces options :

- **Type de totaux** indique le type de total à afficher dans la table. Vous pouvez choisir d'afficher les totaux généraux et les sous-totaux. Vous devez sélectionner au moins un type de total à afficher. Pour afficher tous les totaux dans la table, activez la case à cocher **Sélectionner tout**.
- **Position** indique l'emplacement auquel la ligne de total doit apparaître dans la table. Vous pouvez afficher le total avant ou après les lignes ou les observations utilisées dans le calcul du total.
- **Libellé des totaux** indique le libellé à utiliser pour les lignes contenant les totaux. Par défaut, ce libellé est Total.
- **Libellé des sous-totaux** indique le libellé à utiliser pour les lignes contenant les sous-totaux. Par défaut, ce libellé est Sous-total. Vous pouvez également préciser si le libellé doit être affiché avant ou après le texte de ces lignes. Par défaut, aucun espace n'est inséré entre le libellé et le texte qui le précède ou le suit. Pour insérer un espace, vous devez l'ajouter avant ou après le libellé dans la zone **Libellé des sous-totaux**.

- 4 Cliquez sur **Suivant** pour indiquer des titres et des notes de bas de page.

Assistant Rapport sous forme de liste :

Indiquer les titres et les notes de bas de page

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

- 1 Avant d'indiquer les titres et les notes de bas de page, vous devez afficher les totaux des variables numériques.
- 2 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie. Vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.
- 3 Cliquez sur **Terminer** pour exécuter la tâche.

Régression logistique

<i>À propos de la tâche Régression logistique</i>	353
<i>Régression logistique : Attribuer des variables aux fonctions d'analyse</i>	354
<i>Régression logistique : Créer un modèle</i>	355
Spécifier la réponse	355
Créer un modèle	357
Spécifier une méthode de sélection du modèle	357
<i>Régression logistique : Définir les options du modèle</i>	359
<i>Régression logistique : Générer des graphes</i>	361
<i>Régression logistique : Définir les options de prédiction</i>	363
<i>Indiquer des titres et des notes de bas de page</i>	364
<i>Afficher les propriétés</i>	365

À propos de la tâche Régression logistique

Des réponses binaires comme succès ou échec, et des réponses ordinales comme normale, léger ou grave, apparaissent dans de nombreux compartiments de l'étude. L'analyse de régression logistique étudie les relations entre ces réponses discrètes et un ensemble de variables explicatives.

Vous pouvez utiliser la régression logistique pour déterminer les effets d'une étude clinique. Par exemple, supposons que vous utilisiez la tâche Régression logistique pour analyser les résultats d'une étude des effets analgésiques d'un traitement sur des patients âgés souffrant de névralgie. Deux traitements tests et un placebo sont comparés. La variable de réponse indique si le patient a ou non signalé une douleur. Les chercheurs enregistrent l'âge et le sexe des patients et la durée de la plainte avant le début du traitement. À partir des résultats de la tâche de régression logistique, vous pouvez déterminer si les effets de diminution de la douleur sont différents pour les deux traitements testés.

Remarque : lorsque vous ouvrez la tâche Régression logistique, vous remarquerez peut-être un léger décalage du fait que SAS Enterprise Guide détermine quelles variables ont des valeurs binaires. Une fois qu'elle est terminée, la tâche fournit des informations sur la variable attribuée à la fonction **Variable dépendante**.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	LOGISTIC
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Régression logistique : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variable dépendante	indique la variable à utiliser en tant que variable dépendante (de réponse). Vous ne devez attribuer qu'une seule variable à cette fonction.
Variables quantitatives	indique quelles variables utiliser comme variables indépendantes (explicatives). Vous pouvez également spécifier les unités de changement pour les variables quantitatives. Par exemple, pour générer des rapport de cotes estimés dans les résultats, tapez -2 dans la zone Unités pour qu'un rapport de cotes représente le changement des cotes lorsque la variable est augmentée de deux unités.
Variables de classification	indique les variables de classification à utiliser dans l'analyse. Ce peut être des variables caractères ou numériques. Si elles sont numériques, les valeurs doivent être discrètes plutôt que continues.

Nom du rôle	Description
	Pour sélectionner la méthode de paramétrage pour chaque variable, sélectionnez la variable de classification. Dans la zone Style de codification pour <i>nom_variable</i> , indiquez si la méthode de paramétrage des variables doit être du type codage des effets (Effets) ou codage des cellules de référence (Référence).
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : Vous ne pouvez pas regrouper l'analyse avec une variable que vous avez sélectionnée en tant que variable de réponse.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Poids relatif	spécifie une variable à utiliser comme pondération relative pour les valeurs correspondantes de la variable sélectionnée comme variable de réponse. Si une valeur de variable est négative ou manquante, elle est exclue de l'analyse. Pour normaliser la pondération de la variable, activez la case à cocher Normaliser.

Régression logistique : Créer un modèle

Spécifier la réponse

Dans la liste, sous l'en-tête Modèle, cliquez sur **Réponse** pour accéder à ces options.

Nom du rôle	Description
Type de réponse	Les types de réponse disponibles varient selon le nombre de niveaux de la variable de réponse.

Nom du rôle	Description
	■ Si la variable de réponse ne comporte que deux niveaux, le type de réponse est Binaire. ■ Si la variable de réponse comporte plus de deux niveaux, vous avez le choix entre les types de réponse Classée et Non classée.
Type de modèle	Vous pouvez choisir parmi ces types de modèles : ■ logit affiche la sortie de la fonction log-odds qui est la sélection par défaut. Le modèle logit binaire est utilisé lorsqu'il existe deux catégories de réponses, tandis que le modèle logit cumulé est utilisé lorsqu'il en existe plus de deux. ■ probit - affiche la sortie de l'inverse de la fonction de répartition de la loi normale. Le modèle probit binaire est utilisé lorsqu'il existe deux catégories de réponses, et le modèle probit cumulé, lorsqu'il en existe plus de deux. ■ log-log complémentaire - affiche la sortie de la fonction logarithmique complémentaire. Le modèle logarithmique complémentaire binaire s'utilise lorsqu'il y a deux catégories de réponse. Le modèle log-log complémentaire cumulé est utilisé lorsqu'il y a plus de deux catégories de réponse. ■ log-log affiche la sortie de la fonction log-log. Le modèle log-log binaire est utilisée lorsqu'il y a deux catégories de réponse. Le modèle log-log cumulé est utilisé lorsqu'il y a plus de deux catégories de réponse. ■ glogit - affiche la sortie de la fonction de logit généralisé. Dans le modèle logit généralisé, chaque catégorie de non-référence est comparée à la catégorie de référence. Remarque : cette option n'est disponible que si vous avez sélectionné le type de réponse Non classée.
Niveaux de réponse	La case Niveaux de réponse répertorie par ordre croissant les niveaux formatés pour la variable de réponse.
Ajuster le modèle au niveau	Vous pouvez spécifier la manière d'ajuster le modèle à un niveau. ■ Si vous avez choisi le type de réponse Binaire, sélectionnez la catégorie d'événement pour le modèle de réponse binaire dans la liste déroulante Ajuster le modèle au niveau. ■ Si vous avez choisi le type de réponse Classée, vous pouvez utiliser les valeurs de la liste déroulante Ajuster le modèle au niveau pour préciser l'ordre des catégories de réponse. Si vous sélectionnez la première valeur triée, les catégories de réponse sont classées telles qu'elles apparaissent dans la zone Niveaux de réponse. Si vous sélectionnez la dernière

Nom du rôle	Description
	valeur triée, les catégories de réponse sont classées dans l'ordre inverse.
Niveau de référence	Cette option n'est disponible que si vous avez sélectionné le type de réponse Non classée . Pour le modèle logit généralisé, vous pouvez sélectionner la catégorie de référence dans la liste déroulante Niveau de référence . Dans ce modèle, chaque logit compare une catégorie de non-référence à la catégorie de référence.

Créer un modèle

Vous pouvez créer des modèles dans plusieurs tâches SAS. Pour plus d'informations, voir [Annexe 3, "Créer un modèle" page 749](#).

Spécifier une méthode de sélection du modèle

Dans la liste, sous l'en-tête Modèle, cliquez sur **Sélection** pour accéder à ces options.

Nom d'option	Description
Méthode de sélection de modèle	Sélectionnez une méthode dans la liste déroulante des méthodes de sélection des modèles. Si vous ne choisissez pas de méthode, le modèle entier est utilisé. En fonction de la méthode de sélection du modèle, des options supplémentaires deviennent disponibles. Voici les méthodes disponibles : ■ Modèle complet ajusté (aucune sélection) est la valeur par défaut et ne fournit aucune sélection d'effet. Le modèle est celui créé lors de l'attribution des variables dépendantes et quantitatives. ■ La sélection ascendante commence sans effets dans le modèle. Pour chacune des variables explicatives, cette méthode calcule les statistiques F reflétant la contribution de la variable dans le modèle. Les p-values de ces statistiques F sont comparées au niveau de significativité spécifié pour inclure une variable dans le modèle. La valeur par défaut est 0,05. Pour changer le niveau de significativité, saisissez une valeur dans le champ Pour entrer dans le modèle.

Nom d'option	Description
	Si aucune statistique F n'a un niveau de significativité supérieur à cette valeur, la sélection ascendante s'arrête. Sinon, la méthode de sélection ascendante ajoute la variable dont la statistique F est la plus élevée dans le modèle. La méthode de sélection ascendante recalcule ensuite les statistiques F pour les variables restant en dehors du modèle, et le processus d'évaluation est répété. Ainsi, les variables sont ajoutées une à une au modèle jusqu'à ce qu'aucune variable restante ne produise une statistique F significative. Une fois qu'une variable a été ajoutée au modèle, elle y reste. ■ La sélection descendante commence par calculer les statistiques F pour un modèle, en incluant toutes les variables explicatives. Les variables sont ensuite supprimées du modèle une à une jusqu'à ce que toutes les variables restant dans le modèle produisent des statistiques F significatives. Le niveau de significativité est spécifié dans la zone de texte Pour rester dans le modèle. La valeur par défaut est 0,05. À chaque étape, la variable montrant la plus faible contribution au modèle est supprimée. ■ La sélection pas à pas est une variante de la méthode ascendante. Dans la méthode pas à pas, les variables qui sont déjà dans le modèle n'y restent pas nécessairement. De même que dans la méthode de sélection ascendante, les variables sont ajoutées une à une au modèle si la statistique F est significative. Le niveau de significativité est spécifié dans la zone de texte Pour entrer dans le modèle. Néanmoins, après l'ajout d'une variable, la méthode pas à pas vérifie toutes les variables qui sont déjà incluses dans le modèle et supprime celles qui ne produisent pas une statistique F significative. Le niveau de significativité est spécifié dans la zone de texte Pour rester dans le modèle. Ce n'est qu'après cette vérification et après les suppressions requises qu'une autre variable peut être ajoutée au modèle. Le processus pas à pas prend fin lorsque l'une de ces conditions est remplie : □ aucune variable en dehors du modèle n'a une statistique F significative, et chaque variable du modèle est significative au niveau de significativité qui est spécifié pour rester dans le modèle. □ la variable à ajouter dans le modèle est la variable qui vient d'en être supprimée. ■ Meilleur sous-ensemble utilise un algorithme pour rechercher un nombre donné de modèles ayant la statistique Khi-2 la plus élevée pour toutes les tailles de modèle possibles, depuis les modèles d'effet 1, 2, 3, etc., jusqu'au modèle unique contenant la totalité des variables quantitatives.

Nom d'option	Description
Niveaux de significativité	■ Pour la méthode de sélection ascendante, indiquez le niveau de significativité à utiliser pour saisir une variable quantitative (explicative) dans le modèle. ■ Pour la méthode de sélection descendante, spécifiez un niveau de significativité de sorte qu'une variable quantitative reste dans le modèle. ■ Pour la méthode de sélection pas à pas, indiquez les deux niveaux de significativité.
Effets à forcer dans le modèle	Sélectionnez l'effet que vous voulez ajouter au modèle. La variable est ajoutée dans la zone Effets à forcer dans le modèle. Vous pouvez réorganiser ces variables en les sélectionnant et en utilisant les boutons fléchés. Remarque : cette option n'est pas disponible si vous sélectionnez Ajustement complet du modèle (aucune sélection) comme méthode de sélection de modèle.

Régression logistique : Définir les options du modèle

Dans la liste, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Détails sur les estimations	spécifie s'il faut afficher la matrice de corrélation des estimations de paramètres, la matrice de covariance des estimations de paramètres ou les deux.
Évaluation de l'ajustement du modèle	spécifie ces tests d'ajustement du modèle : ■ Statistiques d'influence affiche les mesures de diagnostic permettant d'identifier les observations influentes dans le cas d'un modèle de réponse binaire. ■ Test d'adéquation de Hosmer et Lemeshow affiche le test d'adéquation de Hosmer et Lemeshow dans le cas d'un modèle de réponse binaire. ■ Tests de la déviance et d'adéquation de Pearson la statistique du Chi-deux de Pearson, l'écart, le degré de liberté, le rapport de chaque statistique divisé par son degré de liberté et la <i>p</i>-value correspondante.

Nom d'option	Description
	■ R carré généralisé affiche un coefficient généralisé de détermination pour le modèle ajusté. Remarque : les options Statistiques d'influence, Test d'adéquation de Hosmer et Lemeshow, et Tests d'ajustement de l'écart et de Pearson ne sont disponibles que si vous avez attribué une variable binaire à la fonction Variable dépendante.
Afficher la table de classification	classifie les observations de réponse binaire en entrée selon que les probabilités d'événement prédites sont supérieures ou inférieures à une valeur du point de rupture comprise dans l'intervalle (0,1). Une observation est prédite comme un événement si la probabilité d'événement prédite dépasse le point de rupture. La table par défaut affiche la classification d'un intervalle de probabilités, depuis la probabilité estimée la plus faible (arrondie au 0,02 inférieur) jusqu'à la probabilité estimée la plus élevée (arrondie au 0,02 supérieur) par incréments de 0,02. Pour fournir des points de rupture autres que ceux de la liste par défaut, saisissez une ou plusieurs valeurs dans le champ Valeurs de p-values (points de rupture). Chaque valeur doit être comprise entre 0 et 1. Remarque : cette option est disponible si la variable que vous avez attribuée à la fonction Variable dépendante est binaire.
Intervalle de confiance	spécifie les intervalles de confiance suivants pour les paramètres dans la zone Paramètres, ainsi que les rapports de cotes dans la zone Rapports de cotes conditionnels : ■ Wald calcule les intervalles de confiance sur la base des tests individuels de Wald. ■ Vraisemblance du profil calcule les intervalles de confiance sur la base de la vraisemblance du profil. La création de cet intervalle est dérivée de la distribution du Chi-deux asymptotique du test du rapport des vraisemblances généralisé. Vous pouvez également sélectionner une valeur dans la liste déroulante Niveau de conf. pour définir le niveau de significativité des paramètres de régression ou des rapports de cotes.
Méthodes d'ajustement des modèles	spécifie comment ajuster le modèle. Par défaut, la tâche détermine automatiquement la méthode la mieux adaptée au modèle. Dans la liste déroulante Technique d'ajustement, vous pouvez choisir la technique d'optimisation Scoring de Fisher ou Newton-Raphson pour estimer les paramètres de régression. Remarque : si vous avez sélectionné le type de réponse Non classée, la technique d'optimisation Scoring de Fisher n'est pas disponible.

Nom d'option	Description
	Activez la case à cocher Vraisemblance pénalisée de Firth pour effectuer une estimation du maximum de vraisemblance pénalisée de Firth, qui réduit le biais dans les résultats estimés des paramètres. Cette méthode est utile dans les cas de séparabilité, fréquents lorsque l'événement est rare, et constitue une alternative à la réalisation d'une régression logistique exacte. Note Remarque : la case à cocher Vraisemblance pénalisée de Firth n'est disponible que si vous avez attribué une variable binaire à la fonction Variable dépendante.

Régression logistique : Générer des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Par défaut, la sortie comprend tous les graphes appropriés pour la sélection de données actuelle. Vous pouvez néanmoins choisir les graphes à inclure dans la sortie en sélectionnant l'option **Liste personnalisée de graphes**. Vous avez le choix entre ces options :

Nom du graphe	Description
Graphes du rapport de cotes	génère des graphes de l'intervalle de confiance de vraisemblance du profil et de l'intervalle de confiance de Wald pour les rapports de cotes. Cette option n'est disponible que si vous avez sélectionné une option Rapports de cotes conditionnels dans le volet Modèle > Options.
Courbes ROC	génère un graphe de la (Receiving Operating Characteristic). La zone située sous la courbe ROC est déterminée par la statistique c de la table "Association des probabilités prédites et des réponses observées".
Courbes d'influence	affiche les graphes suivants : ■ graphiques d'index du résidu de Pearson (Khi 2) permettant d'identifier les observations peu représentées par le modèle ; ■ résidu d'écart permettant d'identifier les observations dont l'ajustement est insatisfaisant ; ■ niveau moyen ;

Nom du graphe	Description
	■ diagnostic de décalage de l'intervalle de confiance qui mesure l'influence de chaque observation sur les estimations de régression ; ■ diagnostic de décalage de l'intervalle de confiance qui mesure l'ensemble des modifications des estimations de régression globale imputables à la suppression d'une observation donnée ; ■ modification de la statistique critère qualificatif de lissage du Khi-2 imputable à la suppression d'une observation donnée ; ■ modification de l'écart imputable à la suppression d'une observation donnée ;
Graphes DFBETAS	affiche des graphiques de la différence normalisée dans l'estimation des paramètres liée à la suppression de l'observation. Est utile pour évaluer l'effet d'une observation individuelle sur chaque paramètre estimé du modèle ajusté.
Phat plots	affiche ces courbes : ■ modification de la statistique critère qualificatif de lissage du Khi-2 imputable à la suppression d'une observation donnée ; ■ modification de l'écart imputable à la suppression d'une observation donnée ; ■ décalage de l'intervalle de confiance ; ■ niveau moyen comparé à la probabilité prédite de l'événement.
Graphes d'effet de levier	affiche ces courbes : ■ modification de la statistique critère qualificatif de lissage du Khi-2 imputable à la suppression d'une observation donnée ; ■ modification de l'écart imputable à la suppression d'une observation donnée ; ■ décalage de l'intervalle de confiance ; ■ probabilité prédite comparée au niveau moyen.
Graphiques DCP	affiche ces courbes : ■ modification de la statistique critère qualificatif de lissage du Khi-2 imputable à la suppression d'une observation donnée comparée à la probabilité prédite de l'événement ; ■ modification de l'écart imputable à la suppression d'une observation donnée comparée à la probabilité prédite de l'événement. Dans ces graphes, la couleur des symboles de données dépend de la valeur du décalage de l'intervalle de confiance.

Nom du graphe	Description
Graphiques d'effet	affiche des graphiques des probabilités prédites pour une valeur de la variable indépendante par rapport aux effets du modèle.

Remarque : les options Courbes ROC, Graphiques d'influence, Graphes DFBETAS, Graphes d'effet de levier, Phat plots et Graphiques DCP ne sont disponibles que si vous avez attribué une variable binaire à la fonction Variable dépendante.

Régression logistique : Définir les options de prédiction

Dans la liste, cliquez sur **Prédictions** pour accéder à ces options.

Nom d'option	Description
Données à prédire	identifie les sources de données à utiliser pour prédire les données de sortie. Activez la case à cocher Données supplémentaires pour inclure ces données.
Enregistrer les données de sortie	indique si les prédictions et statistiques de diagnostic doivent être enregistrées de façon permanente dans une table de sortie. ■ Prédictions enregistre la probabilité prédite pour chaque niveau de réponse. Pour une variable de réponse Y à trois niveaux, 1, 2 et 3, les probabilités individuelles sont $\Pr(Y=1)$, $\Pr(Y=2)$, et $\Pr(Y=3)$. ■ Statistiques de diagnostic enregistre les statistiques de diagnostic DFBETAS, les diagnostics de décalage de l'intervalle de confiance (C et CBAR), la matrice H, et les estimations du prédicteur linéaire avec les estimations de l'erreur type. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.

Nom d'option	Description
Statistiques supplémentaires	■ Sélectionnez Résidus pour enregistrer ces statistiques : ☐ résidu d'écart ☐ résidu de Pearson (Khi-2) ☐ modification de l'écart imputable à la suppression d'une observation donnée ; ☐ modification de la statistique critère qualitatif de lissage du Khi-2 imputable à la suppression d'une observation donnée ; ■ Activez Limites de prédiction pour inclure bornes supérieure et inférieure d'un intervalle de confiance pour chaque prédiction et la moyenne attendue de la variable dépendante dans la sortie.
Afficher la sortie et les graphes	affiche les résultats. Pour supprimer les résultats, désactivez la case à cocher. Si vous désactivez cette case à cocher, les options du volet Graphes ne sont pas disponibles. Par défaut, les résultats sont affichés. Utilisez la case à cocher Afficher les prédictions pour inclure les prédictions dans les résultats. Cette option n'est disponible que si une source de données est sélectionnée dans la zone Données à prédire et si la case à cocher Afficher la sortie et les graphes est activée.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Carte

<i>À propos de la tâche Carte</i>	367
<i>Carte : Sélectionner le type de carte</i>	368
<i>Carte : Attribuer des variables aux fonctions d'analyse</i>	369
<i>Carte : Définir les options d'apparence</i>	370
Définir les options de représentation cartographique	370
Définir la présentation	371
Personnaliser les couleurs	372
Options d'apparence supplémentaires	372
<i>Indiquer des titres et des notes de bas de page</i>	372
<i>Carte : Afficher un aperçu des résultats</i>	373
<i>Afficher les propriétés</i>	373

À propos de la tâche Carte

La tâche Carte crée une représentation en couleurs à deux dimensions (choropleth) ou à trois dimensions (bloc et prisme), qui présente la variation de la valeur d'une variable de réponse pour différents secteurs géographiques tels que des régions, des départements ou des pays.

L'exemple suivant montre les différences entre chaque type de carte. Sur une carte des États-Unis, vous voulez afficher les sites de déchets dangereux dans chaque état.

- Sur une carte choroplète, le nombre de sites dans chaque état (la variable de réponse) est représenté par une différente teinte de rouge. Les états avec le plus grand nombre de sites de déchets dangereux apparaissent en rouge foncé. Les états qui n'en ont pas apparaissent en blanc.
- Sur une carte de blocs, le nombre de sites dans chaque état est représenté par la hauteur et la couleur du bloc qui apparaît au centre de chaque zone de la carte.

- Sur une carte avec prismes, le nombre de sites dans chaque état est représenté par la hauteur et la couleur des polyèdres (polygones en relief) pour chaque zone de la carte.

Pour plus d'informations, voir ["Sélectionner un format de sortie pour les graphiques" page 6](#).

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GMAP
Produits SAS requis	SAS Base, SAS/GRAPH
Autres produits SAS recommandés	Néant

Carte : Sélectionner le type de carte

Dans la liste, cliquez sur **Carte** pour accéder à ces options.

Nom du rôle	Description
Carte 2D	Crée une carte 2D dans laquelle les valeurs des variables de réponse sont représentées par différentes couleurs. Pour visualiser rapidement votre table de cartographie, activez la case Créer une carte vide . Quand vous générez une carte à blanc, vous ne spécifiez pas de table de réponse, vous obtenez donc une carte 2D simple.
Carte 3D	Crée une carte de blocs en 3D dans laquelle les niveaux de grandeur des variables de réponse spécifiées sont représentées par des blocs de hauteur et de couleur différentes.
Carte 3D avec prismes	Crée des cartes en 3D avec prismes, dans lesquelles les niveaux de grandeur des variables de réponse spécifiées sont représentés par des polyèdres (polygones en relief) de hauteur et de couleur variables.

Carte : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

La tâche Carte requiert deux tables d'entrée :

- La source des données de réponse contenant les informations qui sont superposées sur la carte. Les données de réponse peuvent, par exemple, être les chiffres des ventes ou les démographies. Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données de réponse de la tâche.
- La source de données de la carte contenant les coordonnées X et Y pour la carte. Une fois que vous avez ouvert la tâche, vous devez spécifier une source de données.

Le nom de chaque source de données en entrée apparaît dans les champs **Source de données de la carte** et **Source des données de réponse**. Si vous créez une carte vide, la source de données de la carte et la source des données de réponse sont les mêmes.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Id	Identifie les variables des tables de cartographie et de réponse qui définissent l'aire de la carte. Chaque colonne que vous affectez à la fonction de la variable ID doit apparaître à la fois dans les tables de cartographie et de réponse. Les variables ID peuvent être numériques ou alphanumériques et doivent avoir les mêmes nom, type, et longueur dans les tables de cartographie et de réponse.
Réponse	Spécifie la variable de la table de réponse contenant les valeurs de réponse qui ne sont pas représentées sur la carte. La façon dont la variable de réponse est affichée dépend du type de la carte. Vous ne pouvez attribuer qu'une seule variable de réponse. ■ Dans les cartes choroplètes, les valeurs des variables de réponse sont représentées par différentes couleurs qui remplissent les aires. ■ Dans les cartes de blocs, les valeurs des variables de réponse sont représentées par diverses hauteurs, motifs ou couleurs.

Nom du rôle	Description
	■ Dans les cartes avec prismes, les valeurs des variables de réponse sont représentées par des aires de la carte en relief.
Grouper les graphiques par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de graphes créés. Un graphique distinct est créé pour chaque valeur unique.

Carte : Définir les options d'apparence

Définir les options de représentation cartographique

Dans la liste, sous l'en-tête Apparence, cliquez sur **Mise en correspondance** pour accéder à ces options.

Nom d'option	Description
Inclure toutes les aires de la carte à partir de la table des correspondances	Indique si les sous-divisions géographiques ne contenant pas de données doivent être incluses dans la sortie.
Génère un niveau de réponse séparé pour chaque valeur différente de la variable de réponse	Indique si vous souhaitez utiliser une couleur différente pour chaque valeur unique de la variable de réponse.
Accepter les valeurs manquantes comme variable de réponse valide	Indique si vous voulez inclure les valeurs manquantes en tant que niveaux valides de la variable de réponse.
Collecter les valeurs de réponse comme pourcentage de l'ensemble	Indique si toutes les valeurs de réponse (ou leurs statistiques) doivent être collectées et si chaque région doit être représentée comme pourcentage de l'ensemble. Activez l'option Indiquer les statistiques à représenter si vous voulez spécifier la statistique qui est utilisée pour calculer

Nom d'option	Description
	ce pourcentage. Si vous ne spécifiez pas de statistique à représenter, la valeur de réponse de la première observation seulement de chaque région sera calculée.
Indiquer les statistiques à représenter	Indique les statistiques à calculer pour la variable de réponse. Par défaut, la valeur de réponse de la première observation seulement de chaque région est calculée. Néanmoins, si votre variable de réponse contient des valeurs caractère, vous pouvez choisir de représenter la première observation ou la fréquence. Si votre variable de réponse contient des valeurs numériques, vous pouvez choisir de représenter la première observation, la somme, la fréquence, ou les valeurs de la moyenne de la variable de réponse.
Spécifier le nombre de niveaux de réponse à représenter	Indique le nombre de couleurs utilisées pour représenter les valeurs des données mises en correspondance.
Spécifier la densité des observations de la carte qui sont utilisées	Indique la proximité d'un point de la carte par rapport aux autres points. Si de nombreux points se trouvent très rapprochés (haute densité), il est possible d'éliminer certains d'entre eux sans nuire à la qualité de la carte. Vous pouvez indiquer un nombre entier entre 0 et 6 pour cette option. Si vous ne spécifiez pas de valeur, toutes les observations de la table seront affichées. Remarque : cette option n'est disponible que si votre table contient une variable appelée <code>DENSITY</code> .

Définir la présentation

Dans la liste, sous l'en-tête Apparence, cliquez sur **Présentation** pour accéder à ces options.

Pour tous les types de cartes, vous pouvez spécifier des couleurs pour les zones suivantes :

- remplissage des zones géographiques qui ne comportent pas de valeurs ("valeurs vides")
- contour des zones géographiques qui ne comportent pas de valeurs
- contour des zones géographiques qui comportent des valeurs

Vous pouvez également spécifier l'épaisseur du contour.

Si vous créez une carte de blocs, vous pouvez également spécifier la forme du bloc à utiliser sur la carte, la couleur du contour des blocs, ainsi que la taille (ou largeur) des blocs.

Personnaliser les couleurs

Dans la liste, sous l'en-tête Apparence, cliquez sur **Couleurs** pour accéder à ces options.

La tâche Carte utilise un jeu de couleurs fourni avec SAS. Vous pouvez cependant créer un jeu de couleurs personnalisé en activant la case à cocher **Utiliser des couleurs personnalisées** et en sélectionnant les couleurs souhaitées. La tâche utilise les couleurs dans l'ordre de leur apparition dans la liste des couleurs personnalisées.

Options d'apparence supplémentaires

Certaines options sont communes aux tâches graphiques. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir les options de légende" page 733](#)
- ["Personnaliser l'aire" page 734](#)

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Carte : Afficher un aperçu des résultats

Dans la liste, cliquez sur **Aperçu** pour voir un aperçu des résultats.

Grâce à cet aperçu, vous pouvez décider si vous avez besoin d'apporter des modifications supplémentaires avant de générer les résultats. Comme les tables de cartographie peuvent être volumineuses, l'utilisation de cet aperçu peut vous faire gagner du temps. Si vous changez des options dans la tâche, cliquez sur **Actualiser** afin de générer un nouvel aperçu des résultats. Si une erreur survient lors de la génération des résultats, le journal SAS apparaît.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Carte de la moyenne et de l'étendue

<i>À propos de la Carte de la moyenne et de l'étendue</i>	375
<i>Carte de la moyenne et de l'étendue : Attribuer des variables aux fonctions d'analyse</i>	376
<i>Cartes de contrôle : Indiquer des limites de contrôle</i>	377
<i>Carte de contrôle : Sélectionner les tests à effectuer</i>	379
<i>Cartes de contrôle : Définir les options d'apparence du graphe</i>	380
<i>Cartes de contrôle : Enregistrer les résultats d'analyse</i>	380
<i>Indiquer des titres et des notes de bas de page</i>	381
<i>Afficher les propriétés</i>	382

À propos de la Carte de la moyenne et de l'étendue

La carte de contrôle de Shewhart est un outil graphique et analytique qui permet de déterminer si un processus est sous contrôle statistique. La tâche Carte de la moyenne et de l'étendue crée une carte de la moyenne et de l'étendue pour les moyennes et les étendues des sous-groupes. Ces cartes sont utiles pour analyser la tendance centrale et la variabilité d'un processus.

Vous pouvez utiliser cette tâche pour analyser votre contrôle qualité. Par exemple, supposons que, dans une usine de fabrication de tranches de silicium, des lots de cinq tranches soient échantillonnés et que leur diamètre soit mesuré en millimètres. Les mesures issues de 25 lots sont enregistrées dans une table SAS, utilisée pour créer les cartes de la moyenne et de l'étendue. Chaque point de la carte de la moyenne représente la moyenne des mesures d'un lot particulier. Chaque point de la carte de l'étendue représente l'étendue des mesures d'un lot particulier. Si tous

les points sont compris dans les limites de contrôle, vous pouvez en conclure que le processus est sous contrôle statistique.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	SHEWHART
Produits SAS requis	SAS Base, SAS/QC, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Carte de la moyenne et de l'étendue : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Mesure	indique les variables de traitement à analyser. Ces variables contiennent les mesures individuelles. Vous devez attribuer au moins une variable à cette fonction.
Identificateur du sous-groupe	indique la variable qui identifie les sous-groupes dans les données. Les valeurs de la variable indiquent de quelle manière les observations de la table d'entrée sont organisées en sous-groupes rationnels. En général, les valeurs sont les suivantes : ■ les indices qui donnent l'ordre dans lequel les sous-échantillons sont collectés

Nom du rôle	Description
	■ les dates ou heures auxquelles les sous-échantillons ont été collectés ■ les libellés qui identifient de façon unique les échantillons de sous-groupe Un identifiant de la mesure est requis et vous ne pouvez attribuer qu'une variable à cette fonction. Pour trier les données selon les sous-groupes identifiés, activez la case à cocher Trier par sous-groupe.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.
Variables de bloc	indique deux variables maximum, qui regroupent les données par blocs de sous-groupes consécutifs. Ces blocs sont libellés dans une légende, et chaque variable de bloc fournit un niveau pour les libellés de la légende.

Cartes de contrôle : Indiquer des limites de contrôle

Dans la liste, cliquez sur **Limites de contrôle** pour accéder à ces options.

Nom d'option	Description
Limites Sigma	spécifie la largeur des limites de contrôle commun multiple de l'erreur type (sigma) des statistiques récapitulatives du sous-groupe tracé dans le graphique. La largeur doit être positive. Le multiple par défaut est 3.
Sélectionner une méthode de calcul	spécifie la méthode de calcul pour les limites de contrôle. Vous avez le choix entre ces options :

Nom d'option	Description
	■ Calculer les limites de contrôle à partir des données actives ■ Calculer les limites de contrôle à partir de la table sélectionnée Pour désigner la table contenant les limites de contrôle ou les paramètres à partir desquels les limites de contrôle peuvent être calculées, cliquez sur Parcourir. Sélectionnez la table à utiliser et cliquez sur Ouvrir. Remarque : la table que vous sélectionnez doit être une table LIMITS=. Pour plus d'informations sur les variables requises dans une table LIMITS=, consultez la documentation des procédures du graphique que vous êtes en train de créer. Pour représenter les statistiques descriptives de tous les sous-groupes, activez la case à cocher Trace les statistiques descriptives de tous les sous-groupes, que la taille de l'échantillon du sous-groupe soit égale à la taille nominale des limites de contrôle ou non. Vous pouvez représenter ces statistiques lorsque la taille de l'échantillon est la même pour pratiquement tous les sous-groupes et lorsque vous voulez afficher les limites de contrôle fixes (non soumises à variation) correspondant à la taille d'échantillon nominale n. Si vous sélectionnez cette option alors que les tailles des échantillons de vos sous-groupes sont très différentes, l'interprétation des limites de contrôle n'a de sens que pour les sous-groupes dont la taille de l'échantillon est égale à n. Si vous souhaitez ajouter des symboles spéciaux pour les points de données où la taille de l'échantillon est différente de n, activez la case à cocher Ajouter des symboles spéciaux pour les points correspondant aux tailles d'échantillon différentes de la taille nominale pour les limites de contrôle fixes. ■ Spécifier les limites de contrôle Dans les champs Supérieure, Centrale, et Inférieure, indiquez respectivement la limite de contrôle supérieure, la ligne centrale et la limite de contrôle inférieure. Si l'effectif du sous-groupe varie, vous devez indiquer un effectif nominal. Remarque : si vous attribuez des variables à la fonction Grouper l'analyse par, vous ne pouvez pas saisir des valeurs pour la limite de contrôle.

Carte de contrôle : Sélectionner les tests à effectuer

Dans la liste, cliquez sur **Tests** pour accéder à ces options.

Remarque : ces tests ne s'appliquent pas aux limites de contrôle qui ne sont pas trois limites sigma ou qui varient en fonction de la taille de l'échantillon du sous-groupe.

Nom d'option	Description
Sélectionner un test	Utilisez les cases à cocher pour demander un ou plusieurs tests pour des sujets spéciaux, également connus sous les noms de test de dérive : Run test, Pattern et Western Electric tests. Ces tests détectent les dérives particulières non aléatoires dans les points représentés sur le graphique de contrôle principal. L'occurrence du modèle non aléatoire est appelé signal. Les tests peuvent indiquer la présence d'une cause spéciale de variation. Lorsque vous sélectionnez l'un des tests, une description de ce test apparaît dans la zone Description.
Libellé	Utilisez les champs de saisie afin de saisir des libellés pour les points où les tests pour causes spéciales sont positifs. Les libellés ne peuvent contenir plus de 16 caractères. Chaque libellé apparaît à tous les points du test où un signal est rencontré.
Texte identifiant les signaux de test	■ Utilisez la liste déroulante afin de sélectionner une couleur pour les libellés que vous avez indiqués. ■ Utilisez la case à cocher Afficher lignes de zone pour inclure des lignes qui représentent les zones A, B et C sur un graphique principal. Ces zones se construisent en divisant l'intervalle entre les limites de contrôle en six sous-intervalles espacés de façon égale. ■ Utilisez la case Surpasser la limite 3 sigma pour tester des sujets spéciaux si vous spécifiez une limite de contrôle sigma différente du multiple par défaut de 3. Pour changer la limite sigma, cliquez sur Limites de contrôle dans le volet de sélection.

Nom d'option	Description
	■ Cochez la case Appliquer le test aux dérives des points se chevauchant afin d'appliquer des tests pour les causes spéciales aux dérives de points se chevauchant.

Cartes de contrôle : Définir les options d'apparence du graphe

Certaines options d'apparence sont communes à plusieurs tâches de la carte de contrôle. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir les options d'un axe" page 743](#)
- ["Ajouter des lignes de référence" page 744](#)
- ["Définir d'autres options de tracé" page 744](#)
- ["Définir les options de variable de bloc" page 746](#)

Cartes de contrôle : Enregistrer les résultats d'analyse

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Pour enregistrer les résultats de l'analyse, vous pouvez créer des tables de sortie contenant plusieurs statistiques. Activez les cases à cocher pour les données de sortie que vous voulez enregistrer de façon permanente.

Nom du rôle	Description
Statistiques de sous-groupe et table des limites de contrôle	crée une table de sortie contenant les informations représentées sur le graphe. Cela inclut les valeurs de la variable de sous-groupe ainsi que les statistiques descriptives correspondantes et les limites de contrôle. Remarque : cette option crée une table OUTTABLE=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option TABLE= décrite par la procédure SHEWHART. Pour plus

Nom du rôle	Description
	d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des statistiques de sous-groupe	crée une table de sortie contenant les statistiques descriptives du sous-groupe. Remarque : cette option crée une table OUTHISTORY=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option HISTORY= décrite par la procédure SHEWHART. Pour plus d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des limites de contrôle	crée une table de sortie qui enregistre les limites de contrôle. Vous pouvez utiliser cette table lorsque vous sélectionnez la méthode de calcul pour les limites de contrôle lors d'une utilisation ultérieure de cette tâche.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Carte de la moyenne et de l'écart-type

<i>À propos de la Carte de la moyenne et de l'écart-type</i>	383
<i>Carte de la moyenne et de l'écart-type : Attribuer des variables aux fonctions d'analyse</i>	384
<i>Cartes de contrôle : Indiquer des limites de contrôle</i>	385
<i>Carte de contrôle : Sélectionner les tests à effectuer</i>	387
<i>Cartes de contrôle : Définir les options d'apparence du graphe</i>	388
<i>Cartes de contrôle : Enregistrer les résultats d'analyse</i>	388
<i>Indiquer des titres et des notes de bas de page</i>	389
<i>Afficher les propriétés</i>	390

À propos de la Carte de la moyenne et de l'écart-type

La carte de contrôle de Shewhart est un outil graphique et analytique qui permet de déterminer si un processus est sous contrôle statistique. La tâche Carte de la moyenne et de l'écart-type crée une carte de la moyenne et de l'écart-type pour les moyennes et les écarts-types des sous-groupes. Ces cartes sont utiles pour analyser la tendance centrale et la variabilité d'un processus.

Vous pouvez utiliser cette tâche pour rechercher la distribution des sorties et déterminer si un processus est sous contrôle statistique. Par exemple, supposons qu'une compagnie pétrolière utilise une turbine qui vaporise l'eau pompée dans le sol afin de réduire la viscosité du pétrole et en faciliter l'extraction. Ce processus est répété 20 fois par jour, et la puissance utilisée (en kW) pour réchauffer l'eau à la température souhaitée est enregistrée. Chaque point de la carte de la moyenne représente la moyenne des mesures d'une journée. Chaque point de la carte de

l'écart-type représente l'écart-type des mesures d'une journée. Si tous les points sont compris dans les limites de contrôle, vous pouvez en conclure que le processus est sous contrôle statistique.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	SHEWHART
Produits SAS requis	SAS Base, SAS/QC, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Carte de la moyenne et de l'écart-type : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Mesure	indique les variables de traitement à analyser. Ces variables contiennent les mesures individuelles. Vous devez attribuer au moins une variable à cette fonction.
Identificateur du sous-groupe	indique la variable qui identifie les sous-groupes dans les données. Les valeurs de la variable indiquent de quelle manière les observations de la table d'entrée sont organisées en sous-groupes rationnels. En général, les valeurs sont les suivantes : ■ les indices qui donnent l'ordre dans lequel les sous-échantillons sont collectés

Nom du rôle	Description
	■ les dates ou heures auxquelles les sous-échantillons ont été collectés ■ les libellés qui identifient de façon unique les échantillons de sous-groupe Un identifiant de la mesure est requis et vous ne pouvez attribuer qu'une variable à cette fonction. Pour trier les données selon les sous-groupes identifiés, activez la case à cocher Trier par sous-groupe.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.
Variables de bloc	indique deux variables maximum, qui regroupent les données par blocs de sous-groupes consécutifs. Ces blocs sont libellés dans une légende, et chaque variable de bloc fournit un niveau pour les libellés de la légende.

Cartes de contrôle : Indiquer des limites de contrôle

Dans la liste, cliquez sur **Limites de contrôle** pour accéder à ces options.

Nom d'option	Description
Limites Sigma	spécifie la largeur des limites de contrôle commun multiple de l'erreur type (sigma) des statistiques récapitulatives du sous-groupe tracé dans le graphique. La largeur doit être positive. Le multiple par défaut est 3.
Sélectionner une méthode de calcul	spécifie la méthode de calcul pour les limites de contrôle. Vous avez le choix entre ces options :

Nom d'option	Description
	■ Calculer les limites de contrôle à partir des données actives ■ Calculer les limites de contrôle à partir de la table sélectionnée Pour désigner la table contenant les limites de contrôle ou les paramètres à partir desquels les limites de contrôle peuvent être calculées, cliquez sur Parcourir. Sélectionnez la table à utiliser et cliquez sur Ouvrir. Remarque : la table que vous sélectionnez doit être une table LIMITS=. Pour plus d'informations sur les variables requises dans une table LIMITS=, consultez la documentation des procédures du graphique que vous êtes en train de créer. Pour représenter les statistiques descriptives de tous les sous-groupes, activez la case à cocher Trace les statistiques descriptives de tous les sous-groupes, que la taille de l'échantillon du sous-groupe soit égale à la taille nominale des limites de contrôle ou non. Vous pouvez représenter ces statistiques lorsque la taille de l'échantillon est la même pour pratiquement tous les sous-groupes et lorsque vous voulez afficher les limites de contrôle fixes (non soumises à variation) correspondant à la taille d'échantillon nominale n. Si vous sélectionnez cette option alors que les tailles des échantillons de vos sous-groupes sont très différentes, l'interprétation des limites de contrôle n'a de sens que pour les sous-groupes dont la taille de l'échantillon est égale à n. Si vous souhaitez ajouter des symboles spéciaux pour les points de données où la taille de l'échantillon est différente de n, activez la case à cocher Ajouter des symboles spéciaux pour les points correspondant aux tailles d'échantillon différentes de la taille nominale pour les limites de contrôle fixes. ■ Spécifier les limites de contrôle Dans les champs Supérieure, Centrale, et Inférieure, indiquez respectivement la limite de contrôle supérieure, la ligne centrale et la limite de contrôle inférieure. Si l'effectif du sous-groupe varie, vous devez indiquer un effectif nominal. Remarque : si vous attribuez des variables à la fonction Grouper l'analyse par, vous ne pouvez pas saisir des valeurs pour la limite de contrôle.

Carte de contrôle : Sélectionner les tests à effectuer

Dans la liste, cliquez sur **Tests** pour accéder à ces options.

Remarque : ces tests ne s'appliquent pas aux limites de contrôle qui ne sont pas trois limites sigma ou qui varient en fonction de la taille de l'échantillon du sous-groupe.

Nom d'option	Description
Sélectionner un test	Utilisez les cases à cocher pour demander un ou plusieurs tests pour des sujets spéciaux, également connus sous les noms de test de dérive : Run test, Pattern et Western Electric tests. Ces tests détectent les dérives particulières non aléatoires dans les points représentés sur le graphique de contrôle principal. L'occurrence du modèle non aléatoire est appelé signal. Les tests peuvent indiquer la présence d'une cause spéciale de variation. Lorsque vous sélectionnez l'un des tests, une description de ce test apparaît dans la zone Description.
Libellé	Utilisez les champs de saisie afin de saisir des libellés pour les points où les tests pour causes spéciales sont positifs. Les libellés ne peuvent contenir plus de 16 caractères. Chaque libellé apparaît à tous les points du test où un signal est rencontré.
Texte identifiant les signaux de test	■ Utilisez la liste déroulante afin de sélectionner une couleur pour les libellés que vous avez indiqués. ■ Utilisez la case à cocher Afficher lignes de zone pour inclure des lignes qui représentent les zones A, B et C sur un graphique principal. Ces zones se construisent en divisant l'intervalle entre les limites de contrôle en six sous-intervalles espacés de façon égale. ■ Utilisez la case Surpasser la limite 3 sigma pour tester des sujets spéciaux si vous spécifiez une limite de contrôle sigma différente du multiple par défaut de 3. Pour changer la limite sigma, cliquez sur Limites de contrôle dans le volet de sélection.

Nom d'option	Description
	■ Cochez la case Appliquer le test aux dérives des points se chevauchant afin d'appliquer des tests pour les causes spéciales aux dérives de points se chevauchant.

Cartes de contrôle : Définir les options d'apparence du graphe

Certaines options d'apparence sont communes à plusieurs tâches de la carte de contrôle. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir les options d'un axe" page 743](#)
- ["Ajouter des lignes de référence" page 744](#)
- ["Définir d'autres options de tracé" page 744](#)
- ["Définir les options de variable de bloc" page 746](#)

Cartes de contrôle : Enregistrer les résultats d'analyse

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Pour enregistrer les résultats de l'analyse, vous pouvez créer des tables de sortie contenant plusieurs statistiques. Activez les cases à cocher pour les données de sortie que vous voulez enregistrer de façon permanente.

Nom du rôle	Description
Statistiques de sous-groupe et table des limites de contrôle	crée une table de sortie contenant les informations représentées sur le graphe. Cela inclut les valeurs de la variable de sous-groupe ainsi que les statistiques descriptives correspondantes et les limites de contrôle. Remarque : cette option crée une table OUTTABLE=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option TABLE= décrite par la procédure SHEWHART. Pour plus

Nom du rôle	Description
	d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des statistiques de sous-groupe	crée une table de sortie contenant les statistiques descriptives du sous-groupe. Remarque : cette option crée une table OUTHISTORY=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option HISTORY= décrite par la procédure SHEWHART. Pour plus d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des limites de contrôle	crée une table de sortie qui enregistre les limites de contrôle. Vous pouvez utiliser cette table lorsque vous sélectionnez la méthode de calcul pour les limites de contrôle lors d'une utilisation ultérieure de cette tâche.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Modèles mixtes

<i>À propos de la tâche Modèles mixtes</i>	391
<i>Modèles mixtes : Attribuer des variables aux fonctions d'analyse</i>	392
<i>Modèles mixtes : Créer un modèle d'effets fixes</i>	393
<i>Modèles mixtes : Définir les options du modèle d'effets fixes</i>	393
<i>Modèles mixtes : Définir des effets aléatoires</i>	395
<i>Modèles mixtes : Définir des effets répétitifs</i>	397
<i>Modèles mixtes : Spécifier les moyennes des moindres carrés</i>	399
<i>Modèles mixtes : Générer des graphes</i>	400
<i>Modèles mixtes : Définir les options de prédiction</i>	402
<i>Indiquer des titres et des notes de bas de page</i>	402
<i>Afficher les propriétés</i>	403

À propos de la tâche Modèles mixtes

Un modèle linéaire mixte est une généralisation du modèle linéaire standard, la généralisation permettant aux données de présenter une corrélation et une variabilité non constantes. Le modèle linéaire mixte vous offre une flexibilité de modélisation des moyennes de vos données, mais également de leurs variances et covariances.

Vous pouvez utiliser cette tâche pour modéliser une interaction entre deux effets principaux. Par exemple, supposons que vous ayez mesuré la taille de 18 personnes, classées par famille et par sexe. En utilisant la hauteur comme valeur dépendante, le genre comme effet de classification fixe, et la famille comme effet de classification aléatoire, vous pouvez utiliser la tâche Modèles mixtes pour fixer un modèle linéaire mixte sur ces données.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	MIXED
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Modèles mixtes : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variable dépendante	indique la variable à utiliser en tant que variable dépendante (de réponse). Vous devez attribuer une variable à cette fonction.
Variables quantitatives	spécifie les variables à utiliser comme variables indépendantes continues (explicatives).
Variables de classification	indique les variables à utiliser en tant qu'effets indépendants discrets. Les variables que vous attribuez à cette fonction peuvent être numériques ou alphanumériques, mais doivent avoir un nombre limité de valeurs discrètes.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables .

Nom du rôle	Description
	Remarque : vous ne pouvez pas grouper en fonction d'une variable que vous avez déjà sélectionnée comme variable dépendante ou variable indépendante.
Poids relatif	spécifie la variable utilisée pour pondérer les valeurs correspondantes de la variable sélectionnée comme variable de réponse. Vous ne pouvez attribuer qu'une seule variable à cette fonction.
Libellé d'identification	indique les variables supplémentaires à inclure dans les tables de valeurs prédites.

Modèles mixtes : Créer un modèle d'effets fixes

Dans la liste, cliquez sur **Modèle d'effets fixes** pour accéder à ces options.

Vous pouvez créer des modèles dans plusieurs tâches SAS. Pour plus d'informations, voir [Annexe 3, "Créer un modèle" page 749](#).

Modèles mixtes : Définir les options du modèle d'effets fixes

Dans la liste, cliquez sur **Options du modèle d'effets fixes** pour accéder à ces options.

Nom d'option	Description
Afficher les tests associés à l'intercept	Sélectionnez cette option pour afficher les résultats à partir de tests d'hypothèse réalisés sur la constante.
Type de test d'hypothèse	Chaque test d'hypothèse utilise au moins une fonction estimable pour tester l'hypothèse. Les fonctions estimables sont des combinaisons linéaires des paramètres du modèle. Le Type 3 est le test d'hypothèse par défaut. Vous pouvez ajouter ou

Nom d'option	Description
	supprimer un test d'hypothèse en activant la case à cocher correspondante.
Afficher les paramètres estimés	La valeur de l'estimation du paramètre dépend de la méthode d'estimation sélectionnée.
Intervalles de confiance pour les valeurs estimées du paramètre	Un intervalle de confiance se compose de bornes inférieure et supérieure qui constituent un intervalle de valeurs raisonnables pour le paramètre estimé avec un niveau de confiance prédéfini. Vous pouvez soit saisir une valeur pour un niveau de confiance soit en sélectionner une dans la liste déroulante. La valeur par défaut est 95 pour cent.
Méthode d'estimation	Une méthode d'estimation est la technique utilisée pour estimer les éléments de la partie de la variance et de la matrice de la covariance du modèle. La méthode par défaut est le Maximum de vraisemblance résiduelle . Vous pouvez sélectionner une autre méthode d'estimation dans la liste déroulante.
Méthode des degrés de liberté	Les degrés de liberté constituent des statistiques qui indiquent le nombre d'éléments d'information indépendants contenus dans une somme. Vous pouvez sélectionner l'une des méthodes suivantes dans la liste déroulante. La méthode par défaut est Défaut. Si vous sélectionnez cette option, la tâche Modèle mixte sélectionne la méthode par défaut en fonction du modèle que vous avez spécifié et selon que votre modèle contient des effets aléatoires et répétitifs. La liste suivante répertorie comment la méthode par défaut est sélectionnée en fonction de la présence d'un effet aléatoire ou répétitif : ■ Si votre modèle contient des effets aléatoires, la méthode par défaut est la Méthode Containment. ■ Si votre modèle ne contient que des effets répétitifs, la méthode par défaut est la méthode Entre et dans les portions spécialisées. ■ Si votre modèle contient à la fois des effets aléatoires et des effets répétitifs, la méthode par défaut est la Méthode Containment. ■ Si votre modèle ne contient pas d'effet aléatoire ou répétitif, la méthode par défaut est la méthode des résidus.
Fonctions estimables	Activez les cases Type I, Type II, et Type III pour afficher les coefficients de la matrice des fonctions estimables de Type I à Type III pour chaque effet du modèle. Activez la case à cocher face au type pour afficher les coefficients de la matrice pour une fonction estimable spécifique.

Nom d'option	Description
	Remarque : vous pouvez aussi afficher les fonctions estimables pour l'intercept si vous activez la case Afficher les tests associés à l'intercept .
Utiliser les paramètres de mise au point	Remarque : ces paramètres de mise au point ne sont utiles que lorsque des problèmes sont mal conditionnés. Pour changer la sensibilité dans les fonctions de type III, saisissez la nouvelle valeur dans la zone Sensibilité pour les fonctions de Type III. Tout élément contenu dans la base de la fonction estimable ayant une valeur absolue inférieure à la valeur spécifiée est défini à 0. La valeur par défaut est 1E-08. Pour changer la sensibilité utilisée lors de l'analyse d'une matrice, saisissez la valeur voulue dans la zone Valeur du test de la singularité. Si un élément pivot diagonal est inférieur à l'élément diagonal original de la matrice, multiplié par la valeur spécifiée, la colonne associée est déclarée linéairement dépendante des colonnes précédentes, et le paramètre associé est défini à 0. La valeur par défaut est 1E-07.

Modèles mixtes : Définir des effets aléatoires

Dans la liste, cliquez sur **Effets aléatoires** pour accéder à ces options.

- 1 Cliquez sur **Ajouter** pour ajouter un effet à la zone **Effets aléatoires à estimer**.
- 2 Pour chaque effet de la liste, indiquez des options dans la zone **Effets aléatoires et options**. Ces options sont disponibles :

Nom d'option	Description
Effets aléatoires	spécifie les effets aléatoires du modèle mixte. Par défaut, aucun effet aléatoire n'est utilisé. Vous pouvez également inclure un intercept pour un effet aléatoire du modèle. Utiliser la boîte de dialogue Générateur d'effets pour spécifier les effets aléatoires de votre modèle mixte. Pour spécifier un effet, vous devez attribuer au moins une variable à la fonction Variables quantitatives ou Variables de classification. Vous pouvez sélectionner des combinaisons de variables pour créer des effets croisés, imbriqués, factoriels, ou polynomiaux. Une fois que vous avez spécifié les effets aléatoires, cliquez sur OK

Nom d'option	Description
	pour enregistrer vos modifications et retourner dans la tâche Modèles mixtes. Pour les étapes sur la façon de créer un modèle, voir "Modèles mixtes : Créer un modèle d'effets fixes" page 393.
Intervalle de confiance	Indique si l'intervalle de confiance de type t doivent apparaître pour chacune des estimations des effets aléatoires.
Sujets de modèle	Spécifiez l'identifiant du sujet à inclure dans l'analyse. Une indépendance totale entre les sujets est supposée. Cette option produit une structure diagonale en blocs dans la matrice G, avec des blocs identiques. La matrice Z est modifiée pour s'adapter à la diagonalité des blocs. Utilisez la boîte de dialogue Générateur d'effets pour spécifier l'identifiant du sujet de votre modèle mixte. Pour spécifier un effet, vous devez attribuer au moins une variable à la fonction Variables de classification. Vous pouvez sélectionner des combinaisons de variables pour créer des effets croisés, imbriqués, factoriels, ou polynomiaux. L'effet que vous avez spécifié apparaît dans la zone Identifiant du sujet. Pour supprimer cet effet, cliquez sur Supprimer les effets. Pour enregistrer cet identifiant de sujet et revenir à la tâche Modèles mixtes, cliquez sur OK.
Effet de groupe	spécifie l'identifiant du groupe à inclure dans l'analyse. Cette option permet de définir un effet qui spécifie l'hétérogénéité dans la structure de covariance de la matrice G. Toutes les observations ayant le même niveau d'effet du groupe ont les mêmes paramètres de covariance. Chaque nouveau niveau de l'effet de groupe produit un nouvel ensemble de paramètres de covariance ayant la même structure que le groupe original. Utilisez la boîte de dialogue Générateur d'effets pour spécifier l'identifiant du groupe de votre modèle mixte. Pour spécifier un effet, vous devez attribuer au moins une variable à la fonction Variables de classification. Vous pouvez sélectionner des combinaisons de variables pour créer des effets croisés, imbriqués, factoriels, ou polynomiaux. L'effet que vous avez spécifié apparaît dans la zone Identifiant du groupe. Pour supprimer cet effet, cliquez sur Supprimer les effets. Pour enregistrer cet identifiant de groupe et revenir à la tâche Modèles mixtes, cliquez sur OK.
Structure de la covariance	Indiquez la structure de la covariance pour la matrice G .

Nom d'option	Description
Options supplémentaires	Indiquez si vous voulez afficher la matrice G dans l'analyse et les options à utiliser. Ces options sont disponibles : ■ Affiche la matrice G affiche la matrice G estimée. ■ Affiche la racine de Cholesky pour la matrice G affiche la racine triangulaire inférieure de Cholesky pour la matrice G estimée. ■ Affiche la racine inverse de Cholesky pour la matrice G affiche la racine inverse de Cholesky pour la matrice G estimée. ■ Affiche l'inverse de la matrice G affiche l'inverse de la matrice G estimée. ■ Affiche la matrice de corrélation affiche la matrice de corrélation correspondant à la matrice G estimée. ■ Inclut les résultats estimés affiche la solution des paramètres des effets aléatoires.

Modèles mixtes : Définir des effets répétitifs

Dans la liste, cliquez sur **Effets répétitifs** pour accéder à ces options.

Pour chaque effet de la liste, spécifiez les options dans la zone **Effets répétitifs et options**. Ces options sont disponibles :

Nom d'option	Description
Effets à utiliser	Sélectionnez l'effet intra-sujets à utiliser dans l'analyse. Les effets intra-sujets sont les variables que vous avez attribuées à la fonction Variables de classification. Vous pouvez éventuellement spécifier un effet intra-sujets si vos données ne sont pas triées en fonction de la mesure répétitive appropriée dans un sujet. Par exemple, si Heure est votre effet répétitif et si les données ne sont pas triées par heure dans chacun des sujets, vous devez indiquer Heure comme effet intra-sujets pour que votre modèle interprète les données correctement. Par défaut, aucun effet intra-sujets n'est utilisé.
Sujets de modèle	Spécifiez l'identifiant du sujet à inclure dans l'analyse. Pour la plupart des modèles de mesures répétitives, aucun effet

Nom d'option	Description
	répétitif n'est nécessaire. Si vous indiquez un identifiant du sujet, l'instruction répétitive est utilisée pour spécifier la matrice R dans le modèle mixte. L'identifiant du sujet définit les blocs de la matrice R. Une indépendance totale entre les sujets est supposée. Utilisez la boîte de dialogue Générateur d'effets pour spécifier l'identifiant du sujet de votre modèle mixte. Pour spécifier un effet, vous devez attribuer au moins une variable à la fonction Variables de classification. Vous pouvez sélectionner des combinaisons de variables pour créer des effets croisés, imbriqués, factoriels, ou polynomiaux. L'effet que vous avez spécifié apparaît dans la zone Identifiant du sujet. Pour supprimer cet effet, cliquez sur Supprimer les effets. Pour enregistrer cet identifiant de sujet et revenir à la tâche Modèles mixtes, cliquez sur OK.
Structure de la covariance	Indiquez la structure de la covariance pour la matrice R .
Effet de groupe	Spécifiez l'identifiant du groupe à inclure dans l'analyse. Cette option permet de définir un effet qui spécifie l'hétérogénéité dans la structure de covariance de la matrice G. Toutes les observations ayant le même niveau d'effet du groupe ont les mêmes paramètres de covariance. Chaque nouveau niveau de l'effet de groupe produit un nouvel ensemble de paramètres de covariance ayant la même structure que le groupe original. Utilisez la boîte de dialogue Générateur d'effets pour spécifier l'identifiant du groupe de votre modèle mixte. Pour spécifier un effet, vous devez attribuer au moins une variable à la fonction Variables de classification. Vous pouvez sélectionner des combinaisons de variables pour créer des effets croisés, imbriqués, factoriels, ou polynomiaux. L'effet que vous avez spécifié apparaît dans la zone Identifiant du groupe. Pour supprimer cet effet, cliquez sur Supprimer les effets. Pour enregistrer cet identifiant de groupe et revenir à la tâche Modèles mixtes, cliquez sur OK.
Options de la matrice R	Indiquez si vous voulez afficher la matrice R dans l'analyse et les options à utiliser pour cette matrice. Ces options sont disponibles : ■ Affiche la matrice R affiche la matrice R estimée. ■ Affiche la matrice de corrélation affiche la matrice de corrélation correspondant à la matrice R estimée. ■ Affiche l'inverse de la matrice R affiche l'inverse de la matrice R estimée.

Nom d'option	Description
	■ Affiche la racine de Cholesky pour la matrice R affiche la racine triangulaire inférieure de Cholesky pour la matrice R estimée.

Modèles mixtes : Spécifier les moyennes des moindres carrés

Dans la liste, cliquez sur **Tests a posteriori (ou tests post hoc) de comparaison des moyennes** pour accéder à ces options.

- 1 Cliquez sur **Ajouter** pour ajouter un effet dans la zone **Effets à estimer**.
- 2 Pour chaque effet de la liste, définissez les options de la zone **Test et options de la moyenne des moindres carrés**. Ces options sont disponibles :

Nom d'option	Description
Effets à utiliser	Sélectionnez les effets de classe à utiliser dans l'analyse. Les effets de classe sont les variables que vous avez attribuées à la fonction Variables de classification . Par défaut, aucun effet de classe n'est utilisé.
Comparaisons	Sélectionnez la méthode de comparaison à utiliser dans l'analyse. ■ Afficher les p-values des différences affiche les différences des moyennes des moindres carrés. Ces options sont disponibles : ☐ Aucun n'affiche pas les différences. ☐ Défaut affiche toutes les différences appariées. ☐ Toutes les différences appariées affiche toutes les différences appariées. Il s'agit de la présentation par défaut. ☐ Contrôle utilisant le premier niveau affiche les différences avec un contrôle qui, par défaut, est le premier niveau de chacun des effets spécifiés de la moyenne des moindres carrés. Les tests bilatéraux et l'intervalle de confiance sont associés à ces types de différence. ☐ Test pour inférieur au contrôle teste si les niveaux de non contrôle sont sensiblement inférieurs au contrôle.

Nom d'option	Description
	Vous pouvez utiliser ce test pour les résultats unilatéraux. ☐ Test pour inférieur au contrôle teste si les niveaux de non contrôle sont sensiblement supérieurs au contrôle. Vous pouvez utiliser ce test pour les résultats unilatéraux. ■ Méthode d'ajustement pour la comparaison permet de réaliser plusieurs ajustements pour la comparaison sur les p-values et l'intervalle de confiance pour les différences des moindres carrés.
Intervalle de confiance	Indiquez si vous souhaitez inclure l'intervalle de confiance de type t créé pour les moyennes des moindres carrés.
Covariables	Indiquez si vous souhaitez inclure les covariables dans le calcul et si vous souhaitez afficher les coefficients de la matrice L. ■ Valeurs de covariables indique comment calculer les valeurs de covariables. Ces options sont disponibles : ☐ Valeurs de la moyenne définit tous les effets des covariables égaux à leurs moyennes pour le calcul des moyennes des moindres carrés standard. ☐ Valeurs de moy. et produits croisés corrigés définit les effets des covariables comme étant identiques à leurs valeurs moyennes et intègre un ajustement aux produits croisés des covariables. ☐ Par variable permet de spécifier des valeurs arbitraires pour les covariables. Les colonnes supplémentaires dans la sortie indiquent les valeurs des covariables. ■ Coefficients de la matrice L affichés affiche les coefficients de la matrice L pour tous les effets des moindres carrés.

Modèles mixtes : Générer des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Par défaut, la sortie comprend tous les graphes appropriés pour la sélection de données actuelle. Vous pouvez néanmoins choisir les graphes à inclure dans la sortie en sélectionnant l'option **Liste personnalisée de graphes**. Vous avez le choix entre les options suivantes :

Nom d'option	Description
Boîte à moustaches	demande des boîtes à moustaches pour les effets de votre modèle qui sont composés des effets de classification uniquement. Remarque : ces effets peuvent impliquer plus d'une variable de classification (interaction et effets imbriqués), mais ils ne peuvent pas contenir des variables continues. Par défaut, les boîtes à moustaches sont créées sur la base des résidus bruts pour les effets qualifiants dans les modèles d'effets fixes, les modèles d'effets aléatoires et les modèles d'effets répétitifs.
Graphique de distance	crée un graphe de la vraisemblance ou de la distance de vraisemblance restreinte.
Graphes d'estimation de l'influence	crée les panels des estimations de suppression dans une analyse d'influence.
Panel de statistiques d'influence	crée les panels de statistiques d'influence. Pour l'analyse d'influence itérative, le panel affiche le D de Cook et les statistiques du rapport de covariance pour les effets fixes et les paramètres de covariance, vous permettant de mesurer l'impact sur les estimations et la précision pour les deux types d'estimation. Dans l'analyse non-itérative, seules les statistiques des effets fixes sont représentées.
Graphique des résidus bruts	crée un panel des résidus bruts. Par défaut, les résidus conditionnels sont générés.
Graphe du résidu 'studentisé'	crée un panel du résidu 'studentisé'. Par défaut, les résidus conditionnels sont générés.
Graphique des résidus de Pearson	crée un panel des résidus de Pearson. Par défaut, les résidus conditionnels sont générés.
Graphe du résidu du PRESS	crée un graphe des résidus du PRESS ou des statistiques du PRESS. Ces derniers sont basés sur une prédiction "leave-one-out" ou "leave-set-out" de la moyenne marginale.
Graphe du résidu à l'échelle	crée un panel des graphiques de résidus basés sur les résidus mis à l'échelle--les résidus de réponse et résidus marginaux qui ont été mis à l'échelle par la racine de Cholesky inverse pour la variance marginale-matrice de la covariance.

Modèles mixtes : Définir les options de prédiction

Dans la liste, cliquez sur **Prédictions** pour accéder à ces options.

Nom d'option	Description
Données à prédire	identifie les sources de données à utiliser pour prédire les données de sortie. Si vous activez la case Données supplémentaires, indiquez l'emplacement de ces données en cliquant sur Parcourir.
Enregistrer les données de sortie	spécifie s'il faut enregistrer les prévisions et les moyennes prédites dans une table de sortie : Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Afficher les prédictions	affiche les prédictions dans les résultats.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

.....

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

.....

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

.....

Scoring du modèle

<i>À propos de l'Assistant Scoring du modèle</i>	405
<i>Scoring du modèle : Sélectionner les données</i>	406
<i>Scoring du modèle : Sélectionner un modèle de scoring</i>	406
<i>Scoring du modèle : Mettre en correspondance les variables de données avec les champs obligatoires du modèle.</i>	407
<i>Scoring du modèle : Sélectionner les colonnes des données de sortie.</i>	407
<i>Scoring du modèle : Enregistrer les données de sortie</i>	408
<i>Scoring du modèle : Confirmer vos sélections</i>	408
<i>Scoring du modèle : Interpréter les résultats</i>	409

À propos de l'Assistant Scoring du modèle

L'Assistant **Scoring du modèle** permet d'évaluer une table par rapport à un modèle de prédiction SAS Enterprise Miner. Le classement ou le scoring des données brutes constitue le principal objectif du data mining. Vous pouvez utiliser un modèle de prédiction pour évaluer des tables contenant, par exemple, une liste d'emprunteurs, de répondants à une campagne de marketing direct, ou des tentatives de fraude. Vous pouvez également utiliser cette tâche pour créer une ou plusieurs variables de segment pour la table d'entrée.

Avant d'accéder à un modèle à partir de l'Assistant **Scoring du modèle**, le modèle qui crée le code du scoring doit d'abord être créé et validé dans SAS Enterprise Miner. Le code du scoring qui a été généré par le modèle doit également être enregistré dans le référentiel de métadonnées SAS.

Vous pouvez utiliser cet Assistant pour évaluer des tables de données d'attributs de clients. À partir de cette analyse, vous pouvez créer une liste de diffusion des clients les plus intéressés par un achat. Vous pouvez utiliser le code du scoring d'un

autre modèle pour attribuer chaque client à un segment de marché en fonction d'un ou de plusieurs scores d'attribut.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	Néant
Produits SAS requis	SAS Enterprise Miner
Autres produits SAS recommandés	Aucun

Scoring du modèle : Sélectionner les données

Par défaut, l'Assistant **Scoring du modèle** utilise les données actives du projet.

Cliquez sur **Suivant** pour sélectionner un modèle de scoring.

Scoring du modèle : Sélectionner un modèle de scoring

- 1 Avant de sélectionner un modèle de scoring, vous devez sélectionner les données.
- 2 Cliquez sur **Parcourir** pour sélectionner un modèle de scoring. La fenêtre **Scoring du modèle** apparaît. Sous l'onglet **Parcourir**, sélectionnez le modèle à utiliser.

Sélectionnez un modèle, puis cliquez sur **OK**. Le chemin d'accès au répertoire du modèle apparaît dans la zone **Modèle de scoring** et les propriétés du modèle, dans la zone **Attributs du modèle**. Pour afficher le code SAS généré par le modèle, cliquez sur **Source**.

- 3 Cliquez sur **Suivant** pour mettre en correspondance les variables de données avec les champs obligatoires du modèle.

Scoring du modèle : Mettre en correspondance les variables de données avec les champs obligatoires du modèle.

- 1 Avant de mettre en correspondance les variables de données avec les champs obligatoires du modèle, vous devez sélectionner un modèle de scoring.
- 2 Pour chaque champ obligatoire, sélectionnez la variable de données à utiliser dans la source de données. Par défaut, les champs du modèle et les variables de données portant le même nom et de type identique sont mis en correspondance. Pour modifier une variable de données mise en correspondance avec un champ du modèle, sélectionnez la variable et cliquez sur **Modifier**. Dans la liste déroulante qui apparaît, sélectionnez la variable de données que vous voulez utiliser.
- 3 Cliquez sur **Suivant** pour spécifier les colonnes de sortie.

Remarque : le bouton **Suivant** n'est disponible que lorsque toutes les variables de données ont été mises en correspondance avec les champs obligatoires du modèle.

Scoring du modèle : Sélectionner les colonnes des données de sortie.

- 1 Avant de spécifier les colonnes des données de sortie, vous devez mettre en correspondance les variables de données avec les champs obligatoires du modèle.
- 2 Dans les onglets **Variables de données** et **Sorties du modèle**, sélectionnez les éléments à inclure dans la sortie. L'onglet **Variables de données** comprend toutes les variables de la table d'entrée. L'onglet **Sorties du modèle** comprend les variables **Date**, **DateHeure** et **ModelKey**. Dans chaque onglet, sélectionnez les éléments à inclure dans la sortie et faites-les glisser vers le panneau **Colonnes sélectionnées**. Les variables et les sorties du modèle figurant dans le panneau **Colonnes sélectionnées** apparaîtront dans la sortie. Utilisez les flèches vers le haut et vers le bas pour définir l'ordre des variables et des sorties.

- 3 Cliquez sur **Suivant** pour enregistrer les données de sortie.

Scoring du modèle : Enregistrer les données de sortie

Avant d'enregistrer les données de sortie, vous devez sélectionner les colonnes des données de sortie.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches** > **Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Cliquez sur **Suivant** pour confirmer vos sélections.

La table de sortie contient les données d'origine ainsi que les scores tels que les valeurs prédites, les résidus et les résultats de classification. Pour plus d'informations sur les variables de la table de sortie, voir Interpréter les résultats.

Scoring du modèle : Confirmer vos sélections

Avant de confirmer vos sélections, vous devez sélectionner les colonnes des données de sortie.

La dernière étape contient un récapitulatif de vos sélections. Pour modifier l'une d'elles, cliquez sur **Précédent**. Pour exécuter la tâche, cliquez sur **Terminer**.

Scoring du modèle : Interpréter les résultats

La table de sortie contient les données d'origine ainsi que les scores tels que les valeurs prédites, les résidus et les résultats de classification. Le nom d'une variable de score est généralement formé par l'ajout d'un préfixe au nom de la variable cible, de la variable en entrée, ou de la table d'entrée.

Le tableau suivant répertorie chacun de ces préfixes, fournit une brève description du préfixe, indique si une variable cible est requise, et spécifie si cette variable est dans la matrice de décision. Pour plus d'informations sur l'interprétation des résultats, veuillez consulter l'aide de SAS Enterprise Miner.

Préfixe	Description	Cible Requise ?	Matrice de décision ?
AOV_	Identifie le numéro de classe de la variable de l'intervalle qui est divisé en 16 groupes à espace égal	Non	Non
B_	Variable binaire	Oui	Non
BL_	Meilleure perte possible des décisions, $-B(i)$	Oui	Oui
BP_	Meilleur profit possible des décisions, $B(i)$	Oui	Oui
CL_	Perte calculée à partir de la valeur cible, $-C(i)$	Oui	Oui
CP_	Profit calculé à partir de la valeur cible, $C(i)$	Oui	Oui
D_	Libellé de la décision choisie par le modèle	Non	Oui
EL_	Perte attendue pour la décision choisie par le modèle, $-E(i)$	Non	Oui
EM_CCF	Facteur du coût de crédit moyen	Non	Non

Préfixe	Description	Cible Requise ?	Matrice de décision ?
EM_CLASSIFICATION	Catégorie normalisée dans laquelle le cas est classifié	Non	Oui
EM_DECISION	Libellé de la décision choisie par le modèle	Non	Oui
EM_EVENTPROBABILITY	Probabilité a posteriori de l'événement cible	Non	Non
EM_EXPOSURE	Valeur moyenne d'exposition	Non	Non
EM_LGD	Perte moyenne étant donnée la valeur par défaut	Non	Non
EM_PD	Valeur prédite moyenne	Non	Non
EM_PREDICTION	Valeur prédite	Non	Non
EM_PROBABILITY	Probabilité a posteriori max. correspondant au niveau qui a été attribué à la variable EM_CLASSIFICATION	Non	Non
EM_PROFITLOSS	Profit ou perte attendu(e)	Non	Oui
EM_SEGMENT	Variable de segment	Non	Non
EP_	Profit attendu pour la décision choisie par le modèle, $E(i)$	Non	Oui
F_	Catégorie normalisée dont est issu le cas	Oui	Non
G_	Numéro d'identification du groupe auquel appartient l'observation	Non	Non
GRP_	Numéro d'identification du groupe auquel appartient l'observation	Non	Non
I_	Catégorie normalisée dans laquelle le cas est classifié	Non	Non
IC_	Coût d'investissement, $IC(i)$	Non	Oui

Préfixe	Description	Cible Requise ?	Matrice de décision ?
IMP_	Variable imputée	Non	Non
LBL	Libellé du texte de la variable à utiliser dans les tables et dans les graphes	Non	Non
M_	Variable de type indicateur identifiant les observations imputées manquantes ou les observations filtrées	—	Non
_N	Variable de voisinage répertoriant les observations de voisinage	Oui	Non
NODE	ID numérique pour chaque feuille d'une table de sortie de l'arbre de décision	Oui	Non
P_	Valeurs prédites et probabilités a posteriori	Non	Non
R_	Résidus simples ; cible moins prédiction	Oui	Non
REPL_	Variable remplacée	Oui	Non
ROI_	Retour sur investissement, $ROI(i)$	Oui	Oui
SCORECARD_BIN	Classe attribuée à chaque observation	Non	Non
SCORECARD_POINTS	Total de points de scorecard pour chaque individu	Non	Non
SEGMENT	Variable de segment	Non	Non
SOM_DIMENSION1	Identifie les lignes dans une map auto-organisatrice (SOM)	Non	Non
SOM_DIMENSION2	Identifie les lignes dans SOM	Non	Non
SOM_ID	Variable SOM	Non	Non

Préfixe	Description	Cible Requise ?	Matrice de décision ?
SOM_SEGMENT	Identifie les classes générées par une SOM	Non	Non
U_	Catégorie non formatée dans laquelle le cas est classifié	Non	Non
WOE_	Numéro d'identification du groupe auquel appartient l'observation	Non	Non

Les tables de sortie du score contiennent également une variable nommée `_WARN_`, qui indique les problèmes de calcul des valeurs prédites ou de prise de décisions. `_WARN_` est une variable alphanumérique qui soit est vide, ce qui indique qu'il n'y a pas eu de problèmes, soit contient l'un des codes caractère suivants :

Code	Signification
C	Variable de coût manquante
M	Entrées manquantes
P	Probabilité a posteriori incorrecte (par exemple, <0 ou >1)
S	Catégorie d'entrée non reconnue

Régression non linéaire

<i>À propos de la tâche Régression non linéaire</i>	413
<i>Régression non linéaire : Attribuer des variables aux fonctions d'analyse</i>	414
<i>Régression non linéaire : Définir les options du modèle</i>	415
<i>Régression non linéaire : Définir les options de calcul</i>	415
Spécifier des options d'itération	415
Spécifier une méthode de recherche par paliers	416
<i>Régression non linéaire : Générer des graphes</i>	417
<i>Régression non linéaire : Définition des options des résultats</i>	418
<i>Indiquer des titres et des notes de bas de page</i>	419
<i>Afficher les propriétés</i>	419

À propos de la tâche Régression non linéaire

La tâche Régression non linéaire produit des évaluations de moindres carrés ou des évaluations de moindres carrés pondérés des paramètres d'un modèle non linéaire. Les modèles non linéaires sont plus difficiles à indiquer et à estimer que les modèles linéaires. Plutôt que d'établir simplement la liste des variables de régression, vous devez sélectionner une expression de régression, déclarer les noms des paramètres et fournir les valeurs initiales de ces paramètres. Certains modèles sont difficiles à ajuster et rien ne garantit que la procédure pourra s'ajuster correctement au modèle.

Vous pouvez utiliser cette tâche si vous avez besoin de créer un modèle non linéaire pour votre expérience. Par exemple, supposons que vous souhaitiez étudier la relation entre la concentration et la vitesse d'un couple particulier d'enzyme et de substrat. Vous enregistrez le taux de réaction (vitesse) observé pour différentes concentrations de substrats et stockez ces valeurs dans une table SAS. La tâche

Régression non linéaire vous permet de calculer les estimations de chaque paramètre, l'erreur type asymptotique associée, ainsi que l'intervalle de confiance.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	NLIN
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	Aucun

Régression non linéaire : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variable dépendante	indique la variable à utiliser en tant que variable dépendante (de réponse). Vous devez attribuer une seule variable à cette fonction.
Variable explicative	indique quelles variables utiliser comme variables indépendantes (explicatives). Vous devez attribuer une seule variable à cette fonction.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : Vous ne pouvez pas grouper l'analyse en fonction d'une variable déjà attribuée à une autre fonction.

Régression non linéaire : Définir les options du modèle

Dans la liste, cliquez sur **Modèle** pour accéder à ces options.

Sélectionnez une expression de modèle non linéaire dans la liste déroulante. Cette expression est utilisée pour évaluer les valeurs prédites.

Dans la zone **Paramètres du modèle sélectionné**, indiquez les types de paramètre a, b et c comme valeur unique ou comme plage de valeurs. Si vous spécifiez une valeur unique comme type de paramètre, le paramètre reçoit la valeur du champ **Valeur unique**. Si vous spécifiez une plage de valeurs comme type de paramètre, vous pouvez également spécifier la première et la dernière valeur ainsi que la valeur d'incrément.

Pour définir des restrictions ou des bornes sur les paramètres, cliquez sur **Bornes du paramètre** et indiquez les limites dans la fenêtre **Bornes du paramètre**.

Remarque : la fenêtre **Bornes du paramètre** valide uniquement la syntaxe des limites que vous saisissez. Elle ne vérifie pas si les bornes ont une signification mathématique. Par exemple, vous pouvez saisir une borne telle que $2 < b < 1$, même si elle ne permet pas au modèle de converger.

Régression non linéaire : Définir les options de calcul

Spécifier des options d'itération

Dans la liste, cliquez sur **Calcul** pour accéder à ces options.

Nom du graphe	Description
Méthode	Faites une sélection dans la liste déroulante Méthode . Les options disponibles sont les suivantes : Gauss-Newton ,

Nom du graphe	Description
	Marquardt, Newton et Méthode dite de "Steepest descent" (Gradient).
Itérations maximales	Spécifiez le nombre d'itérations qu'effectue la tâche avant d'arrêter les tentatives de convergence. Cette valeur doit être un entier positif. Par défaut, le nombre maximal d'itérations est de 100.
Utiliser l'inversion de Moore-Penrose dans une estimation de paramètre	Cochez cette case pour utiliser l'inversion de Moore-Penrose dans une estimation de paramètre.
Remplacer le carré moyen de l'erreur	Cochez la case, puis spécifiez une valeur pour remplacer l'erreur du carré moyen lors du calcul des erreurs types des estimations dans le champ de saisie.
Mise au point de la minimalisation	Spécifiez dans les champs de saisie des critères de convergence ou un critère de singularité pour le modèle. La mesure de convergence du décalage relatif de Bates et Watts est utilisée par défaut pour déterminer la convergence. La valeur par défaut est $1.0E5$. Le critère de singularité est la grandeur absolue de la valeur pivot la plus faible autorisée lors d'une inversion ou d'une approximation de la matrice hessienne. La valeur par défaut est $1.0E8$.

Spécifier une méthode de recherche par paliers

Dans la liste, cliquez sur **Calcul** pour accéder à ces options.

Nom du graphe	Description
Méthode de recherche par paliers	Sélectionnez une méthode de recherche. Les options suivantes sont disponibles : ■ Aucune recherche par paliers supprime la restriction imposant la diminution de la valeur objective à chaque itération. La méthode Step halving est toujours utilisée pour restreindre les résultats estimés des paramètres dans les limites spécifiées et pour garantir que le nombre d'observations évaluables ne diminue pas. Cette option est utile pour les problèmes de moindres carrés repondérés de façon itérative. Il s'agit de la présentation par défaut.

Nom du graphe	Description
	■ Step halving recherche la plus petite somme des carrés des erreurs (SSE). ■ interpolation cubique estime le palier à l'aide d'une interpolation cubique. Si le palier estimé n'entraîne pas une diminution de la somme des carrés des erreurs (SEE), la méthode Step halving est utilisée. ■ Rechercher dans la Golden section détermine le palier par une recherche dans la Golden Section. Le paramètre Tau détermine la longueur de l'intervalle initial à rechercher. Le paramètre Rho spécifie le niveau de détail de la recherche.
Nbre max. de niveaux (de partitionnement)	Spécifiez le nombre maximal de niveaux de partitionnement à effectuer. Cette valeur doit être un entier positif.

Régression non linéaire : Générer des graphes

Dans la liste, ces options apparaissent sous l'en-tête **Graphiques**. Sélectionnez le type de graphe à générer. Vous pouvez créer un graphique de chaque type.

Remarque : la zone **Récapitulatif des graphes demandés** affiche la liste de tous les graphes que vous avez demandés.

Nom du graphe	Description
Prédit	Vous pouvez générer des nuages de points pour comparer les valeurs observées et prédites de chaque variable dépendante, ainsi que les valeurs observées de chaque variable dépendante et de chaque variable explicative (indépendante). Pour les graphes des variables observées par rapport aux variables indépendantes, vous pouvez également choisir de superposer l'intervalle de confiance ou les limites de prédiction sur tous les graphes des valeurs prédites. ■ Si vous spécifiez un intervalle de confiance, la courbe affiche les intervalles de confiance pour la moyenne. ■ Si vous spécifiez des limites de prédiction, la courbe affiche les intervalles de prédiction pour les réponses individuelles.

Nom du graphe	Description
	Pour les deux courbes, alpha est 0,05.
Résidu	Vous pouvez représenter les résidus ordinaires, standardisés et "studentisés" de chaque variable dépendante par rapport aux valeurs prédites de chaque variable dépendante ou aux valeurs de chaque variable explicative (indépendante). Les résidus sont calculés à l'aide des formules suivantes : ■ Un résidu ordinaire représente la différence entre la valeur observée et la valeur prédite. ■ Un résidu standardisé représente la valeur du résidu divisée par l'erreur type. ■ Un résidu 'studentisé' représente la valeur du résidu divisée par l'erreur type, avec suppression de l'observation courante.
Influence	Vous pouvez tracer l'effet levier par rapport aux valeurs prédites de la variable dépendante ou aux valeurs de la variable explicative (indépendante). L'effet levier pour une observation correspond à la valeur d'un élément diagonal correspondant dans la matrice H.

Régression non linéaire : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Vous pouvez choisir d'enregistrer de façon permanente les statistiques, les résultats estimés des paramètres ou bien les deux. Si vous choisissez d'enregistrer les résultats estimés des paramètres, vous pouvez également choisir d'enregistrer les estimations finales si la limite itérative est dépassée. Vous pouvez spécifier la limite itérative dans l'option Itérations maximales.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Vous pouvez également choisir d'afficher l'historique des itérations, la mesure du coefficient d'asymétrie (Skewness) de Hougaard, ou les deux, dans les résultats. La mesure du coefficient du coefficient d'asymétrie (Skewness) de Hougaard permet de déterminer si un paramètre est proche de la linéarité ou s'il contient un grand nombre de relations non linéaires.

Remarque : Ces options ne sont pas disponibles si vous activez la case à cocher **Supprimer l'affichage des résultats**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche

- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Analyse de variance non-paramétrique à une dimension

<i>À propos de la tâche Analyse de variance non paramétrique à une dimension</i>	421
<i>Analyse de variance non-paramétrique à une dimension : Attribuer des variables aux fonctions d'analyse</i>	422
<i>Analyse de variance non-paramétrique à une dimension : Sélectionner les analyses</i> ..	423
<i>Analyse de variance non-paramétrique à une dimension : Calculer les p-values exactes</i>	424
<i>Analyse de variance non-paramétrique à une dimension : Définition des options des résultats</i>	424
<i>Indiquer des titres et des notes de bas de page</i>	425
<i>Afficher les propriétés</i>	426

À propos de la tâche Analyse de variance non paramétrique à une dimension

La tâche Analyse de variance non paramétrique à une dimension se compose de tests de position non paramétriques et de différences d'échelle sur une dimension. La tâche fournit également une analyse de variance standard sur les données brutes et des statistiques basées sur la fonction de distribution empirique.

Vous pouvez utiliser cette tâche pour tester l'hypothèse nulle par rapport à une hypothèse alternative. Par exemple, supposons que 59 patientes souffrant de polyarthrite et ayant participé à un essai clinique soient réparties en deux groupes, l'un actif et l'autre placebo. La réponse (excellente=5, bonne=4, moyenne=3, légère=2, faible=1) de chaque patiente est enregistrée. En utilisant la tâche Analyse de variance non paramétrique à une dimension, vous pouvez tester l'hypothèse nulle selon laquelle les réponses des patientes ne présentent aucune différence

entre les deux groupes de traitement, par rapport à une autre hypothèse selon laquelle la réponse des patientes diffère entre les deux groupes.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	NPAR1WAY
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Analyse de variance non-paramétrique à une dimension : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables dépendantes	Indique les variables qui doivent être utilisées comme variables dépendantes (de réponse). Vous devez attribuer au moins une variable à cette fonction.
Variable indépendante	Indique la variable à utiliser comme variable indépendante (explicative). Vous devez attribuer une seule variable à cette fonction. Différentes analyses sont effectuées pour les sous-groupes définis par la variable de cette liste.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les

Nom du rôle	Description
	données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper en fonction d'une variable que vous avez déjà sélectionnée comme variable dépendante ou variable indépendante.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.

Analyse de variance non-paramétrique à une dimension : Sélectionner les analyses

Dans la liste, cliquez sur **Analyse** pour accéder à ces options.

Nom d'option	Description
Scores calculés dans les différents tests	Sélectionnez les analyses à effectuer.
Fonction de distribution empirique	Activez la case à cocher Calculer la loi empirique de la fonction de répartition (EDF) si vous voulez que la Fonction de distribution empirique soit réalisée.
Valeurs manquantes	Vous pouvez choisir d'inclure les valeurs manquantes en tant que niveau de classification possible en activant la case à cocher.
Correction de continuité	Incorpore une correction de continuité lors du calcul des statistiques de test standardisées pour les scores Wilcoxon et Siegel-Tukey. Si vous ne voulez pas utiliser la correction de continuité pour les tests, cochez la case Supprimer la correction de continuité .

Analyse de variance non-paramétrique à une dimension : Calculer les p -values exactes

Dans la liste, cliquez sur **p-values exactes** pour accéder à ces options.

Nom d'option	Description
P-values exactes	Par défaut, la tâche calcule les p-values asymptotiques au lieu des p-values exactes. Vous pouvez choisir de calculer les p-values exactes pour les différentes analyses en activant les cases à cocher appropriées. Remarque : les Scores de test ont été sélectionnés pour que les cases à cocher de la zone P-values exactes soient activées. Pour sélectionner les scores de test, cliquez sur Analyse dans la liste. Pour annuler tous les calculs de p-values exactes, cliquez sur Réinitialiser.
Durée du calcul	Dans la zone Délai limite de calcul, indiquez le délai limite de calcul de chaque p-value exacte, ou bien définissez le délai par Illimité. Le calcul des p-values exactes peut prendre beaucoup de temps et nécessiter une importante quantité de mémoire.

Analyse de variance non-paramétrique à une dimension : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Enregistrer les données de sortie	enregistre certains ou tous vos résultats dans une table de sortie. Cochez les cases des statistiques que vous voulez inclure dans la sortie. Remarque : les Scores de test ont été sélectionnés pour que les cases à cocher de la zone Statistiques à inclure soient activées. Pour sélectionner les scores de test, cliquez sur Analyse dans la liste. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Supprimer les rapports	N'affiche pas les résultats.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Carte np

<i>À propos de la tâche Carte np</i>	427
<i>Carte np : Attribuer des variables aux fonctions d'analyse</i>	428
<i>Cartes de contrôle : Indiquer des limites de contrôle</i>	429
<i>Carte de contrôle : Sélectionner les tests à effectuer</i>	431
<i>Cartes de contrôle : Définir les options d'apparence du graphe</i>	432
<i>Cartes de contrôle : Enregistrer les résultats d'analyse</i>	432
<i>Indiquer des titres et des notes de bas de page</i>	433
<i>Afficher les propriétés</i>	434

À propos de la tâche Carte np

La carte de contrôle de Shewhart est un outil graphique et analytique qui permet de déterminer si un processus est sous contrôle statistique. La tâche Carte np crée des cartes np présentant le nombre de non-conformités (défauts) des échantillons de sous-groupe.

Vous pouvez utiliser cette tâche pour suivre le nombre de défauts d'un processus de fabrication. Par exemple, supposons qu'un fabricant de composants électroniques fabrique des circuits par lots de 500 pièces et utilise une carte np pour suivre le nombre de circuits défectueux. Trente lots sont analysés, et le nombre de défauts est compté dans chacun d'eux. Le nombre de défauts est enregistré dans une table SAS, utilisée pour créer la carte np. Chaque point de la carte np représente le nombre d'éléments non conformes d'un sous-groupe particulier. Par exemple, la valeur représentée pour le premier lot est 5. Si tous les points sont compris dans les limites du contrôle (égales à 3 sigma par défaut), le processus est sous contrôle statistique.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	SHEWHART
Produits SAS requis	SAS Base, SAS/QC, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Carte np : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Nombre de non-conformités	indique les variables contenant le nombre de non-conformités. Vous devez attribuer au moins une variable à cette fonction
Identificateur du sous-groupe	indique la variable qui identifie les sous-groupes dans les données. Les valeurs de la variable indiquent de quelle manière les observations de la table d'entrée sont organisées en sous-groupes rationnels. En général, les valeurs sont les suivantes : ■ les indices qui donnent l'ordre dans lequel les sous-échantillons sont collectés ■ les dates ou heures auxquelles les sous-échantillons ont été collectés ■ les libellés qui identifient de façon unique les échantillons de sous-groupe Un identificateur de sous-groupe est requis et vous ne pouvez attribuer qu'une seule variable à cette fonction.

Nom du rôle	Description
	Pour trier les données selon les sous-groupes identifiés, activez la case à cocher Trier par sous-groupe .
Taille de l'échantillon du sous-groupe	spécifie la taille de l'échantillon du sous-groupe comme valeur de la variable. La taille de l'échantillon du sous-groupe est nécessaire et vous n'attribuez qu'une seule variable numérique à cette fonction. Activez la case à cocher Utiliser une valeur numérique pour le sous-groupe (au lieu d'une var.) lorsque la taille de l'échantillon de sous-groupe est déterminée et que vous devez définir une valeur constante pour la taille. Saisissez la valeur à utiliser dans la zone Taille de l'échantillon du sous-groupe. La taille de l'échantillon du sous-groupe doit être une valeur comprise entre la valeur maximale de la variable de traitement et 99999. Remarque : cette fonction n'est pas nécessaire si vous avez activé la case à cocher Utiliser une valeur numérique pour le sous-groupe (au lieu d'une var.) dans le volet Options.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.
Variables de bloc	indique deux variables maximum, qui groupent les données par blocs de sous-groupes consécutifs. Ces blocs sont libellés dans une légende, et chaque variable de bloc fournit un niveau pour les libellés de la légende.

Cartes de contrôle : Indiquer des limites de contrôle

Dans la liste, cliquez sur **Limites de contrôle** pour accéder à ces options.

Nom d'option	Description
Limites Sigma	spécifie la largeur des limites de contrôle commun multiple de l'erreur type (sigma) des statistiques récapitulatives du sous-groupe tracé dans le graphique. La largeur doit être positive. Le multiple par défaut est 3.
Sélectionner une méthode de calcul	spécifie la méthode de calcul pour les limites de contrôle. Vous avez le choix entre ces options : ■ Calculer les limites de contrôle à partir des données actives ■ Calculer les limites de contrôle à partir de la table sélectionnée Pour désigner la table contenant les limites de contrôle ou les paramètres à partir desquels les limites de contrôle peuvent être calculées, cliquez sur Parcourir. Sélectionnez la table à utiliser et cliquez sur Ouvrir. Remarque : la table que vous sélectionnez doit être une table LIMITS=. Pour plus d'informations sur les variables requises dans une table LIMITS=, consultez la documentation des procédures du graphique que vous êtes en train de créer. Pour représenter les statistiques descriptives de tous les sous-groupes, activez la case à cocher Trace les statistiques descriptives de tous les sous-groupes, que la taille de l'échantillon du sous-groupe soit égale à la taille nominale des limites de contrôle ou non. Vous pouvez représenter ces statistiques lorsque la taille de l'échantillon est la même pour pratiquement tous les sous-groupes et lorsque vous voulez afficher les limites de contrôle fixes (non soumises à variation) correspondant à la taille d'échantillon nominale n. Si vous sélectionnez cette option alors que les tailles des échantillons de vos sous-groupes sont très différentes, l'interprétation des limites de contrôle n'a de sens que pour les sous-groupes dont la taille de l'échantillon est égale à n. Si vous souhaitez ajouter des symboles spéciaux pour les points de données où la taille de l'échantillon est différente de n, activez la case à cocher Ajouter des symboles spéciaux pour les points correspondant aux tailles d'échantillon différentes de la taille nominale pour les limites de contrôle fixes. ■ Spécifier les limites de contrôle Dans les champs Supérieure, Centrale, et Inférieure, indiquez respectivement la limite de contrôle supérieure, la ligne centrale et la limite de contrôle

Nom d'option	Description
	inférieure. Si l'effectif du sous-groupe varie, vous devez indiquer un effectif nominal. Remarque : si vous attribuez des variables à la fonction Grouper l'analyse par , vous ne pouvez pas saisir des valeurs pour la limite de contrôle.

Carte de contrôle : Sélectionner les tests à effectuer

Dans la liste, cliquez sur **Tests** pour accéder à ces options.

Remarque : ces tests ne s'appliquent pas aux limites de contrôle qui ne sont pas trois limites sigma ou qui varient en fonction de la taille de l'échantillon du sous-groupe.

Nom d'option	Description
Sélectionner un test	Utilisez les cases à cocher pour demander un ou plusieurs tests pour des sujets spéciaux, également connus sous les noms de test de dérive : Run test, Pattern et Western Electric tests. Ces tests détectent les dérives particulières non aléatoires dans les points représentés sur le graphique de contrôle principal. L'occurrence du modèle non aléatoire est appelé signal. Les tests peuvent indiquer la présence d'une cause spéciale de variation. Lorsque vous sélectionnez l'un des tests, une description de ce test apparaît dans la zone Description .
Libellé	Utilisez les champs de saisie afin de saisir des libellés pour les points où les tests pour causes spéciales sont positifs. Les libellés ne peuvent contenir plus de 16 caractères. Chaque libellé apparaît à tous les points du test où un signal est rencontré.
Texte identifiant les signaux de test	■ Utilisez la liste déroulante afin de sélectionner une couleur pour les libellés que vous avez indiqués. ■ Utilisez la case à cocher Afficher lignes de zone pour inclure des lignes qui représentent les zones A, B et C

Nom d'option	Description
	sur un graphique principal. Ces zones se construisent en divisant l'intervalle entre les limites de contrôle en six sous-intervalles espacés de façon égale.
	■ Utilisez la case Surpasser la limite 3 sigma pour tester des sujets spéciaux si vous spécifiez une limite de contrôle sigma différente du multiple par défaut de 3. Pour changer la limite sigma, cliquez sur Limites de contrôle dans le volet de sélection.
	■ Cochez la case Appliquer le test aux dérives des points se chevauchant afin d'appliquer des tests pour les causes spéciales aux dérives de points se chevauchant.

Cartes de contrôle : Définir les options d'apparence du graphe

Certaines options d'apparence sont communes à plusieurs tâches de la carte de contrôle. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir les options d'un axe" page 743](#)
- ["Ajouter des lignes de référence" page 744](#)
- ["Définir d'autres options de tracé" page 744](#)
- ["Définir les options de variable de bloc" page 746](#)

Cartes de contrôle : Enregistrer les résultats d'analyse

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Pour enregistrer les résultats de l'analyse, vous pouvez créer des tables de sortie contenant plusieurs statistiques. Activez les cases à cocher pour les données de sortie que vous voulez enregistrer de façon permanente.

Nom du rôle	Description
Statistiques de sous-groupe et table des limites de contrôle	crée une table de sortie contenant les informations représentées sur le graphe. Cela inclut les valeurs de la variable de sous-groupe ainsi que les statistiques descriptives correspondantes et les limites de contrôle. Remarque : cette option crée une table OUTTABLE=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option TABLE= décrite par la procédure SHEWHART. Pour plus d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des statistiques de sous-groupe	crée une table de sortie contenant les statistiques descriptives du sous-groupe. Remarque : cette option crée une table OUTHISTORY=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option HISTORY= décrite par la procédure SHEWHART. Pour plus d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des limites de contrôle	crée une table de sortie qui enregistre les limites de contrôle. Vous pouvez utiliser cette table lorsque vous sélectionnez la méthode de calcul pour les limites de contrôle lors d'une utilisation ultérieure de cette tâche.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Analyse de variance à une dimension

<i>À propos de la tâche Analyse de variance à une dimension</i>	435
<i>Analyse de variance à une dimension : Attribuer des variables aux fonctions d'analyse</i>	436
<i>Analyse de variance à une dimension : sélection des tests</i>	437
<i>Analyse de variance à une dimension : Comparer les moyennes de groupe</i>	438
<i>Analyse de variance à une dimension : Sélectionner des statistiques pour les variables quantitatives</i>	439
<i>Analyse de variance à une dimension : Générer des graphes</i>	440
<i>Analyse de variance à une dimension : Définition des options des résultats</i>	441
<i>Indiquer des titres et des notes de bas de page</i>	442
<i>Afficher les propriétés</i>	442

À propos de la tâche Analyse de variance à une dimension

Une analyse de la variance simple (ANOVA) considère un facteur de traitement avec au moins deux niveaux de traitement. Le but de l'analyse est de tester les différences parmi les moyennes des niveaux et de quantifier ces différences. S'il y a deux niveaux de traitement, cette analyse équivaut à un test t comparant les moyennes de deux groupes.

Vous pouvez utiliser cette tâche pour réaliser les opérations suivantes :

- Étudier l'effet des bactéries sur la quantité d'azote contenue dans les plants de trèfle rouge. Le facteur utilisé est la souche bactérienne ; il possède six niveaux.

- Analyser un dispositif complet en blocs aléatoires. Par exemple, supposons que vous souhaitiez savoir si trois types de traitement ont des effets différents sur le rendement et la valeur d'une récolte. Vous choisissez d'utiliser un facteur bloquant visant à remédier à l'hétérogénéité probable des unités expérimentales en les rendant homogènes au sein de chaque bloc. Les trois types de traitement sont ensuite attribués aléatoirement dans chaque bloc.
- Comparer la durée de vie de trois batteries de marques différentes. Le facteur utilisé est la marque ; il possède trois niveaux.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	ANOVA
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Analyse de variance à une dimension : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables dépendantes	Indique les variables qui doivent être utilisées comme variables dépendantes (de réponse). Vous devez attribuer au moins une variable à cette fonction.

Nom du rôle	Description
Variable indépendante	Indique la variable à utiliser comme variable indépendante (explicative). Vous devez attribuer une seule variable à cette fonction.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper en fonction d'une variable que vous avez déjà sélectionnée comme variable dépendante ou variable indépendante.

Analyse de variance à une dimension : sélection des tests

Dans la liste, cliquez sur **Tests** pour accéder à ces options.

Vous pouvez utiliser l'analyse de variance à une dimension (ANOVA) lorsque la variable attribuée au rôle Variable indépendante a au moins deux valeurs distinctes. Vous pouvez tester votre modèle pour rechercher l'hypothèse des variances égales.

Nom d'option	Description
ANOVA de la variance pondérée de Welch	teste les moyennes de groupe en utilisant une variance pondérée. Vous pouvez utiliser ce test si l'hypothèse des variances égales est rejetée.
Tests pour la variance égale	sélectionnez l'un de ces tests : ■ Le Test de Bartlett est une modification du test normal du rapport de vraisemblance. Ce test calcule des taux d'erreur précis de Type I lorsque la distribution des données est normale. ■ Le test de Brown-Forsythe est une variante du test de Levene. Les variances égales sont déterminées à l'aide des écarts absolus des médianes de groupes. Bien que ce soit un bon test pour déterminer les différences de variances, il peut nécessiter une quantité importante de ressources si vos données contiennent plusieurs grands groupes.

Nom d'option	Description
	■ Le test de Levene est considéré comme le test d'homogénéité de variance standard. Il calcule les résidus au carré pour déterminer la variance égale.

Analyse de variance à une dimension : Comparer les moyennes de groupe

Dans la liste, sous l'en-tête Moyennes, cliquez sur **Comparaison** pour accéder à ces options.

Sélectionnez les tests de comparaison multiple de l'effet principal (c'est-à-dire, la variable indépendante) à stocker dans la table de sortie. Utilisez la liste déroulante **Niveau de conf.** afin de sélectionner un niveau de confiance pour tous les tests sélectionnés. La valeur par défaut est 95 pour cent.

Ces tests sont disponibles :

- Test t de Bonferroni
- Test de l'étendue 'studentisée' de Tukey (HSD)
- Test multivarié de Duncan
- Test t de Dunnett
- Test de la différence la moins significative de Fisher
- Procédure de comparaison multiple de Gabriel
- Tests multivariés de Student-Newman-Keuls
- Test t du rapport k de Waller-Duncan
- Procédure de comparaison multiple de Scheffe
- Test multivarié de Ryan-Einot-Gabriel-Welsch

Analyse de variance à une dimension : Sélectionner des statistiques pour les variables quantitatives

Dans la liste sous l'en-tête Moyennes, cliquez sur **Répartition** pour accéder à ces options.

Sélectionnez les statistiques descriptives à calculer et à afficher pour les variables dépendantes à chaque niveau de la variable indépendante.

Nom d'option	Description
Moyenne	moyenne arithmétique calculée en additionnant les valeurs d'une variable d'échantillon et en divisant cette somme par le nombre d'observations.
Écart-type	mesure statistique de la variabilité d'un groupe de valeurs de données. Cette mesure, qui est la mesure de dispersion d'une distribution de fréquence la plus largement utilisée, est égale à la racine carrée de la variance.
Erreur type	écart-type de la moyenne de l'échantillon. L'erreur type est définie comme le rapport entre l'écart-type de l'échantillon et la racine carrée de la taille de l'échantillon.
Variance	mesure statistique de la dispersion des données. Cette mesure est une moyenne de la dispersion totale au carré entre chaque observation et la moyenne de l'échantillon.
Nombre d'observations non manquantes	nombre d'observations sans valeur manquante.
Nombre d'observations manquantes	nombre d'observations pour lesquelles aucune valeur n'a été saisie.
Minimum	valeur la plus faible d'une observation.
Maximum	valeur la plus élevée d'une observation.

Analyse de variance à une dimension : Générer des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Utilisez les cases à cocher pour générer une boîte à moustaches, un graphe des moyennes ou bien les deux.

Nom du graphe	Description
Boîte à moustaches	Une boîte à moustaches affiche la moyenne, les quartiles et les observations minimales et maximales pour un groupe. Le graphique représente ces statistiques : ■ Le point extrême de la moustache supérieure représente la valeur maximale. ■ Le sommet de la boîte représente le troisième quartile (ou 75ème centile). ■ La longueur de la boîte représente l'intervalle interquartile (distance entre le troisième et le premier quartile). ■ Le point dans la boîte représente la moyenne. ■ La ligne horizontale dans la boîte représente la médiane (ou 50ème centile). ■ Le sommet de la boîte représente le premier quartile (ou 25ème centile). ■ Le point extrême de la moustache inférieure représente la valeur minimale.
Moyennes	Un graphe des moyennes affiche la moyenne et l'erreur type de la variable indépendante. Chaque point du graphique est entouré d'une barre de variance. Si vous choisissez Moyennes, vous pouvez également sélectionner ces options : ■ Indiquer si vous souhaitez utiliser l'erreur type de la moyenne ou l'écart-type pour tracer les barres. L'erreur type de la moyenne est l'écart-type de la moyenne. L'écart-type est la racine carrée de la variance de l'échantillon. ■ Utiliser la variance globale pour le calcul de la longueur des barres. La variance globale est la variance combinée de deux échantillons indépendants où les populations sont égales.

Nom du graphe	Description
	■ Choisissez une hauteur pour les barres : 1, 2 ou 3 unités standard. ■ Indiquer si l'axe vertical doit démarrer à 0.

Analyse de variance à une dimension : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Enregistrer les données de sortie des statistiques	enregistre les résultats dans une table de sortie. La table de sortie contient les sommes des carrés, les degrés de liberté, les statistiques de F, ainsi que les niveaux de probabilité pour chaque effet du modèle. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Supprimer l'affichage des résultats	N'affiche pas les résultats.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Fréquences à un critère de classification

<i>À propos de la tâche Fréquences à un critère de classification</i>	445
<i>Fréquences à un critère de classification : Attribuer des variables aux fonctions d'analyse</i>	446
<i>Fréquences à un critère de classification : Définir les options statistiques</i>	447
<i>Fréquences à un critère de classification : Générer des graphes</i>	448
<i>Fréquences à un critère de classification : Définition des options des résultats</i>	448
<i>Indiquer des titres et des notes de bas de page</i>	450
<i>Afficher les propriétés</i>	450

À propos de la tâche Fréquences à un critère de classification

Vous pouvez utiliser la tâche Fréquences à un critère de classification pour générer des tables de fréquences à partir de vos données. Vous pouvez également vous en servir pour réaliser les tests binomiaux et du Chi-deux.

Vous pouvez utiliser cette tâche pour analyser l'efficacité d'un nouveau médicament. Par exemple, supposons qu'un groupe de chercheurs souhaite évaluer l'efficacité d'un nouveau traitement dermatologique. Les dermatologues des cliniques participantes reçoivent une formation pour mener l'étude et évaluer le traitement. Après la formation, deux dermatologues examinent les patients atteints de l'affection à traiter dans le cadre d'une étude pilote et les évaluent. Ils peuvent utiliser la tâche Fréquences à un critère de classification pour établir leur diagnostic.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	FREQ
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Fréquences à un critère de classification : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables d'analyse	Indique les variables faisant l'objet de l'analyse. Pour chaque variable que vous attribuez à cette fonction, la tâche crée une table de fréquences à un critère de classification. Vous devez attribuer au moins une variable à cette fonction.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables .

Nom du rôle	Description
	Remarque : Vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà attribuée à une autre fonction.

Fréquences à un critère de classification : Définir les options statistiques

Dans la liste, cliquez sur **Statistiques** pour accéder à ces options.

Nom d'option	Description
Options de la table de fréquence	Sélectionnez les statistiques à inclure dans la table de fréquences à un critère de classification. Ces options sont disponibles : ■ L'option Fréquences et pourcentages avec cumuls permet de créer une table contenant les fréquences, fréquences cumulées, pourcentages des fréquences totales et pourcentages cumulés de chaque valeur de la variable d'analyse. ■ L'option Fréquences et fréquences cumulées permet de créer une table contenant les fréquences et fréquences cumulées de chaque valeur de la variable d'analyse. ■ L'option Fréquences et pourcentages permet de créer une table contenant les fréquences et pourcentages des fréquences totales pour chaque valeur de la variable d'analyse. ■ L'option Fréquences uniquement permet de créer une table contenant uniquement les fréquences de chaque valeur de la variable d'analyse.
Valeurs manquantes	Sélectionnez Afficher les fréquences pour inclure les valeurs manquantes dans les tables de fréquences. Sélectionnez Inclure dans les calculs pour inclure les fréquences des valeurs manquantes dans les tests de proportion binomiale et du Chi-deux.
Proportions binomiales	Sélectionnez le test asymptotique, les p-values exactes, ou les deux. Pour les proportions binomiales, indiquez une proportion

Nom d'option	Description
	test (valeur de proportion de l'hypothèse nulle) et un niveau de confiance.
Critère qualitatif de lissage du Khi-2	Sélectionnez le test asymptotique, les p -values exactes, ou les deux.
Calculs exacts	Le calcul des p-values exactes peut prendre beaucoup de temps et nécessiter une importante quantité de mémoire. Dans la zone Calculs exacts, activez la case Délai limite de calcul et spécifiez un délai limite de calcul (en secondes) pour chaque p-value de chaque table à double entrée. La valeur par défaut est 900 secondes. Pour calculer l'estimation Monte Carlo des p-values exactes au lieu de calculer directement ces valeurs, activez la case à cocher Utiliser l'estimation Monte Carlo. L'estimation Monte Carlo peut être utile pour résoudre des problèmes importants dont les calculs exacts exigent beaucoup de temps et de mémoire, mais pour lesquels des approximations asymptotiques risquent de ne pas être suffisantes.

Fréquences à un critère de classification : Générer des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Utilisez les cases à cocher pour afficher les histogrammes des fréquences. Vous pouvez afficher un histogramme horizontal, un histogramme vertical ou bien les deux.

Fréquences à un critère de classification : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Enregistrer les données de sortie	Pour enregistrer les tables de fréquence dans une table de sortie, activez la case à cocher Créer une table avec des fréquences et des pourcentages. Pour inclure les effectifs et le pourcentage cumulés des tables à un critère de classification dans les données de sortie, activez la case à cocher Ajouter des fréquences et des pourcentages cumulés à la table. Sélectionnez les variables d'analyse à inclure dans la table, puis indiquez l'emplacement dans lequel vous voulez enregistrer la table. Saisissez le chemin d'accès à la table sous la forme <code>\serveur\bibliothèque\membre</code>. Vous pouvez également cliquer sur Parcourir pour sélectionner un emplacement. Si vous avez sélectionné l'un des tests de proportion binomiale ou du Chi-deux, vous pouvez enregistrer définitivement les tables de fréquence dans une table de sortie en activant la case à cocher Ajouter la proportion binomiale et la statistique du Chi-deux dans la table. Indiquez l'emplacement dans lequel vous voulez enregistrer la table. Saisissez le chemin d'accès à la table sous la forme <code>\\serveur\bibliothèque\membre</code>. Vous pouvez également cliquer sur Parcourir pour sélectionner un emplacement. Remarque : vous pouvez enregistrer de façon permanente les proportions binomiales et statistiques du Chi-deux pour seulement la dernière variable de la fonction Variables d'analyse.
Supprimer l'affichage des résultats	N'affiche pas les résultats. Remarque : pour exécuter la tâche, vous devez afficher la sortie ou enregistrer les résultats dans une table. Si vous avez désactivé la case à cocher Créer une table avec des fréquences et des pourcentages et activé la case à cocher Supprimer l'affichage des résultats et des graphes, vous ne pouvez pas exécuter ni enregistrer la tâche.
Sélectionner l'ordre des données de sortie	Pour indiquer l'ordre des valeurs dans la table, sélectionnez l'ordre des données de sortie dans la liste déroulante Trier données de sortie par. Ces options sont disponibles : ■ L'option Ordre de la table trie les valeurs en fonction de leur ordre dans la table d'entrée. ■ L'option Valeurs formatées trie les valeurs en fonction de leur valeur formatée. ■ L'option Fréquences décroissantes trie les valeurs par nombre d'occurrences décroissant. ■ Valeurs non formatées trie les valeurs en fonction de leur valeur non formatée.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Carte p

<i>À propos de la Carte p</i>	453
<i>Carte p : Attribuer des variables aux fonctions d'analyse</i>	454
<i>Cartes de contrôle : Indiquer des limites de contrôle</i>	455
<i>Carte de contrôle : Sélectionner les tests à effectuer</i>	457
<i>Cartes de contrôle : Définir les options d'apparence du graphe</i>	458
<i>Cartes de contrôle : Enregistrer les résultats d'analyse</i>	458
<i>Indiquer des titres et des notes de bas de page</i>	459
<i>Afficher les propriétés</i>	460

À propos de la Carte p

La carte de contrôle de Shewhart est un outil graphique et analytique qui permet de déterminer si un processus est sous contrôle statistique. La tâche Carte p crée des cartes p pour la proportion d'unités non conformes (défauts) par échantillon.

Vous pouvez utiliser cette tâche pour suivre le nombre de défauts d'un processus de fabrication. Par exemple, supposons qu'un fabricant de composants électroniques fabrique des circuits par lots de 500 pièces et utilise une carte p pour suivre le nombre de circuits défectueux. Trente lots sont analysés, et le nombre de défauts est compté dans chacun d'eux. Le nombre de défauts est enregistré dans une table SAS, utilisée pour créer la carte p. Chaque point de la carte p représente la proportion d'éléments non conformes d'un sous-groupe particulier. En l'occurrence, si le nombre de défauts du premier circuit est 5, la valeur tracée pour le premier lot correspond à $5/500 = 0,01$. Si tous les points sont compris dans les limites du contrôle (égales à 3 sigma par défaut), le processus est sous contrôle statistique.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	SHEWHART
Produits SAS requis	SAS Base, SAS/QC, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Carte p : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Proportion d'éléments non conformes	indique les variables qui contiennent le nombre d'éléments non conformes. Vous devez attribuer au moins une variable à cette fonction
Identificateur du sous-groupe	indique la variable qui identifie les sous-groupes dans les données. Les valeurs de la variable indiquent de quelle manière les observations de la table d'entrée sont organisées en sous-groupes rationnels. En général, les valeurs sont les suivantes : ■ les indices qui donnent l'ordre dans lequel les sous-échantillons sont collectés ■ les dates ou heures auxquelles les sous-échantillons ont été collectés ■ les libellés qui identifient de façon unique les échantillons de sous-groupe Un identificateur de sous-groupe est requis et vous ne pouvez attribuer qu'une seule variable à cette fonction.

Nom du rôle	Description
	Pour trier les données selon les sous-groupes identifiés, activez la case à cocher Trier par sous-groupe .
Taille de l'échantillon du sous-groupe	spécifie la taille de l'échantillon du sous-groupe comme valeur de la variable. La taille de l'échantillon du sous-groupe est nécessaire et vous n'attribuez qu'une seule variable numérique à cette fonction. Activez la case à cocher Utiliser une valeur numérique pour le sous-groupe (au lieu d'une var.) lorsque la taille de l'échantillon de sous-groupe est déterminée et que vous devez définir une valeur constante pour la taille. Saisissez la valeur à utiliser dans la zone Taille de l'échantillon du sous-groupe. La taille de l'échantillon du sous-groupe doit être une valeur comprise entre la valeur maximale de la variable de traitement et 99999. Remarque : cette fonction n'est pas nécessaire si vous avez activé la case à cocher Utiliser une valeur numérique pour le sous-groupe (au lieu d'une var.) dans le volet Options.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.
Variables de bloc	indique deux variables maximum, qui regroupent les données par blocs de sous-groupes consécutifs. Ces blocs sont libellés dans une légende, et chaque variable de bloc fournit un niveau pour les libellés de la légende.

Cartes de contrôle : Indiquer des limites de contrôle

Dans la liste, cliquez sur **Limites de contrôle** pour accéder à ces options.

Nom d'option	Description
Limites Sigma	spécifie la largeur des limites de contrôle commun multiple de l'erreur type (sigma) des statistiques récapitulatives du sous-groupe tracé dans le graphique. La largeur doit être positive. Le multiple par défaut est 3.
Sélectionner une méthode de calcul	spécifie la méthode de calcul pour les limites de contrôle. Vous avez le choix entre ces options : ■ Calculer les limites de contrôle à partir des données actives ■ Calculer les limites de contrôle à partir de la table sélectionnée Pour désigner la table contenant les limites de contrôle ou les paramètres à partir desquels les limites de contrôle peuvent être calculées, cliquez sur Parcourir. Sélectionnez la table à utiliser et cliquez sur Ouvrir. Remarque : la table que vous sélectionnez doit être une table LIMITS=. Pour plus d'informations sur les variables requises dans une table LIMITS=, consultez la documentation des procédures du graphique que vous êtes en train de créer. Pour représenter les statistiques descriptives de tous les sous-groupes, activez la case à cocher Trace les statistiques descriptives de tous les sous-groupes, que la taille de l'échantillon du sous-groupe soit égale à la taille nominale des limites de contrôle ou non. Vous pouvez représenter ces statistiques lorsque la taille de l'échantillon est la même pour pratiquement tous les sous-groupes et lorsque vous voulez afficher les limites de contrôle fixes (non soumises à variation) correspondant à la taille d'échantillon nominale n. Si vous sélectionnez cette option alors que les tailles des échantillons de vos sous-groupes sont très différentes, l'interprétation des limites de contrôle n'a de sens que pour les sous-groupes dont la taille de l'échantillon est égale à n. Si vous souhaitez ajouter des symboles spéciaux pour les points de données où la taille de l'échantillon est différente de n, activez la case à cocher Ajouter des symboles spéciaux pour les points correspondant aux tailles d'échantillon différentes de la taille nominale pour les limites de contrôle fixes. ■ Spécifier les limites de contrôle Dans les champs Supérieure, Centrale, et Inférieure, indiquez respectivement la limite de contrôle supérieure, la ligne centrale et la limite de contrôle

Nom d'option	Description
	inférieure. Si l'effectif du sous-groupe varie, vous devez indiquer un effectif nominal. Remarque : si vous attribuez des variables à la fonction Grouper l'analyse par , vous ne pouvez pas saisir des valeurs pour la limite de contrôle.

Carte de contrôle : Sélectionner les tests à effectuer

Dans la liste, cliquez sur **Tests** pour accéder à ces options.

Remarque : ces tests ne s'appliquent pas aux limites de contrôle qui ne sont pas trois limites sigma ou qui varient en fonction de la taille de l'échantillon du sous-groupe.

Nom d'option	Description
Sélectionner un test	Utilisez les cases à cocher pour demander un ou plusieurs tests pour des sujets spéciaux, également connus sous les noms de test de dérive : Run test, Pattern et Western Electric tests. Ces tests détectent les dérives particulières non aléatoires dans les points représentés sur le graphique de contrôle principal. L'occurrence du modèle non aléatoire est appelé signal. Les tests peuvent indiquer la présence d'une cause spéciale de variation. Lorsque vous sélectionnez l'un des tests, une description de ce test apparaît dans la zone Description .
Libellé	Utilisez les champs de saisie afin de saisir des libellés pour les points où les tests pour causes spéciales sont positifs. Les libellés ne peuvent contenir plus de 16 caractères. Chaque libellé apparaît à tous les points du test où un signal est rencontré.
Texte identifiant les signaux de test	■ Utilisez la liste déroulante afin de sélectionner une couleur pour les libellés que vous avez indiqués. ■ Utilisez la case à cocher Afficher lignes de zone pour inclure des lignes qui représentent les zones A, B et C

Nom d'option	Description
	sur un graphique principal. Ces zones se construisent en divisant l'intervalle entre les limites de contrôle en six sous-intervalles espacés de façon égale.
	■ Utilisez la case Surpasser la limite 3 sigma pour tester des sujets spéciaux si vous spécifiez une limite de contrôle sigma différente du multiple par défaut de 3. Pour changer la limite sigma, cliquez sur Limites de contrôle dans le volet de sélection.
	■ Cochez la case Appliquer le test aux dérives des points se chevauchant afin d'appliquer des tests pour les causes spéciales aux dérives de points se chevauchant.

Cartes de contrôle : Définir les options d'apparence du graphe

Certaines options d'apparence sont communes à plusieurs tâches de la carte de contrôle. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir les options d'un axe" page 743](#)
- ["Ajouter des lignes de référence" page 744](#)
- ["Définir d'autres options de tracé" page 744](#)
- ["Définir les options de variable de bloc" page 746](#)

Cartes de contrôle : Enregistrer les résultats d'analyse

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Pour enregistrer les résultats de l'analyse, vous pouvez créer des tables de sortie contenant plusieurs statistiques. Activez les cases à cocher pour les données de sortie que vous voulez enregistrer de façon permanente.

Nom du rôle	Description
Statistiques de sous-groupe et table des limites de contrôle	crée une table de sortie contenant les informations représentées sur le graphe. Cela inclut les valeurs de la variable de sous-groupe ainsi que les statistiques descriptives correspondantes et les limites de contrôle. Remarque : cette option crée une table OUTTABLE=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option TABLE= décrite par la procédure SHEWHART. Pour plus d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des statistiques de sous-groupe	crée une table de sortie contenant les statistiques descriptives du sous-groupe. Remarque : cette option crée une table OUTHISTORY=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option HISTORY= décrite par la procédure SHEWHART. Pour plus d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des limites de contrôle	crée une table de sortie qui enregistre les limites de contrôle. Vous pouvez utiliser cette table lorsque vous sélectionnez la méthode de calcul pour les limites de contrôle lors d'une utilisation ultérieure de cette tâche.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.
.....

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

.....
Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.
.....

Graphique de Pareto

<i>À propos de la tâche Graphique de Pareto</i>	461
<i>Graphique de Pareto : Attribuer des variables aux fonctions d'analyse</i>	462
<i>Graphique de Pareto : Limiter le nombre de catégories</i>	463
<i>Graphique de Pareto : Définir les options d'apparence du graphe</i>	465
Définir les options des barres	465
Graphique de Pareto : Définir les options d'un axe	466
Graphique de Pareto : Ajouter des lignes de référence	466
Graphique de Pareto : Ajouter un encart	466
Définir d'autres options de tracé	467
<i>Indiquer des titres et des notes de bas de page</i>	468
<i>Afficher les propriétés</i>	469

À propos de la tâche Graphique de Pareto

Le graphique de Pareto constitue l'un des sept outils de base utilisés pour l'amélioration de la qualité dans l'industrie américaine et japonaise. Le graphique de Pareto vertical présente la fréquence relative des problèmes apparaissant dans un processus sous forme de barres de hauteur décroissante de la gauche vers la droite. Le graphique de Pareto horizontal présente les informations sous forme de barres de longueur décroissante du haut vers le bas.

Les graphiques de Pareto constituent un outil permettant de visualiser le principe de Pareto, qui déclare qu'un petit ensemble de problèmes, les 'vital few', se produisent beaucoup plus fréquemment que les problèmes restants, les 'trivial many'. Le graphique de Pareto vous aide à établir des priorités pour les activités d'amélioration de qualité en identifiant les problèmes qui méritent le plus d'attention et dont les solutions sont susceptibles d'avoir le plus grand impact.

Vous pouvez utiliser la tâche Graphique de Pareto pour créer un graphique de Pareto permettant de comparer les résultats qui précèdent et suivent un

événement. Par exemple, supposons que pendant la fabrication d'un condensateur semi-conducteur à oxyde métallique, des causes de défaut soient enregistrées avant et après le nettoyage d'un tube dans le four à diffusion. Dans la sortie, les deux graphiques de Pareto sont empilés de manière à déterminer si le processus de nettoyage a permis d'augmenter ou de diminuer le nombre de défauts.

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	PARETO
Produits SAS requis	SAS Base, SAS/QC, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Graphique de Pareto : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Traiter variable	indique la variable du processus à analyser. Cette variable détermine les catégories de Pareto pour le diagramme. Une variable de traitement est nécessaire et vous ne pouvez attribuer qu'une variable à cette fonction.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.

Nom du rôle	Description
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Poids relatif	spécifie la variable à utiliser comme variable de poids relatif. Lorsque vous attribuez une variable à cette fonction, les hauteurs des barres représentent les fréquences pondérées des catégories. Une variable de poids relatif est utile pour effectuer une analyse de Pareto basée sur le coût plutôt que sur la fréquence des occurrences. Vous ne pouvez attribuer qu'une seule variable à cette fonction.
Variables de classification	indique une ou deux variables à utiliser en tant que variables de classification. Lorsque vous attribuez une variable à cette fonction, la tâche crée un graphique de Pareto comparatif qui vous permet de comparer les fréquences des différents niveaux de classification.

Graphique de Pareto : Limiter le nombre de catégories

Dans la liste, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Afficher catégories	Vous pouvez limiter le nombre de catégories de Pareto à afficher dans le diagramme. Par défaut, toutes les catégories du processus sont représentées par des barres sur le graphique de Pareto. Pour modifier le nombre de catégories affichées, sélectionnez l'une des options suivantes : ■ L'option N valeurs les plus élevées permet de créer un graphique qui affiche uniquement les catégories Pareto comportant les n valeurs les plus élevées. La valeur par défaut de n est 5. ■ L'option Jusqu'à N % de la fréquence totale crée un graphique qui affiche les catégories Pareto dont la somme permet d'obtenir le pourcentage cumulé inférieur ou égal à n pour cent. La valeur par défaut de n est 90.

Nom d'option	Description
Inclure les valeurs manquantes pour les variables de traitement	■ L'option Au moins N % de la fréquence totale permet de créer un graphique qui affiche les catégories Pareto dont les pourcentages sont supérieurs ou égaux à n pour cent. La valeur par défaut de n est 5 pour cent. La valeur par défaut de n est 5. Après avoir sélectionné cette option, saisissez une valeur pour n dans la zone N. Pour créer une nouvelle catégorie fusionnant toutes les catégories que vous voulez exclure, activez la case à cocher Afficher les catégories non incluses ci-dessus. Par défaut, cette catégorie prend le nom de Autre. Vous pouvez indiquer un autre libellé pour la catégorie dans la zone Libellé ; il doit contenir un maximum de 32 caractères. Activez cette case à cocher pour traiter les valeurs manquantes en tant que catégorie. Cette catégorie est représentée par une barre sur le graphique de Pareto. Si la variable est alphanumérique, une valeur manquante est définie en tant que valeur interne vide (non formatée). Si la variable est numérique, une valeur manquante est définie comme n'importe quelle valeur manquante SAS. Par défaut, les valeurs manquantes sont exclues de l'analyse.
Enregistrer la variable de sortie dans la table	Activez cette case à cocher pour enregistrer vos résultats dans une table de sortie. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.

Graphique de Pareto : Définir les options d'apparence du graphe

Définir les options des barres

Dans la liste, sous l'en-tête Apparence, cliquez sur **Barres** pour accéder à ces options.

Nom d'option	Description
Orientation et format des barres	Sélectionnez Vertical pour afficher les barres verticalement. C'est l'orientation par défaut. Sélectionnez Horizontal pour afficher les barres horizontalement. Utilisez la liste déroulante Barres pour changer la couleur des barres. Utilisez la liste déroulante Contour pour changer la couleur du contour des barres.
Attribuer un libellé aux barres	Vous pouvez placer des libellés au-dessus des barres. Sélectionnez Hauteur de la barre pour attribuer un libellé à la hauteur de chaque barre. Sélectionnez Pourcentage cumulé pour attribuer un libellé au pourcentage cumulé de chaque barre. L'option par défaut est de ne pas mettre de libellé.
Afficher la courbe	Par défaut, le graphique de Pareto affiche la courbe des pourcentages cumulés. Pour modifier la couleur de la courbe, utilisez la liste déroulante Couleur de la courbe. Pour attribuer un libellé aux points de la courbe avec leurs valeurs, activez la case à cocher Libeller les points avec leurs valeurs. Pour supprimer cette courbe et l'axe secondaire, désactivez la case à cocher Afficher courbe.
Barres en surbrillance	Activez la case à cocher Les plus élevées pour mettre en surbrillance les barres avec les n valeurs les plus élevées. Sélectionnez Les plus basses pour mettre en surbrillance les barres avec les n valeurs les plus basses. Utilisez la liste déroulante Couleur pour changer la couleur de la surbrillance. Pour sélectionner plusieurs barres, utilisez les flèches vers le haut et vers le bas afin d'indiquer la valeur de n dans la zone Nombre de barres. La valeur de n doit être inférieure ou égale au nombre de catégories du processus.

Graphique de Pareto : Définir les options d'un axe

Dans la liste, sous l'en-tête Apparence, cliquez sur **Axes** pour accéder à ces options.

L'orientation des barres détermine si les options de l'axe concernent l'axe vertical principal (Y) ou l'axe horizontal principal (X).

Vous pouvez définir ces options :

- la couleur et la largeur de l'axe.
- le libellé de l'axe principal.
- si un cadre est dessiné autour du graphique de Pareto et de la couleur de remplissage de l'arrière-plan.
- l'échelle pour l'axe principal. Vous avez le choix entre ces options :
 - ☐ **Pourcentage** – L'échelle est un pourcentage de la fréquence totale. Si vous attribuez une variable de poids relatif, l'échelle est un pourcentage du poids total. Il s'agit de la présentation par défaut.
 - ☐ **Effectif** – L'échelle est en unités de fréquence (nombre). Cette option n'est pas applicable si vous attribuez une variable de poids relatif.
 - ☐ **Poids** – L'échelle utilise les mêmes unités que la variable de poids relatif. Cette option n'est applicable que si vous attribuez une variable de poids relatif.

Graphique de Pareto : Ajouter des lignes de référence

Dans la liste, sous l'en-tête Apparence, cliquez sur **Lignes de référence** pour accéder à ces options.

L'orientation des barres détermine si les lignes de référence apparaissent le long de l'axe vertical ou de l'axe horizontal. Pour utiliser les lignes de référence, activez la case à cocher **Utiliser les lignes de référence**.

Graphique de Pareto : Ajouter un encart

Dans la liste, sous l'en-tête Apparence, cliquez sur **Encart** pour accéder à ces options.

Vous pouvez améliorer un graphique de Pareto en insérant dans le graphe une case ou un tableau (appelé encart) contenant les statistiques descriptives. Si vous créez un graphique de Pareto comparatif, un encart apparaît dans chaque graphe.

Remarque : si vous ajoutez un encart au graphique de Pareto, vous devez désactiver la case à cocher **Afficher la légende de la taille d'échantillon** de la page **Apparence > Options**. Si vous insérez à la fois un encart et la légende de la taille de l'échantillon dans le graphique de Pareto, il est possible que les zones d'encart et de légende se chevauchent dans la sortie. Il est donc conseillé de choisir d'insérer un encart ou d'afficher la légende de la taille de l'échantillon, mais pas les deux à la fois.

Cochez la case **Inclure l'encart** pour ajouter les statistiques descriptives au graphique de Pareto. Sélectionnez ensuite les statistiques que vous voulez inclure dans l'encart. Les statistiques disponibles varient selon les options préalablement sélectionnées.

- Par défaut, **Effectif** est la seule statistique apparaissant dans l'encart.
- Si vous avez attribué une variable à la fonction **Poids relatif**, vous pouvez sélectionner **Somme des poids**.
- Si vous avez limité le nombre de catégories, ces statistiques sont également disponibles :
 - ☐ **Observations exclues du graphique restreint** est disponible s'il n'existe pas d'**autre** catégorie représentant les catégories exclues.
 - ☐ **Nombre d'observations dans OTHER=** est disponible s'il y a une **Autre** catégorie représentant les catégories exclues.
 - ☐ **Nombre de catégories OTHER=** est disponible s'il y a une **Autre** catégorie représentant les catégories exclues.

Définir d'autres options de tracé

Dans la liste, sous l'en-tête Apparence, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Légende de la taille de l'échantillon	Activez la case à cocher Afficher la légende de la taille d'échantillon pour afficher une légende de taille d'échantillon. Par défaut, la légende affiche $N=n$, où n est le nombre total de catégories de Pareto. Si vous créez un graphique de Pareto comparatif, une légende apparaît dans chaque graphe. L'orientation des barres détermine la position de la légende sur le graphique. Remarque : si vous ajoutez un encart au graphique de Pareto, vous devez désactiver la case à cocher Afficher la légende de la taille d'échantillon. Si vous insérez à la fois un encart et la légende de la taille de l'échantillon dans le graphique de Pareto, il est possible que les zones d'encart et de légende se chevauchent dans la sortie. Il est donc conseillé de choisir d'insérer un encart ou d'afficher la légende de la taille de l'échantillon, mais pas les deux à la fois.

Nom d'option	Description
	Pour créer un cadre autour de la légende, sélectionnez une couleur dans la liste déroulante Arrière-plan . Cette couleur est utilisée pour remplir le cadre. Pour attribuer un libellé à la légende de l'effectif, saisissez une chaîne de texte de 32 caractères maximum dans le champ Libellé .
Options comparatives	Si vous attribuez au moins une variable de classification à la fonction d'analyse, vous pouvez utiliser les cases Par page pour modifier la disposition des colonnes et des lignes dans le graphique de Pareto comparatif. Si vous attribuez une variable de classification, la disposition par défaut est d'une colonne et de deux lignes par page. Si vous attribuez deux variables de classification, la disposition par défaut est de deux colonnes et deux lignes par page. Vous pouvez également utiliser les listes déroulantes Arrière-plan afin d'indiquer les couleurs utilisées pour remplir le cadre des libellés de lignes et des libellés de colonnes.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Diagramme circulaire

<i>À propos de la tâche Diagramme circulaire</i>	471
<i>Diagramme circulaire : Sélectionner un type de graphique</i>	472
<i>Diagramme circulaire : Attribuer des variables aux fonctions d'analyse</i>	473
<i>Diagramme circulaire : Définir les options d'apparence</i>	474
Définir les options du diagramme circulaire	474
Définir la présentation	475
Définir les options avancées	476
Options d'apparence supplémentaires	476
<i>Indiquer des titres et des notes de bas de page</i>	476
<i>Afficher les propriétés</i>	477

À propos de la tâche Diagramme circulaire

La tâche Diagramme circulaire crée des diagrammes simples, groupés ou empilés qui représentent la contribution relative des parties à l'ensemble en affichant les données sous forme de "secteurs" d'un cercle. Chaque secteur représente une catégorie de données. La taille d'un secteur représente la contribution des données aux statistiques totales du graphique. Par exemple, un diagramme circulaire peut afficher les ventes de chaque magasin sous forme de fraction des ventes totales de la chaîne.

Pour plus d'informations, voir "[Sélectionner un format de sortie pour les graphiques](#)" page 6.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GCHART

Nom du prérequis	Noms des procédures et des produits
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Diagramme circulaire : Sélectionner un type de graphique

Dans la liste, cliquez sur **Diagramme circulaire** pour accéder à ces options.

Nom d'option	Description
Diagramme circulaire simple	crée un graphique qui affiche la contribution relative des parties à l'ensemble. Les données se présentent sous la forme de "secteurs", ou parts, d'un cercle. Chaque secteur représente une catégorie de données. La taille d'un secteur représente la contribution des données aux statistiques totales du graphique.
Diagramme circulaire groupé	produit un diagramme circulaire distinct pour chaque valeur unique de la variable attribuée à la fonction Grouper diagr. circulaires par .
Diagramme circulaire empilé	divise le graphique en cercles concentriques en fonction des valeurs des variables attribuées à la fonction Empiler diagr. circulaires par . La largeur des cercles, qui est identique pour chaque sous-groupe, est déterminée par le rayon du diagramme circulaire.
Diagramme circulaire groupé/empilé	produit un diagramme circulaire distinct pour chaque valeur unique de la variable attribuée à la fonction Grouper diagr. circulaires par . Chaque graphique est divisé en cercles concentriques en fonction des valeurs des variables attribuées à la fonction Empiler diagr. circulaires par .

Diagramme circulaire : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Remarque : les rôles disponibles dépendent du type de graphique sélectionné.

Nom du rôle	Description
Colonne à représenter	Les valeurs de la colonne que vous attribuez à cette fonction déterminent les différents secteurs. La colonne peut être soit alphanumérique soit numérique.
Grouper diagr. circulaires par	Un diagramme circulaire est créé pour chaque valeur unique de la colonne que vous attribuez à cette fonction. Tous les graphiques s'affichent sur la même page.
Empiler diagr. circulaires par	Un cercle concentrique (ou pile) est créé dans le diagramme circulaire pour chaque valeur unique de la colonne que vous attribuez à cette fonction. La colonne est appelée variable de sous-groupe. Remarque : si vous avez sélectionné Java comme format de sortie graphique, les valeurs de la variable de sous-groupe ne sont pas affichées comme cercles concentriques. Au lieu de cela, dans le graphique résultant, vous pouvez détailler un secteur du graphique pour afficher les valeurs du sous-groupe. Néanmoins, si vous avez sélectionné l'image Java comme format de sortie graphique, le graphique est une image statique, et vous perdez la fonctionnalité de détail qui est disponible avec le format de sortie graphique Java.
Somme de	La colonne que vous attribuez à cette fonction détermine la taille des secteurs.

Nom du rôle	Description
Grouper les graphiques par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de graphes créés. Un graphique distinct est créé pour chaque valeur unique. Chaque diagramme circulaire s'affiche sur une nouvelle page.

Diagramme circulaire : Définir les options d'apparence

Définir les options du diagramme circulaire

Dans la liste, sous l'en-tête Apparence, cliquez sur **Diagrammes circulaires** pour accéder à ces options.

Nom d'option	Description
Indiquer les couleurs	indique la couleur des barres. Vous avez le choix entre ces options : ■ Jeu de couleurs par défaut utilise le jeu de couleurs par défaut. Ce dernier est déterminé par le style actuel. Vous pouvez indiquer le style à l'aide du Gestionnaire de styles ou de la boîte de dialogue Propriétés du résultat. Dans SAS Enterprise Guide, vous pouvez également spécifier le style de chaque type de résultat à partir de la boîte de dialogue Options. Dans SAS Add-In for Microsoft Office, vous pouvez appliquer des styles aux formats Rapport SAS et HTML à l'aide de la boîte de dialogue Options de SAS Add-In. ■ Couleurs existantes (Spécifiées dans une version précédente) permet d'utiliser les couleurs qui ont été spécifiées dans une version précédente de SAS Enterprise Guide ou SAS Add-In for Microsoft Office. Cette option n'est disponible que pour le contenu SAS ayant été migré depuis une version antérieure. ■ Couleurs personnalisées permet de spécifier des couleurs personnalisées. En vous servant des listes déroulantes, vous pouvez indiquer jusqu'à 12 couleurs pour un histogramme. Si votre graphique nécessite plus de 12 couleurs, sélectionnez

Nom d'option	Description
	les couleurs supplémentaires dans le jeu de couleurs par défaut actuel. Cliquez sur Réinitialiser pour remplacer les valeurs personnalisées par les valeurs par défaut. Remarque : selon le type de graphique en anneau, il est possible que toutes les couleurs personnalisées ne puissent pas être utilisées.
Spécifier un nombre de secteurs	spécifie le nombre de secteurs dans le graphique. Par défaut, le nombre de secteurs est déterminé automatiquement. Pour les colonnes continues, les valeurs sont divisées par intervalles, à raison d'un secteur par intervalle. Les secteurs qui représenteraient 4 pour cent ou moins du total sont groupés dans un secteur nommé "Autre". Pour indiquer le nombre de secteurs devant apparaître dans le graphique, activez la case à cocher Spécifier un nombre de secteurs.

Définir la présentation

Dans la liste, sous l'en-tête Apparence, cliquez sur **Présentation** pour accéder à ces options.

Indiquez si vous voulez utiliser le nom, le pourcentage, la valeur statistique, ou encore une combinaison des trois dans le libellé de chaque secteur. Pour chacun des composants du libellé, sélectionnez l'une des positions suivantes pour le libellé :

- à l'intérieur du secteur
- à l'extérieur du diagramme circulaire ou de l'anneau
- à l'extérieur du secteur avec une flèche pointant dessus.

Vous pouvez choisir d'utiliser la même couleur de police pour les libellés des secteurs que les secteurs eux-mêmes. Si vous voulez que les libellés apparaissent à l'intérieur des secteurs, n'activez pas cette option.

Vous pouvez indiquer le pourcentage et le libellé pour le secteur "Autre". Ce secteur est l'ensemble de tous les milieux ayant des valeurs statistiques inférieures ou égales au pourcentage du total. La valeur par défaut du pourcentage du total est 4. De ce fait, tout secteur représentant 4% ou moins du total est placé dans le secteur "Autre".

Ce libellé peut être une chaîne de texte comprenant jusqu'à 16 caractères. Le libellé par défaut est Autre. Le secteur "Autre" est le dernier secteur du diagramme, quel que soit l'ordre des secteurs.

Si un seul milieu entre dans la catégorie "Autre", le secteur de ce milieu reste à sa position d'origine sur le graphique et conserve son libellé d'origine. Supposons, par exemple, qu'un diagramme circulaire illustrant le pourcentage de ventes de boissons comporte les secteurs et valeurs suivants : Café 35%, thé 15%, soda 5% et lait 45%. Si le pourcentage du total est 5, le secteur correspondant au soda reste le troisième secteur du graphique plutôt que de devenir le dernier secteur, et il est intitulé "Soda" plutôt que "Autre."

Pour les graphiques groupés, utilisez la zone **Disposition groupée** afin d'organiser les graphiques. Indiquez le nombre de colonnes sur la largeur de la zone de sortie ou le nombre de lignes sur la longueur de la zone de sortie.

Définir les options avancées

Dans la liste, sous l'en-tête Apparence, cliquez sur **Options avancées** pour accéder à ces options.

Sélectionnez dans la liste déroulante la statistique à utiliser pour le calcul des secteurs du graphique. Si vous n'attribuez pas de variable à la fonction **Somme de**, vous pouvez calculer la fréquence ou le pourcentage de chaque secteur. Par défaut, la fréquence est calculée. Si vous n'attribuez pas de variable à la fonction **Somme de**, vous pouvez calculer la somme ou la moyenne de chaque secteur. Par défaut, la somme est calculée.

Options d'apparence supplémentaires

Certaines options sont communes aux tâches graphiques. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir les options de légende" page 733](#)
- ["Personnaliser l'aire" page 734](#)

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

.....

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

.....

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

.....

Assistant Diagramme circulaire

<i>À propos de l'assistant Diagramme circulaire</i>	479
<i>Assistant Diagramme circulaire : Sélectionner les données</i>	479
<i>Assistant Diagramme circulaire : Attribuer des variables aux fonctions</i>	480
<i>Assistant Diagramme circulaire : Définir les options d'apparence</i>	481
<i>Assistant Diagramme circulaire : Indiquer les titres et les notes de bas de page</i>	481

À propos de l'assistant Diagramme circulaire

L'Assistant **Diagramme circulaire** permet de créer des diagrammes circulaires à deux ou à trois dimensions, avec des options pour les diagrammes circulaires groupés ou empilés. Chaque secteur du diagramme représente une catégorie de données. La taille d'un secteur représente la contribution des données aux statistiques totales du graphique.

Assistant Diagramme circulaire : Sélectionner les données

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Cliquez sur **Suivant** pour attribuer des variables aux fonctions.

Assistant Diagramme circulaire : Attribuer des variables aux fonctions

- 1 Avant d'attribuer des variables aux fonctions, vous devez sélectionner les données.
- 2 Attribuez des variables aux fonctions suivantes :
 - **Secteur** - les valeurs de la variable que vous attribuez à cette fonction déterminent les différents secteurs. La variable peut être alphanumérique ou numérique.

Vous pouvez également spécifier ces options :

 - ☐ **Un secteur pour chaque valeur de données unique** - cette option n'a aucune incidence sur les variables alphanumériques. En ce qui concerne les variables numériques, un secteur est tracé pour chaque valeur unique.
 - ☐ **Générer les secteurs automatiquement** - les variables numériques sont traitées comme des variables continues et les valeurs sont divisées en intervalles à raison d'un secteur par intervalle. Le nombre d'intervalles est déterminé automatiquement. Les secteurs qui représenteraient 4 pour cent ou moins du total sont groupés dans un secteur nommé "Autre".
 - ☐ **Taille du secteur** - la variable ou statistique que vous attribuez à cette fonction détermine la taille des secteurs.
 - Si vous attribuez une variable à cette fonction, vous pouvez calculer la somme ou la moyenne de chaque secteur. Par défaut, la somme est calculée.
 - Si vous n'attribuez pas de variable à cette fonction, vous pouvez calculer la fréquence ou le pourcentage de chaque secteur. Par défaut, la fréquence est calculée.
- 3 (Facultatif) Attribuez des variables à ces fonctions :
 - **Grouper par** - les valeurs de la variable que vous attribuez à cette fonction sont organisées en groupes. Un diagramme circulaire distinct est généré pour chaque valeur unique.
 - **Empiler par** - les valeurs de la variable que vous attribuez à cette fonction déterminent le nombre d'anneaux concentriques empilés pour chacun des diagrammes circulaires. Si la sortie graphique est de type Java, la fonctionnalité de détail remplace les cercles concentriques.
 - **Représenter par** - des graphiques distincts sont générés pour chaque groupe BY. Les groupes sont déterminés par les valeurs de la variable que vous attribuez à cette fonction. Une sortie est générée pour chaque groupe BY.

- 4 Cliquez sur **Suivant** pour définir les options d'apparence.

Assistant Diagramme circulaire : Définir les options d'apparence

L'Assistant **Diagramme circulaire** détermine automatiquement les couleurs du graphique. Ce jeu de couleurs est visible dans la liste déroulante **Thèmes**.

Spécifiez ces options d'apparence :

- un graphique à deux dimensions ou trois dimensions
- une légende
- nom de variable sur chaque secteur du graphique
- pourcentage de chaque secteur sur le graphique
- valeur de données pour chaque secteur du graphique

Cliquez sur **Suivant** pour indiquer des titres et des notes de bas de page.

Assistant Diagramme circulaire : Indiquer les titres et les notes de bas de page

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

- 1 Avant d'indiquer les titres et les notes de bas de page, vous devez spécifier les options d'apparence.
- 2 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie. Vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.
- 3 Cliquez sur **Terminer** pour exécuter la tâche.

Courbes P-P

<i>À propos de la tâche Courbes P-P</i>	483
<i>Analyse de capacité : Attribuer des variables aux fonctions d'analyse</i>	484
<i>Courbes P-P : Sélectionner la distribution</i>	486
<i>Courbes P-P : Personnaliser la distribution</i>	487
<i>Analyse de capacité : Définir les options d'apparence du graphe</i>	487
Définir les options d'un axe	487
Ajouter des lignes de référence	488
Ajouter un encart	488
Définir d'autres options de tracé	488
Définir les options des limites de spécification	491
<i>Analyse de capacité : Spécifier les types d'analyse</i>	491
<i>Indiquer des titres et des notes de bas de page</i>	492
<i>Afficher les propriétés</i>	492

À propos de la tâche Courbes P-P

Une analyse de capacité de processus compare la distribution des sorties d'un processus sous contrôle statistique à ses limites de spécification afin de déterminer la cohérence avec laquelle ces spécifications sont respectées. Une courbe probabilité-probabilité (également appelée courbe P-P ou courbe de pourcentage) est utilisée dans une analyse de capacité de processus pour comparer la fonction de distribution cumulative empirique (ECDF) d'une variable à la fonction de distribution cumulative théorique spécifiée comme étant normale. Si les deux distributions coïncident, les points de la courbe forment un motif linéaire qui passe par l'origine et présente une pente unitaire. Vous pouvez ainsi utiliser une courbe P-P pour déterminer avec quelle précision un modèle de distribution théorique ajuste un ensemble de mesures.

Vous pouvez utiliser une courbe P-P pour analyser un processus de fabrication. Supposons par exemple que la distance entre deux trous percés dans 50 feuilles de tôle est mesurée et enregistrée dans une table. Le processus de découpe est sous

contrôle statistique. Vous voulez vérifier que les distances sont normalement distribuées. Vous pouvez utiliser la tâche Courbes P-P pour créer une courbe basée sur la fonction de distribution cumulative normale, qui vous permet de déterminer si les distances sont distribuées normalement ou non.

Remarque : Ne pas confondre les courbes P-P avec les courbes des probabilités, qui comparent un ensemble de mesures ordonnées avec les centiles d'une distribution spécifiée.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	CAPABILITY
Produits SAS requis	SAS Base, SAS/QC, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Analyse de capacité : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables d'analyse	indique les variables numériques faisant l'objet de l'analyse. Vous devez attribuer au moins une variable à cette fonction. Pour toutes les courbes (sauf la courbe P-P, de probabilités théoriques par rapport aux probabilités calculées) et pour chaque variable d'analyse, vous pouvez indiquer une limite supérieure, une valeur cible et une limite inférieure dans la zone Limites de spécification pour <i>nom_variable</i>. Lorsque vous indiquez ces limites, la sortie inclut automatiquement une table

Nom du rôle	Description
	contenant les tests de la normalité, une table contenant les limites de spécification et une table contenant les indices de capacité du processus.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Poids relatif	spécifie la variable à utiliser comme variable de poids relatif. Lorsque vous attribuez une variable à cette fonction, la valeur de la variable de chaque observation est utilisée pour calculer les statistiques pondérées. Vous ne pouvez attribuer qu'une seule variable à cette fonction. Remarque : lorsque vous créez un histogramme, une courbe de probabilité, une courbe P-P ou une courbe CDF, il est déconseillé d'attribuer une variable à la fonction Poids relatif. Si vous attribuez malgré tout une variable à cette fonction, le résultat doit être interprété avec précaution.
Variables de classification (disponible pour la tâche d'histogramme uniquement)	indique une ou deux variables à utiliser en tant que variables de classification. Lorsque vous attribuez une variable à cette fonction, la tâche crée un histogramme comparatif vous permettant de comparer la distribution d'une variable d'analyse sur des niveaux de classification. Remarque : pour créer des courbes de probabilité comparatives ou des courbes Q-Q comparatifs, utilisez la tâche Analyse de distribution.

Courbes P-P : Sélectionner la distribution

Dans la liste, sous l'en-tête Distributions, cliquez sur **Récapitulatif** pour accéder à ces options.

Vous pouvez ajuster une distribution théorique sur une courbe P-P. Après avoir sélectionné la distribution, sélectionnez son nom dans la liste pour en spécifier les valeurs de paramètres et les propriétés de courbe.

Nom d'option	Description
Distributions	Utilisez les cases à cocher pour sélectionner une ou plusieurs de ces courbes de densité : ■ Normal ajuste une densité normale avec une moyenne (μ) et un écart-type (σ). ■ Lognormal ajuste une densité lognormale avec un paramètre d'échelle (ζ), un paramètre de seuil (θ) et un paramètre de forme (σ). ■ Exponentiel ajuste une densité exponentielle avec un paramètre d'échelle (σ) et un paramètre de seuil (θ). ■ Weibull ajuste une densité Weibull à trois paramètres avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et un paramètre de forme (c). ■ Bêta ajuste une densité bêta avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et des paramètres de forme (α et β). ■ Gamma ajuste une densité gamma avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et un paramètre de forme (α).
Style des graphiques	Sélectionnez le style des graphiques à utiliser dans la sortie. Vous avez le choix entre les graphiques SAS traditionnels et les graphiques statistiques ODS. Pour plus d'informations sur les graphiques ODS, voir SAS Output Delivery System: User's Guide.

Courbes P-P : Personnaliser la distribution

Dans la liste, sous l'en-tête Distributions, cliquez sur le nom de la distribution que vous voulez personnaliser pour pouvoir accéder à ces options.

Remarque : ces options de distribution ne sont disponibles que si vous avez sélectionné la distribution au préalable sur la page **Récapitulatif**, ou bien si vous activez la case à cocher **Nom de la distribution** sur la page de cette distribution.

Pour les distributions normales, lognormales, exponentielles, Weibull, bêta et gamma, vous pouvez spécifier les paramètres pour toutes les variables d'analyse ou pour chacune d'entre elles. Par défaut, la tâche estime les valeurs de paramètre de toutes les variables d'analyse.

Pour définir les paramètres de chaque variable d'analyse :

- 1 Désactivez la case à cocher **Appliquer la distribution à toutes les variables**.
- 2 Activez la case à cocher de la variable dans la zone **Variables d'analyse**.
- 3 Spécifiez les valeurs des paramètres.

Remarque : pour revenir aux valeurs des paramètres estimés, cliquez sur **Utiliser les estimations**.

Analyse de capacité : Définir les options d'apparence du graphe

Définir les options d'un axe

Dans la liste, sous l'en-tête **Apparence** ⇒ **Axes**, cliquez sur **Axes** pour accéder à ces options.

Vous pouvez définir ces options :

- la couleur des axes
- la couleur des graduations
- un libellé pour les axes horizontaux et verticaux

Ajouter des lignes de référence

Dans la liste, sous l'en-tête **Apparence** ⇒ **Axes**, cliquez sur **Horizontal** ou **Vertical** pour accéder à ces options.

Pour utiliser les lignes de référence, sélectionnez **Utiliser les lignes de référence**. Vous pouvez formater la ligne de référence et spécifier l'emplacement de chaque ligne de référence.

Ajouter un encart

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Encart** pour accéder à ces options.

Vous pouvez améliorer un graphe en ajoutant une case ou une table (appelée encart) contenant les statistiques descriptives directement sur le graphe. Si vous créez un histogramme comparatif, un encart apparaît dans chaque histogramme.

Sélectionnez **Inclure l'encart** pour ajouter des statistiques descriptives au graphe. Sélectionnez ensuite les statistiques que vous voulez inclure dans l'encart.

Remarque : les statistiques permettant de calculer les indices de capabilité ne sont disponibles que si vous avez créé des limites de spécification pour la variable attribuée à la fonction **Variables d'analyse**. Vous pouvez créer des limites de spécification pour les histogrammes, les courbes CDF, les courbes de probabilité et les diagrammes Q-Q. Pour plus d'informations sur le calcul d'indices de capabilité, consultez l'aide relative à la procédure CAPABILITY.

Définir d'autres options de tracé

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Options** pour accéder à ces options.

Les options d'apparence disponibles dépendent du type de graphe créé par la tâche.

Nom du rôle	Description
Arrière-plan	spécifie la couleur d'arrière-plan du graphe.
Contour de la barre et remplissage de la barre	spécifie les couleurs pour les barres d'un histogramme. Remarque : ces options ne sont disponibles que pour les histogrammes.
Inverser les deux axes en plaçant les centiles sur l'axe opposé	inverse les axes du graphe. Si vous sélectionnez cette option, les options et les lignes de référence que vous avez sélectionnées pour les axes risquent d'être incorrectes en raison du changement d'orientation du graphe. Remarque : cette option n'est disponible que pour les courbes de probabilité et les courbes Q-Q.
Supprimer légendes par défaut	empêche l'affichage de la légende par défaut avec le graphe.
Utiliser des milieux	spécifie les milieux dans l'histogramme. Les options disponibles dans la liste déroulante varient selon que vous créez un histogramme ou un histogramme comparatif. Si vous attribuez une variable à la fonction Variables de classification, vous créez un histogramme comparatif. Pour les histogrammes comparatifs, ces options sont disponibles : ■ Spécifier les milieux permet d'indiquer les milieux de début et de fin. Vous pouvez également spécifier l'intervalle d'escalier à utiliser pour calculer les autres milieux. Le milieu correspond à la valeur située au milieu de la plage de valeurs que représente chaque barre. ■ Clé détermine les milieux des données de la cellule clé. Par défaut, les niveaux de la variable de classification s'affichent de haut en bas (de gauche à droite) dans l'ordre croissant des valeurs internes (non formatées) de la première variable de classification. Si vous n'avez spécifié qu'une seule de ces variables, la cellule clé correspond au niveau qui apparaît le premier dans cet ordre. Si vous en avez spécifié deux, la cellule clé correspond à une combinaison des niveaux des variables 1 et 2 qui apparaissent en premier dans cet ordre. Ainsi, le choix de la cellule clé détermine l'axe horizontal uniforme utilisé pour les autres cellules. La liste des milieux de la cellule clé est donc étendue dans l'une ou l'autre direction si besoin jusqu'à ce qu'elle couvre les données dans les cellules restantes. ■ Uniforme détermine le nombre de milieux en fonction de la taille de l'échantillon complet.

Nom du rôle	Description
	Si vous n'attribuez pas de variable à la fonction Variables de classification, vous créez un histogramme. Les options disponibles sont alors les suivantes : ■ Spécifier les milieux permet d'indiquer les milieux de début et de fin. Vous pouvez également spécifier l'intervalle d'escalier à utiliser pour calculer les autres milieux. Le milieu correspond à la valeur située au milieu de la plage de valeurs que représente chaque barre. ■ Indiquer le nombre de classes permet d'indiquer le nombre de classes (également appelées intervalles d'histogrammes) des données. Par défaut, le nombre de classes affichées dépend de l'étendue des données. Cependant, vous pouvez indiquer un maximum. Remarque : vous ne pouvez indiquer des milieux que si vous utilisez la tâche Histogramme.
Indiquer taille du graphique personnalisé (en pixels)	spécifie la taille de l'image. À titre d'indication, la taille totale de l'écran est affichée dans les zones Largeur et Hauteur .

Remarque : des options de comparaison sont disponibles pour les histogrammes comparatifs. Vous pouvez créer un histogramme comparatif en attribuant une variable à la fonction **Variables de classification** dans la tâche Histogramme.

Nom du rôle	Description
Lignes et Colonnes	Si vous attribuez au moins une variable de classification à la fonction d'analyse, vous pouvez utiliser les cases Lignes par page et Colonnes par page pour modifier la disposition des colonnes et des lignes dans l'histogramme comparatif. Si vous attribuez une variable de classification, la disposition par défaut est de deux lignes et d'une colonne par page. Si vous attribuez deux variables de classification, la disposition par défaut est de deux lignes et de deux colonnes par page. Vous pouvez également utiliser les listes déroulantes Côté du cadre et Haut du cadre afin d'indiquer les couleurs utilisées pour remplir les zones de libellés de ligne et libellés de colonne.
Nombre de classes	Par défaut, le nombre de classes affichées dépend de l'étendue des données. Cependant, vous pouvez indiquer un maximum. Vous pouvez également spécifier un écart-type (par exemple, le nombre de classes affichées), sur la base d'une valeur de

Nom du rôle	Description
	l'écart-type comprise au-dessus et en dessous de la moyenne des données.

Définir les options des limites de spécification

Remarque : ces options sont disponibles pour les histogrammes, les courbes de probabilité, les graphiques Q-Q et les courbes CDF.

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Limites de spécification** pour indiquer la limite de spécification inférieure, la cible, et la limite de spécification supérieure.

Remarque : les options **Limites de spécification** ne sont disponibles que si vous indiquez une ou plusieurs limites pour au moins une variable d'analyse. Pour indiquer une limite de spécification pour une variable d'analyse, cliquez sur **Données** dans la liste, puis sélectionnez la variable d'analyse.

Analyse de capacité : Spécifier les types d'analyse

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Par défaut, le rapport Analyse de capacité se compose des tables suivantes : une table d'intervalles de confiance de base, une table de mesures de base, une table de tests de position et une table de moments.

Pour ajouter une table au rapport, sélectionnez la table dans la zone **Tables**.

Pour spécifier les options d'analyse, sélectionnez la table dans la zone **Tables**. Si les options sont disponibles, une case s'affiche sur la droite avec les options à spécifier.

Pour supprimer toutes les tables de statistiques descriptives du rapport, activez la case à cocher **Supprimer stat. descriptives et indices de capacité**. La sélection de cette option ne supprime pas ces tables :

- tables créées par la distribution sélectionnée. Pour supprimer ces tables, activez la case à cocher **Supprimer les tables de distribution** dans le panneau correspondant à cette distribution.

- Tables créées par l'instruction INTERVALS. Pour supprimer ces tables, vous pouvez spécifier l'option NOPRINT. Pour plus d'informations sur l'instruction INTERVALS, consultez l'aide relative à la procédure CAPABILITY.

Pour enregistrer vos résultats dans une table de sortie, activez la case à cocher **Enregistrer les statistiques de sortie dans une table**. La table inclura les statistiques à une variable et les indices de capabilité.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Préparer les données d'une série chronologique

<i>À propos de la tâche Préparer les données d'une série chronologique</i>	495
<i>Préparer les données d'une série chronologique : Attribuer des variables aux fonctions d'analyse</i>	496
<i>Préparer les données d'une série chronologique : Indiquer la fréquence de sortie</i>	497
<i>Préparer les données d'une série chronologique : Définir les caractéristiques d'une série chronologique</i>	498
<i>Préparer les données d'une série chronologique : Indiquer des transformations pour chaque variable d'une série chronologique</i>	500
<i>Préparer les données d'une série chronologique : Définition des options des résultats</i>	501
<i>Afficher les propriétés</i>	502
<i>Créer des lignes d'observations pour les lignes manquantes de la table</i>	502
<i>Convertir les données d'une série chronologique d'une fréquence d'échantillonnage à une autre</i>	503
<i>Modifier les caractéristiques des observations d'une table d'entrée</i>	504
<i>Convertir des séries apériodiques en estimations périodiques</i>	505
<i>Interpoler les valeurs manquantes</i>	506
<i>Appliquer des transformations génériques aux données</i>	507

À propos de la tâche Préparer les données d'une série chronologique

La tâche Préparer les données d'une série chronologique permet de rendre des données plus exploitables, à des fins d'analyse, par d'autres tâches de série

chronologique. Vous pouvez également vous en servir pour réaliser des transformations génériques sur des données prévues pour être utilisées dans d'autres tâches. Par exemple, vous pouvez appliquer des opérations mathématiques, des fonctions et d'autres transformations aux variables d'une table.

Utilisez cette tâche avec précaution si vous manipulez des séries chronologiques, car les options que vous sélectionnez peuvent affecter la précision des analyses réalisées ultérieurement sur les données.

Vous pouvez utiliser cette tâche pour :

- interpoler des évaluations trimestrielles à partir de séries annuelles ou grouper des valeurs trimestrielles pour produire des séries annuelles
- combiner des séries chronologiques mesurées à différentes fréquences d'échantillonnage. Par exemple, supposons que vous disposiez de données sur les stocks monétaires mensuels, le produit intérieur brut trimestriel et les taux d'intérêt hebdomadaires, et que vous souhaitiez réaliser l'analyse sur un modèle qui utilise toutes ces variables. Pour réaliser l'analyse, vous devez d'abord convertir les séries à une fréquence commune et combiner les variables dans une seule table.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	EXPAND
Produits SAS requis	SAS Base, SAS/ETS
Autres produits SAS recommandés	Aucun

Préparer les données d'une série chronologique : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variable ID temps	indique la variable utilisée pour attribuer des dates aux lignes. Vous devez attribuer une seule variable à cette fonction. La première variable ayant un format date, temps ou date temps SAS est automatiquement attribuée à cette fonction. Vous pouvez aussi créer une variable ID temps.
Variables de la série chronologique	indique les variables que vous voulez préparer pour l'analyse en appliquant les opérations mathématiques disponibles dans cette tâche. Si aucune variable n'est spécifiée, la tâche convertit toutes les variables numériques des données d'entrée en fonction des options que vous avez définies sous l'onglet Caractéristiques.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.

Préparer les données d'une série chronologique : Indiquer la fréquence de sortie

Dans la liste, sous l'en-tête Caractéristiques, cliquez sur **Fréquences** pour accéder à ces options.

- Pour quitter la fréquence de sortie inchangée, sélectionnez **Identique à la fréquence d'entrée**.
- Pour modifier l'intervalle de fréquence, sélectionnez **Intervalle** puis sélectionnez l'intervalle de sortie dans la liste déroulante.

- Pour changer le rapport fréquence d'entrée et fréquence de sortie, sélectionnez **Rapport**. Utilisez les cases **Obs. d'entrée** et **Obs. de sortie** pour indiquer le rapport des observations d'entrée par rapport aux observations de sortie.

Préparer les données d'une série chronologique : Définir les caractéristiques d'une série chronologique

Dans la liste, sous l'en-tête Caractéristiques, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Méthode d'interpolation	spécifie la méthode d'interpolation. Ces méthodes sont disponibles : ■ Spline cubique ajuste une courbe spline cubique aux valeurs d'entrée. Une spline cubique est une fonction segmentée composée de fonctions polynomiales de troisième degré (cubiques) jointes afin que la courbe tout entière et ses première et seconde valeurs dérivées soient continues. ■ L'option Spline linéaire ajuste une courbe continue aux données en reliant les segments de ligne droite successifs. ■ La fonction par paliers ajuste une courbe constante par morceaux discontinue. ■ L'option Agrégation simple effectue une agrégation simple de séries chronologiques sans interpolation des valeurs manquantes. ■ Aucune interpolation signifie qu'aucune interpolation ne sera effectuée.

Remarque : l'option **Aucune interpolation** n'est pas disponible si une conversion de fréquence a été sélectionnée. Pour plus d'informations, voir "[Préparer les données d'une série chronologique : Indiquer la fréquence de sortie](#)" page 497.

Pour la méthode du spline cubique, vous pouvez sélectionner les contraintes sur les points extrêmes inférieur et supérieur. Voici les options de contrainte sur l'extrémité :

- **Pas de division** signifie que les deux premières parties du spline sont contraintes de faire partie de la même courbe

Nom d'option	Description
	cubique, de même que les deux dernières parties. Il s'agit de la présentation par défaut. ■ Spline naturel indique la contrainte de spline naturel. La seconde valeur dérivée de la courbe spline est contrainte d'être égale à zéro au point extrême. ■ Pente indique la première valeur dérivée de la courbe spline au point extrême. ■ L'option Courbure indique la seconde valeur dérivée de la courbe spline au point extrême. Pour les options Pente et Courbure, vous devez spécifier la première valeur dérivée et la seconde valeur dérivée, respectivement. La valeur par défaut des deux valeurs dérivées est 1.
Caractéristiques des observations	spécifie les caractéristiques des observations pour les séries chronologiques d'entrée et de sortie. Une caractéristique d'observation définit le type d'observation pour chaque valeur d'une série chronologique. Cela garantit une bonne réalisation de l'interpolation. Voici les options disponibles : ■ L'option Total indique que les valeurs de données représentent les totaux de l'intervalle de temps correspondant à l'observation. ■ L'option Moyenne indique que les valeurs de données représentent des moyennes de période. ■ Début indique que les valeurs de données sont des valeurs de début de période. Il s'agit de la présentation par défaut. ■ L'option Milieu indique que les valeurs de données sont des valeurs de milieu de période. ■ L'option Fin indique que les valeurs de données sont des valeurs de fin de période. ■ L'option Valeur dérivée indique que les données de sortie doivent être les valeurs dérivées de la courbe spline cubique ajustée aux données d'entrée. Remarque : l'option Dérivée est disponible uniquement à partir de la liste déroulante Sortie. La différence entre l'alignement d'une date et les caractéristiques des observations peut être confuse. L'option d'alignement de date contrôle les valeurs de la date qui sont attribuées à la variable ID temps pour chaque observation. Les caractéristiques des observations contrôlent la façon dont les valeurs de la série chronologique sont mathématiquement interprétées lorsque des interpolations sont effectuées.

Nom d'option	Description
Alignement de la date dans l'intervalle	spécifie l'alignement des dates dans chaque intervalle. Par exemple, si vous avez spécifié un intervalle mensuel, l'option Début démarre l'intervalle de temps au premier jour du mois, l'option Milieu démarre l'intervalle de temps au milieu du mois et l'option Fin le démarre au dernier jour du mois.

Préparer les données d'une série chronologique : Indiquer des transformations pour chaque variable d'une série chronologique

Sous l'en-tête Transformations de la liste, cliquez sur **Avant l'interpolation** ou sur **Après l'interpolation** pour accéder à ces options.

Les transformations uniques peuvent être appliquées à chaque variable attribuée à la fonction **Variables de la série chronologique**.

Nom d'option	Description
Avant interpolation / Après interpolation	Cliquez sur Ajouter pour afficher la liste des transformations disponibles. Cochez les cases des transformations que vous souhaitez ajouter, puis cliquez sur OK. Pour supprimer une transformation, sélectionnez-la puis cliquez sur Supprimer. Les transformations sont effectuées dans l'ordre dans lequel elles apparaissent dans la liste. Vous pouvez changer cet ordre en sélectionnant une transformation puis en utilisant les boutons Monter et Descendre afin de la déplacer dans la liste. Pour appliquer la transformation $y = 5 + \log(x + 10)$ à la série chronologique d'entrée, vous devez saisir les transformations suivantes dans la liste, dans l'ordre affiché : 1 Ajouter une constante (n = 10) 2 log 3 3 Ajouter une constante (n = 5)

Nom d'option	Description
Options pour l'opération sélectionnée	Certaines transformations nécessitent l'indication d'options supplémentaires, qui apparaîtront dans la zone Options de l'opération sélectionnée . Les boutons Exécuter et Enregistrer ne sont disponibles qu'une fois que les paramètres de transformation ont été spécifiés.

Préparer les données d'une série chronologique : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Enregistrer les données de sortie	La table contenant les variables de séries chronologiques converties ou transformées est toujours sauvegardée. Le nom de cette table de sortie apparaît dans la zone Données modifiées. Pour enregistrer de façon permanente les coefficients des courbes spline ayant été ajustés sur la série chronologique d'entrée, activez la case à cocher Résultats estimés des paramètres. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Inclure des graphes de séries dans la sortie	Activez la case à cocher Créer des graphes pour les séries en entrée et en sortie pour inclure dans la sortie des graphes des séries en entrée transformées, des séries en sortie transformées et des séries converties. Des graphes des séries et des séries intermédiaires sont créés pour chaque variable attribuée à la fonction Variable de la série chronologique .

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Créer des lignes d'observations pour les lignes manquantes de la table

La plupart des tâches d'analyse de séries chronologiques supposent que les observations des données d'une série chronologique soient prises à intervalles de temps ou périodes égaux. Si des intervalles de temps sont manquants, les données ne sont pas espacées de façon égale et l'analyse peut échouer ou produire des résultats erronés. Ces données ne sont pas espacées de façon égale :

Ligne	Date	Valeur
-------	------	--------

1	JAN10	100
2	FEB10	110
3	MAR10	120
4	MAY10	140
5	JUN10	150

Ces données ne sont pas espacées de façon égale car l'intervalle entre les observations 3 et 4 n'est pas le même que l'intervalle entre les observations 2 et 3.

Pour créer des observations pour les intervalles manquants :

- 1 Attribuez la variable date à la fonction **Variable ID temps**.
- 2 Définissez l'intervalle pour la fréquence des données. Dans l'exemple, les données sont **Mensuelles**.
- 3 Cliquez sur **Exécuter**.

Une nouvelle table est créée, elle contient toutes les colonnes de la table d'origine. De nouvelles observations sont ajoutées pour les intervalles omis, avec des valeurs mises à manquantes. Ces données peuvent alors être utilisées dans des tâches traitant les valeurs manquantes ou bien la tâche Préparer les données d'une série chronologique peut être utilisée sur la nouvelle table pour interpoler les valeurs manquantes. Pour plus d'informations, voir ["Interpoler les valeurs manquantes" page 506](#).

Vous pouvez également interpoler les valeurs manquantes pour fournir des estimations sur les observations manquantes. Pour plus d'informations, voir ["Convertir des séries apériodiques en estimations périodiques" page 505](#).

Convertir les données d'une série chronologique d'une fréquence d'échantillonnage à une autre

La tâche Préparer les données d'une série chronologique peut s'utiliser pour convertir les données de la série chronologique d'une fréquence d'échantillonnage à une autre. Par exemple, des données hebdomadaires peuvent être agrégées en données mensuelles ou bien des données hebdomadaires peuvent être interpolées à partir de données mensuelles.

Pour convertir la fréquence des données de votre série chronologique :

- 1 Attribuez la variable date à la fonction **Variable ID temps**. Dans la zone **Utiliser la var. ID temps existante**, définissez l'intervalle pour la fréquence des données d'entrée.
- 2 Attribuez les variables de la série chronologique que vous voulez convertir à la fonction **Variables de la série chronologique**.
- 3 Sous l'en-tête Caractéristiques de la liste, cliquez sur **Fréquences** et sélectionnez soit **Rapport** soit **Intervalle** dans la zone **Fréquence de sortie**. Par exemple, sélectionnez **Intervalle** et définissez sa valeur par **Hebdomadaire** afin de convertir les données d'entrée en fréquence hebdomadaire.
- 4 Sous l'en-tête Caractéristiques de la liste, cliquez sur **Options**, puis remplacez la méthode d'interpolation par une méthode appropriée aux données. Par exemple, sélectionnez **Spline cubique** pour ajuster une courbe spline de troisième ordre aux données d'entrée.
- 5 Sélectionnez la caractéristique des observations des données d'entrée et la caractéristique des observations des données de sortie. Par exemple, vous pouvez convertir des données Début de mois en moyennes trimestrielles.
- 6 Sélectionnez l'alignement de date de sortie souhaité.
- 7 Cliquez sur **Exécuter**.

Modifier les caractéristiques des observations d'une table d'entrée

La tâche Préparer les données d'une série chronologique peut s'utiliser pour changer les caractéristiques des observations de la table d'entrée. Par exemple, les totaux mensuels peuvent être convertis en moyennes mensuelles.

Pour changer les caractéristiques des observations des données de votre série chronologique :

- 1 Attribuez la variable date à la fonction **Variable ID temps**. Dans la zone **Utiliser la var. ID temps existante**, définissez l'intervalle pour la fréquence des données d'entrée.
- 2 Attribuez les variables de la série chronologique que vous voulez convertir à la fonction **Variables de la série chronologique**.
- 3 Sous l'en-tête Caractéristiques de la liste, cliquez sur **Fréquences**. Dans la zone **Fréquence de sortie**, gardez la même fréquence de sortie que la fréquence d'entrée, à moins que vous ne changiez également la fréquence des données.
- 4 Sous l'en-tête Caractéristiques de la liste, cliquez sur **Options**, puis remplacez la méthode d'interpolation par une méthode appropriée aux données. Par exemple,

sélectionnez **Spline cubique** pour ajuster une courbe spline de troisième ordre aux données d'entrée.

- 5 Sélectionnez la caractéristique des observations des données d'entrée et la caractéristique des observations des données de sortie. Par exemple, vous pouvez convertir des données Début de mois en moyennes trimestrielles.
- 6 Sélectionnez l'alignement de date de sortie souhaité.
- 7 Cliquez sur **Exécuter**.

Convertir des séries apériodiques en estimations périodiques

La plupart des tâches de l'analyse d'une série chronologique supposent que les observations des données de séries chronologiques soient prises à intervalles de temps ou périodes égaux et qu'aucun intervalle de temps ne soit manquant. Les données apériodiques telles que les données de transaction doivent donc être converties en données périodiques avant d'être analysées à l'aide d'autres tâches pour séries chronologiques.

Pour convertir des données apériodiques en estimations périodiques :

- 1 Attribuez la variable date à la fonction **Variable ID temps**. Dans la zone **Utiliser la var. ID temps existante**, sélectionnez l'option **Données apériodiques**.
- 2 Attribuez les variables de la série chronologique que vous voulez convertir à la fonction **Variables de la série chronologique**.
- 3 Dans le volet de sélection sous l'en-tête Caractéristiques, cliquez sur **Fréquences** et sélectionnez l'intervalle de sortie voulu.
- 4 Sous l'en-tête Caractéristiques de la liste, cliquez sur **Options** et remplacez la méthode d'interpolation par une méthode appropriée aux données. Par exemple, sélectionnez **Spline cubique** pour ajuster une courbe spline de troisième ordre aux données d'entrée.
- 5 Sélectionnez l'alignement de date de sortie souhaité.
- 6 Sélectionnez la caractéristique des observations des données d'entrée et la caractéristique des observations des données de sortie. Par exemple, vous pouvez convertir des valeurs apériodiques en moyennes mensuelles.
- 7 Cliquez sur **Exécuter**.

Interpoler les valeurs manquantes

La tâche Préparer les données de la série chronologique peut s'utiliser pour interpoler les valeurs manquantes de la série chronologique d'entrée. Ces données contiennent des valeurs manquantes :

Ligne	Date	Valeur
1	JAN10	100
2	FEB10	110
3	MAR10	120
4	APR10	.
5	MAY10	140
6	JUN10	150
7	JUL10	.
8	AUG10	150

Pour interpoler les valeurs manquantes des données d'entrée :

- 1 Attribuez la variable date à la fonction **Variable ID temps**. Dans la zone **Utiliser la var. ID temps existante**, définissez l'intervalle pour la fréquence des données d'entrée.
- 2 Attribuez les variables de la série chronologique que vous voulez convertir à la fonction **Variables de la série chronologique**.
- 3 Sous l'en-tête Caractéristiques de la liste, cliquez sur **Fréquences**, puis sélectionnez **Même que fréquence d'entrée** pour la fréquence de la série de sortie.
- 4 Sous l'en-tête Caractéristiques de la liste, cliquez sur **Options** et remplacez la méthode d'interpolation par une méthode appropriée aux données. Par exemple, sélectionnez **Spline cubique** pour ajuster une courbe spline de troisième ordre aux données d'entrée.
- 5 Sélectionnez la caractéristique des observations des données d'entrée et la caractéristique des observations des données de sortie. Par exemple, vous pouvez convertir des données Fin de jour par des moyennes mensuelles.

- 6 Sélectionnez l'alignement de date de sortie souhaité.
- 7 Cliquez sur **Exécuter**.

Appliquer des transformations génériques aux données

La tâche Préparer les données de la série chronologique peut s'utiliser pour réaliser toutes sortes de transformations mathématiques sur la série chronologique d'entrée. Ces transformations peuvent se faire conjointement avec les opérations d'interpolation disponibles, ou peuvent être effectuées sans interpolation.

Pour transformer des données conjointement avec une opération d'interpolation, suivez toutes les étapes de l'option d'interpolation sauf celle de soumission de la tâche. (Pour plus d'informations, voir ["Interpoler les valeurs manquantes" page 506](#).) Utilisez ensuite les pages **Transformations** pour définir les transformations que vous voulez réaliser sur les données avant ou après qu'elles ne soient interpolées. Les deux listes peuvent être vides.

Pour transformer les données sans réaliser d'interpolation :

- 1 Attribuez la variable date à la fonction **Variable ID temps**. Dans la zone **Utiliser la var. ID temps existante**, définissez l'intervalle pour la fréquence des données d'entrée.
- 2 Attribuez les variables de la série chronologique que vous voulez convertir à la fonction **Variables de la série chronologique**.
- 3 Sous l'en-tête Caractéristiques de la liste, cliquez sur **Options** et sélectionnez **Aucune interpolation** comme méthode.
- 4 Sélectionnez l'alignement de date de sortie souhaité.
- 5 Dans le volet de sélection sous l'en-tête Transformations, cliquez sur **Avant l'interpolation** pour insérer les transformations que vous voulez réaliser. Bien que vous laisseriez normalement la liste **Après l'interpolation** vide, vous pouvez aussi définir des transformations sous l'onglet **Après l'interpolation**. Ces transformations seront réalisées après les transformations de la liste **Avant l'interpolation**.
- 6 Cliquez sur **Exécuter**.

Composantes principales

<i>À propos de la tâche Composantes principales</i>	509
<i>Composantes principales : Attribuer des variables aux fonctions d'analyse</i>	510
<i>Composantes principales : Définir les options de l'analyse</i>	511
<i>Composantes principales : Générer des graphes</i>	512
<i>Composantes principales : Définition des options des résultats</i>	513
<i>Indiquer des titres et des notes de bas de page</i>	514
<i>Afficher les propriétés</i>	515

À propos de la tâche Composantes principales

L'analyse en composantes principales est une technique multivariée qui permet d'examiner les relations entre différentes variables quantitatives. Vous pouvez utiliser l'analyse de la composante principale si vous êtes intéressé par l'agrégation des données et par la détection des relations linéaires.

Vous pouvez utiliser la tâche Composantes principales lorsque les variables à représenter simultanément sont trop nombreuses. Par exemple, supposons que vous ayez le taux de criminalité pour 100.000 personnes, dans 7 catégories, pour chacun des 50 États en 1977. Comme il y a sept variables numériques, il est impossible de représenter toutes les variables en même temps. Vous pouvez utiliser l'analyse en composantes principales pour agréger les données en deux ou trois dimensions et en faciliter ainsi la consultation.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	PRINCOMP

Nom du prérequis	Noms des procédures et des produits
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Composantes principales : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables d'analyse	effectue une analyse sur chaque variable de cette liste. Vous devez attribuer au moins une variable à cette fonction.
Variables partielles	spécifie les variables que vous voulez exclure lors de l'analyse d'une matrice de corrélation ou de covariance partielle. La tâche Analyse en composantes principales calcule les composantes principales des résidus issus de la prédiction des variables d'analyse par les variables de cette fonction.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée en tant que variable d'analyse.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.

Nom du rôle	Description
Poids relatif	spécifie les valeurs à utiliser pour pondérer les valeurs correspondantes dans la variable d'analyse. Vous ne pouvez attribuer qu'une seule variable à cette fonction.

Composantes principales : Définir les options de l'analyse

Dans la liste, cliquez sur **Analyse** pour accéder à ces options.

Nom d'option	Description
Analyser	Sélectionnez l'analyse que vous souhaitez effectuer : ■ Corrélations permet de calculer les composantes principales de la matrice de corrélation. ■ Covariances permet de calculer les composantes principales de la matrice de covariance. Si vous définissez cette option, les variables qui présentent des variances importantes sont plus étroitement associées aux composantes comportant des valeurs propres élevées, et les variables qui présentent des variances peu importantes seront plus étroitement associées à des composantes comportant des valeurs propres réduites. Vous ne devez spécifier cette option que si les unités de mesure des variables sont comparables ou si les variables sont standardisées d'une façon ou d'une autre. ■ Corrélations non corrigées permet de calculer les composantes principales de la matrice de corrélation. Toutefois, la constante est omise du modèle. Il est donc possible que la matrice des corrélations et les écarts-types ne soient pas corrigés pour la moyenne. ■ Covariances non corrigées permet de calculer les composantes principales de la matrice de covariance. Toutefois, la constante est omise du modèle. Il est donc possible que la matrice de covariance et les écarts-types ne soient pas corrigés pour la moyenne.
Composantes principales à calculer	Le paramètre par défaut est le nombre de variables. Cliquez sur les flèches allant vers le bas ou le haut pour modifier ce nombre.

Nom d'option	Description
Préfixe pour l'appellation des composantes principales	Indiquez le préfixe des composantes principales. Le nom que vous saisissez peut comporter jusqu'à 26 caractères. Si vous n'indiquez pas de préfixe, la tâche utilise la valeur par défaut. Par défaut, le préfixe du nom des composantes principales est PRIN1, PRIN2, ..., PRIN _n .
Critère de singularité	Spécifiez le critère de singularité où $0 < p < 1$. Si une variable affectée à la fonction Variables partielles a un R^2 aussi grand que $1-p$ lorsqu'il est prédit à partir des variables listées avant dans l'instruction PARTIAL, un coefficient normalisé de 0 est affecté à la variable. La valeur par défaut est 1E-08 .
Diviseur pour variance	Par défaut, le diviseur pour le calcul de la variance est le degré de liberté. Vous pouvez choisir un autre diviseur en faisant une sélection dans la liste déroulante Diviseur pour variance . Remarque : pour la partialisation, vous devez attribuer une variable à la fonction Variables partielles .

Composantes principales : Générer des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Nom d'option	Description
Créer des graphes des valeurs propres et de la variance	Crée un graphe des valeurs propres et un graphe de la proportion de variance. Un graphe des valeurs propres peut vous aider à déterminer le nombre de composantes à utiliser dans l'analyse. Les valeurs propres correspondent à chacune des composantes principales et représentent un partitionnement de l'écart total de l'échantillon. Les valeurs propres sont classées de la plus grande à la plus petite, les premières constituant la plus grande partie de la variation. Le graphe des valeurs propres affiche généralement une rupture entre les premières valeurs propres et les valeurs restantes. Vous pouvez utiliser le nombre de valeurs propres dans cette zone de transition pour déterminer le nombre approprié de composantes à inclure.

Nom d'option	Description
Créer une matrice graphique avec nuages de points des scores des composantes principales	crée une matrice graphique avec nuages de points des scores des composantes principales. L'histogramme de chaque composante apparaît dans l'élément diagonal de la matrice.
Créer un graphe du profil des représentations	Crée un graphe du profil des représentations. Il existe un profil par composante. La valeur sur l'axe Y représente la corrélation entre la variable et la composante principale.
Créer un nuage de points des scores des composantes principales	Crée un nuage de points pour chaque score de composantes principales. Vous pouvez personnaliser ces graphiques en choisissant d'afficher les ellipses de prédiction pour les scores de composantes principales d'une nouvelle observation. Par défaut, aucune ellipse n'apparaît.
Créer des diagrammes des composantes du pattern	Crée les graphes de représentation de la composante appariée. Chaque observation sur le graphe représente la corrélation entre la variable et les deux composantes correspondantes. Vous pouvez choisir de faire la représentation dans un vecteur. Si vous choisissez d'afficher les vecteurs, par défaut un cercle unitaire avec un cercle de variance à 100% est représenté.

Composantes principales : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Créer une table pour les données d'origine et les scores des composantes principales	L'option Créer une table pour les données d'origine et les scores des composantes principales crée une table qui contient vos données d'entrée ainsi que les scores des composantes principales. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

Nom d'option	Description
	SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir .
Créer une table pour les statistiques	L'option Créer une table pour contenir les statistiques crée une table contenant les moyennes, écarts types, nombre d'observations, corrélations ou covariances, valeurs propres et vecteurs propres.
Normaliser les scores des composantes principales	Activez la case à cocher Normaliser scores des composantes principales si vous voulez normaliser les scores des composantes principales pour égaliser la variance. Cette option agit sur la table spécifiée dans la zone Créer une table pour les données d'origine et les scores des composantes principales .
Supprimer le rapport et les graphes	N'affiche pas les résultats.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Courbes de probabilité

<i>À propos de la tâche Courbes de probabilité</i>	517
<i>Analyse de capacité : Attribuer des variables aux fonctions d'analyse</i>	518
<i>Courbes de probabilité : Sélectionner la distribution</i>	520
<i>Courbes de probabilité : Personnaliser la distribution</i>	521
<i>Analyse de capacité : Définir les options d'apparence du graphe</i>	521
Définir les options d'un axe	521
Ajouter des lignes de référence	522
Ajouter un encart	522
Définir d'autres options de tracé	522
Définir les options des limites de spécification	525
<i>Analyse de capacité : Spécifier les types d'analyse</i>	525
<i>Indiquer des titres et des notes de bas de page</i>	526
<i>Afficher les propriétés</i>	526

À propos de la tâche Courbes de probabilité

Une analyse de capacité de processus compare la distribution des sorties d'un processus sous contrôle statistique à ses limites de spécification afin de déterminer la cohérence avec laquelle ces spécifications sont respectées. Une courbe de probabilité est utilisée dans les analyses de capacité de processus pour comparer les valeurs triées d'une variable aux centiles d'une distribution théorique spécifiée comme normale. Si la distribution des données coïncide avec la distribution théorique, les points de la courbe forment un motif linéaire. Vous pouvez donc utiliser la courbe de probabilité pour déterminer avec quelle précision une distribution théorique modélise un ensemble de mesures.

Remarque : les courbes de probabilité sont similaires à des courbes Q-Q. Toutefois, les courbes de probabilité sont préférables pour l'estimation graphique

de centiles, tandis que les courbes Q-Q sont mieux adaptés à l'estimation graphique de paramètres de distribution et d'indices de capabilité.

Vous pouvez utiliser une courbe de probabilité pour analyser un processus de fabrication. Par exemple, supposons que les diamètres de 50 tubes en acier soient mesurés et enregistrés dans une table. Le processus de production des tubes est sous contrôle statistique et vous décidez de vérifier si les diamètres sont normalement distribués. La linéarité du motif de point indique si les mesures sont normalement distribuées.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	CAPABILITY
Produits SAS requis	SAS Base, SAS/QC, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Analyse de capabilité : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables d'analyse	indique les variables numériques faisant l'objet de l'analyse. Vous devez attribuer au moins une variable à cette fonction. Pour toutes les courbes (sauf la courbe P-P, de probabilités théoriques par rapport aux probabilités calculées) et pour chaque variable d'analyse, vous pouvez indiquer une limite supérieure, une valeur cible et une limite inférieure dans la zone Limites de spécification pour <i>nom_variable</i>. Lorsque vous indiquez ces limites, la sortie inclut automatiquement une table

Nom du rôle	Description
	contenant les tests de la normalité, une table contenant les limites de spécification et une table contenant les indices de capacité du processus.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Poids relatif	spécifie la variable à utiliser comme variable de poids relatif. Lorsque vous attribuez une variable à cette fonction, la valeur de la variable de chaque observation est utilisée pour calculer les statistiques pondérées. Vous ne pouvez attribuer qu'une seule variable à cette fonction. Remarque : lorsque vous créez un histogramme, une courbe de probabilité, une courbe P-P ou une courbe CDF, il est déconseillé d'attribuer une variable à la fonction Poids relatif. Si vous attribuez malgré tout une variable à cette fonction, le résultat doit être interprété avec précaution.
Variables de classification (disponible pour la tâche d'histogramme uniquement)	indique une ou deux variables à utiliser en tant que variables de classification. Lorsque vous attribuez une variable à cette fonction, la tâche crée un histogramme comparatif vous permettant de comparer la distribution d'une variable d'analyse sur des niveaux de classification. Remarque : pour créer des courbes de probabilité comparatives ou des courbes Q-Q comparatifs, utilisez la tâche Analyse de distribution.

Courbes de probabilité : Sélectionner la distribution

Dans la liste, sous l'en-tête Distributions, cliquez sur **Récapitulatif** pour accéder à ces options.

Vous pouvez ajuster une distribution théorique sur une courbe de probabilité. Après avoir sélectionné la distribution, sélectionnez son nom dans la liste pour en spécifier les valeurs de paramètres et les propriétés de courbe.

Nom d'option	Description
Distributions	Utilisez les cases à cocher pour sélectionner une ou plusieurs de ces courbes de densité : ■ Normal ajuste une densité normale avec une moyenne (μ) et un écart-type (σ). ■ Lognormal ajuste une densité lognormale avec un paramètre d'échelle (z), un paramètre de seuil (θ) et un paramètre de forme (σ). ■ Exponentiel ajuste une densité exponentielle avec un paramètre d'échelle (σ) et un paramètre de seuil (θ). ■ Bêta ajuste une densité bêta avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et des paramètres de forme (α et β). ■ Gamma ajuste une densité gamma avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et un paramètre de forme (α). ■ Weibull (3 paramètres) ajuste une densité Weibull à trois paramètres avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et un paramètre de forme (c). ■ Weibull (2 paramètres) ajuste une densité de Weibull à deux paramètres avec un paramètre d'échelle (σ) et un paramètre de forme (c). Dans la densité de Weibull à deux paramètres, vos données sont supposées contenir le paramètre de seuil inférieur connu $\theta=0$.
Style des graphiques	Sélectionnez le style des graphiques à utiliser dans la sortie. Vous avez le choix entre les graphiques SAS traditionnels et les graphiques statistiques ODS. Pour plus d'informations sur les graphiques ODS, voir SAS Output Delivery System: User's Guide.

Courbes de probabilité : Personnaliser la distribution

Dans la liste, sous l'en-tête Distributions, cliquez sur le nom de la distribution que vous voulez personnaliser pour pouvoir accéder à ces options.

Remarque : ces options de distribution ne sont disponibles que si vous avez sélectionné la distribution au préalable sur la page **Récapitulatif**, ou bien si vous activez la case à cocher **Nom de la distribution** sur la page de cette distribution.

Pour les distributions normales, lognormales, exponentielles, Weibull, bêta et gamma, vous pouvez spécifier les paramètres pour toutes les variables d'analyse ou pour chacune d'entre elles. Par défaut, la tâche estime les valeurs de paramètre de toutes les variables d'analyse.

Pour définir les paramètres de chaque variable d'analyse :

- 1 Désactivez la case à cocher **Appliquer la distribution à toutes les variables**.
- 2 Activez la case à cocher de la variable dans la zone **Variables d'analyse**.
- 3 Spécifiez les valeurs des paramètres.

Remarque : pour revenir aux valeurs des paramètres estimés, cliquez sur **Utiliser les estimations**.

Analyse de capacité : Définir les options d'apparence du graphe

Définir les options d'un axe

Dans la liste, sous l'en-tête **Apparence** ⇒ **Axes**, cliquez sur **Axes** pour accéder à ces options.

Vous pouvez définir ces options :

- la couleur des axes
- la couleur des graduations
- un libellé pour les axes horizontaux et verticaux

Ajouter des lignes de référence

Dans la liste, sous l'en-tête **Apparence** ⇒ **Axes**, cliquez sur **Horizontal** ou **Vertical** pour accéder à ces options.

Pour utiliser les lignes de référence, sélectionnez **Utiliser les lignes de référence**. Vous pouvez formater la ligne de référence et spécifier l'emplacement de chaque ligne de référence.

Ajouter un encart

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Encart** pour accéder à ces options.

Vous pouvez améliorer un graphe en ajoutant une case ou une table (appelée encart) contenant les statistiques descriptives directement sur le graphe. Si vous créez un histogramme comparatif, un encart apparaît dans chaque histogramme.

Sélectionnez **Inclure l'encart** pour ajouter des statistiques descriptives au graphe. Sélectionnez ensuite les statistiques que vous voulez inclure dans l'encart.

.....
Remarque : les statistiques permettant de calculer les indices de capabilité ne sont disponibles que si vous avez créé des limites de spécification pour la variable attribuée à la fonction **Variables d'analyse**. Vous pouvez créer des limites de spécification pour les histogrammes, les courbes CDF, les courbes de probabilité et les diagrammes Q-Q. Pour plus d'informations sur le calcul d'indices de capabilité, consultez l'aide relative à la procédure CAPABILITY.
.....

Définir d'autres options de tracé

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Options** pour accéder à ces options.

Les options d'apparence disponibles dépendent du type de graphe créé par la tâche.

Nom du rôle	Description
Arrière-plan	spécifie la couleur d'arrière-plan du graphe.
Contour de la barre et remplissage de la barre	spécifie les couleurs pour les barres d'un histogramme. Remarque : ces options ne sont disponibles que pour les histogrammes.
Inverser les deux axes en plaçant les centiles sur l'axe opposé	inverse les axes du graphe. Si vous sélectionnez cette option, les options et les lignes de référence que vous avez sélectionnées pour les axes risquent d'être incorrectes en raison du changement d'orientation du graphe. Remarque : cette option n'est disponible que pour les courbes de probabilité et les courbes Q-Q.
Supprimer légendes par défaut	empêche l'affichage de la légende par défaut avec le graphe.
Utiliser des milieux	spécifie les milieux dans l'histogramme. Les options disponibles dans la liste déroulante varient selon que vous créez un histogramme ou un histogramme comparatif. Si vous attribuez une variable à la fonction Variables de classification, vous créez un histogramme comparatif. Pour les histogrammes comparatifs, ces options sont disponibles : ■ Spécifier les milieux permet d'indiquer les milieux de début et de fin. Vous pouvez également spécifier l'intervalle d'escalier à utiliser pour calculer les autres milieux. Le milieu correspond à la valeur située au milieu de la plage de valeurs que représente chaque barre. ■ Clé détermine les milieux des données de la cellule clé. Par défaut, les niveaux de la variable de classification s'affichent de haut en bas (de gauche à droite) dans l'ordre croissant des valeurs internes (non formatées) de la première variable de classification. Si vous n'avez spécifié qu'une seule de ces variables, la cellule clé correspond au niveau qui apparaît le premier dans cet ordre. Si vous en avez spécifié deux, la cellule clé correspond à une combinaison des niveaux des variables 1 et 2 qui apparaissent en premier dans cet ordre. Ainsi, le choix de la cellule clé détermine l'axe horizontal uniforme utilisé pour les autres cellules. La liste des milieux de la cellule clé est donc étendue dans l'une ou l'autre direction si besoin jusqu'à ce qu'elle couvre les données dans les cellules restantes. ■ Uniforme détermine le nombre de milieux en fonction de la taille de l'échantillon complet.

Nom du rôle	Description
	Si vous n'attribuez pas de variable à la fonction Variables de classification, vous créez un histogramme. Les options disponibles sont alors les suivantes : ■ Spécifier les milieux permet d'indiquer les milieux de début et de fin. Vous pouvez également spécifier l'intervalle d'escalier à utiliser pour calculer les autres milieux. Le milieu correspond à la valeur située au milieu de la plage de valeurs que représente chaque barre. ■ Indiquer le nombre de classes permet d'indiquer le nombre de classes (également appelées intervalles d'histogrammes) des données. Par défaut, le nombre de classes affichées dépend de l'étendue des données. Cependant, vous pouvez indiquer un maximum. Remarque : vous ne pouvez indiquer des milieux que si vous utilisez la tâche Histogramme.
Indiquer taille du graphique personnalisé (en pixels)	spécifie la taille de l'image. À titre d'indication, la taille totale de l'écran est affichée dans les zones Largeur et Hauteur .

Remarque : des options de comparaison sont disponibles pour les histogrammes comparatifs. Vous pouvez créer un histogramme comparatif en attribuant une variable à la fonction **Variables de classification** dans la tâche Histogramme.

Nom du rôle	Description
Lignes et Colonnes	Si vous attribuez au moins une variable de classification à la fonction d'analyse, vous pouvez utiliser les cases Lignes par page et Colonnes par page pour modifier la disposition des colonnes et des lignes dans l'histogramme comparatif. Si vous attribuez une variable de classification, la disposition par défaut est de deux lignes et d'une colonne par page. Si vous attribuez deux variables de classification, la disposition par défaut est de deux lignes et de deux colonnes par page. Vous pouvez également utiliser les listes déroulantes Côté du cadre et Haut du cadre afin d'indiquer les couleurs utilisées pour remplir les zones de libellés de ligne et libellés de colonne.
Nombre de classes	Par défaut, le nombre de classes affichées dépend de l'étendue des données. Cependant, vous pouvez indiquer un maximum. Vous pouvez également spécifier un écart-type (par exemple, le nombre de classes affichées), sur la base d'une valeur de

Nom du rôle	Description
	l'écart-type comprise au-dessus et en dessous de la moyenne des données.

Définir les options des limites de spécification

Remarque : ces options sont disponibles pour les histogrammes, les courbes de probabilité, les graphiques Q-Q et les courbes CDF.

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Limites de spécification** pour indiquer la limite de spécification inférieure, la cible, et la limite de spécification supérieure.

Remarque : les options **Limites de spécification** ne sont disponibles que si vous indiquez une ou plusieurs limites pour au moins une variable d'analyse. Pour indiquer une limite de spécification pour une variable d'analyse, cliquez sur **Données** dans la liste, puis sélectionnez la variable d'analyse.

Analyse de capacité : Spécifier les types d'analyse

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Par défaut, le rapport Analyse de capacité se compose des tables suivantes : une table d'intervalles de confiance de base, une table de mesures de base, une table de tests de position et une table de moments.

Pour ajouter une table au rapport, sélectionnez la table dans la zone **Tables**.

Pour spécifier les options d'analyse, sélectionnez la table dans la zone **Tables**. Si les options sont disponibles, une case s'affiche sur la droite avec les options à spécifier.

Pour supprimer toutes les tables de statistiques descriptives du rapport, activez la case à cocher **Supprimer stat. descriptives et indices de capacité**. La sélection de cette option ne supprime pas ces tables :

- tables créées par la distribution sélectionnée. Pour supprimer ces tables, activez la case à cocher **Supprimer les tables de distribution** dans le panneau correspondant à cette distribution.

- Tables créées par l'instruction INTERVALS. Pour supprimer ces tables, vous pouvez spécifier l'option NOPRINT. Pour plus d'informations sur l'instruction INTERVALS, consultez l'aide relative à la procédure CAPABILITY.

Pour enregistrer vos résultats dans une table de sortie, activez la case à cocher **Enregistrer les statistiques de sortie dans une table**. La table inclura les statistiques à une variable et les indices de capabilité.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Hasards proportionnels

<i>À propos de la tâche Hasards proportionnels</i>	529
<i>Hasards proportionnels : Attribuer des variables aux fonctions d'analyse</i>	530
<i>Hasards proportionnels : Définir les options du modèle</i>	532
<i>Hasards proportionnels : Définir les options de méthode</i>	534
<i>Hasards proportionnels : Générer des graphes</i>	535
<i>Hasards proportionnels : Définition des options des résultats</i>	536
<i>Indiquer des titres et des notes de bas de page</i>	537
<i>Afficher les propriétés</i>	538

À propos de la tâche Hasards proportionnels

La tâche Hasards proportionnels réalise une analyse de régression des données de survie en s'appuyant sur le modèle de hasard proportionnel de Cox. Ce modèle semi-paramétrique est largement utilisé dans l'analyse des données de survie pour expliquer les effets des variables explicatives sur les périodes de survie.

L'analyse des données de survie est fréquemment utilisée dans les études de suivi clinique et épidémiologique. D'autres domaines utilisent des analyses de survie, en particulier la sociologie, l'ingénierie et l'économie. Quel que soit le domaine, l'objectif commun d'une analyse de survie consiste à déterminer si un événement s'est produit et à quel moment il s'est produit. Par exemple, les sujets décédés cinq ans après une intervention chirurgicale sont différents de ceux qui sont décédés un mois après. Une analyse qui ne compterait que les délais après une intervention ignorerait des informations importantes sur la durée de survie. Les analyses de survie peuvent également être utilisées pour analyser des résultats autres que la durée. Par exemple, un ingénieur peut déterminer l'usure des pneus en fonction du nombre de kilomètres ou la réparation d'un moteur en fonction de nombre de tours.

Le point commun entre ces études est l'analyse d'un résultat jusqu'à l'apparition d'un événement.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	PHREG
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Hasards proportionnels : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Temps de survie	indique la variable à utiliser en tant que variable de temps de survie. Le temps de survie est requis et vous pouvez attribuer une seule variable à cette fonction.
Variable de coupure	indique quelle variable utiliser comme variable de coupure. Sélectionnez une valeur dans la zone Valeurs de coupure pour créer une liste d'une ou plusieurs valeurs utilisées pour couper le temps de survie. Utilisez la case à cocher Saisir valeurs personnalisées pour saisir une valeur spécifique et cliquez sur Ajout pour l'ajouter à la liste des valeurs de coupure. Pour supprimer une valeur dans la liste des valeurs de coupure, désactivez la case à cocher face à cette valeur.

Nom du rôle	Description
	Les valeurs de coupure doivent être des valeurs numériques non manquantes. La variable de coupure n'est pas obligatoire. Vous ne pouvez attribuer qu'une seule variable à cette fonction.
Variables explicatives	indique les variables explicatives pour le modèle.
Variables de niveau (de discrétisation)	indique les variables dont les valeurs déterminent l'amplitude des niveaux de discrétisation. Si vous n'indiquez pas d'extrémités pour les variables numériques, les niveaux de discrétisation seront déterminés par la combinaison des niveaux des valeurs uniques des variables de niveau. Pour autoriser les valeurs manquantes à former des amplitudes de niveau valides, activez la case à cocher. Sinon, les observations avec valeurs manquantes sont supprimées de l'analyse. Pour les variables de niveau numérique, vous pouvez indiquer les intervalles de niveaux de discrétisation en cliquant sur Indiquer des intervalles dans la zone Amplitudes de niveau. Définissez ensuite les intervalles en ajoutant des points extrêmes. Par défaut, les valeurs uniques forment les amplitudes de niveau de discrétisation. Si vous attribuez plusieurs variables à cette fonction, définissez les intervalles séparément pour chaque variable.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Variables ID	désigne une ou plusieurs variables à utiliser comme variables ID dans la table de sortie contenant les estimations de survie. Pour créer la table de sortie, indiquez sous la page Résultats que vous voulez enregistrer les estimations de survie.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.

Hasards proportionnels : Définir les options du modèle

Dans la liste, cliquez sur **Modèle** pour accéder à ces options.

Remarque : ces options ne sont disponibles que si vous avez attribué une variable à la fonction **Variables explicatives**.

Nom d'option	Description
Définir la méthode de sélection du modèle	Sélectionnez une méthode dans la liste déroulante Méthode de sélection pour déterminer les variables explicatives qui figureront dans le modèle. Si vous ne choisissez pas de méthode, le modèle entier est utilisé. Modèle complet (aucune sélection) Il s'agit de la méthode par défaut qui ne propose aucune option de sélection de modèle. Toutes les variables explicatives sont incluses dans le modèle. Sélection ascendante La méthode de sélection ascendante commence par estimer les paramètres des variables qui sont imposées au modèle. Ces variables sont les ordonnées à l'origine ainsi que les n premières variables explicatives apparaissant dans la liste Variables à inclure. La valeur par défaut de n est zéro de sorte qu'aucune variable ne soit incluse dans le modèle. Cette méthode calcule ensuite les chi-deux ajustés pour chaque variable ne se trouvant pas dans le modèle et examine les plus élevées de ces statistiques. Les p-values de ces statistiques sont comparées au niveau de significativité spécifié pour l'entrée d'une variable dans le modèle. La valeur par défaut est 0,05. Pour changer le niveau de significativité, saisissez une valeur dans le champ Pour entrer dans le modèle. Si aucun chi-deux n'a un niveau de significativité supérieur à cette valeur, la sélection ascendante s'arrête. Sinon, la méthode de sélection ascendante ajoute dans le modèle la variable ayant le chi-deux ajusté le plus élevé. La méthode de sélection ascendante recalcule ensuite le chi-deux ajusté pour les variables ne faisant pas partie du modèle et le processus d'évaluation se poursuit. Ainsi, les variables sont ajoutées d'une manière séquentielle au modèle jusqu'à ce qu'aucune variable restante ne produise test du chi-deux ajusté significatif. Une fois qu'une variable a été ajoutée au modèle, elle y reste. Élimination descendante La méthode de sélection descendante commence par estimer les paramètres pour tout le modèle, puis élimine les variables non

Nom d'option	Description
	significatives à moins que des variables explicatives ne figurent dans la liste Variables à inclure. Les résultats du test de Wald pour les paramètres individuels sont examinés. La variable la moins significative et qui ne correspond pas au niveau de significativité spécifié dans la zone Pour rester dans le modèle est supprimée. (Par défaut, cette valeur est 0.10). Une fois qu'une variable a été supprimée d'un modèle, elle reste exclue. Le processus est répété jusqu'à ce qu'aucune autre variable du modèle ne réponde au niveau spécifié pour la suppression. Sélection pas à pas La méthode pas à pas est similaire à la technique de sélection ascendante, à ceci près que les variables figurant déjà dans le modèle ne sont pas nécessairement conservées. Les variables sont ajoutées et supprimées dans le modèle de telle sorte que chaque étape de la sélection ascendante puisse être suivie d'une ou plusieurs étapes de sélection descendante. Le processus de sélection pas à pas se termine s'il n'a plus de variable à ajouter dans le modèle ou si la variable qui vient d'être ajoutée dans le modèle est la seule variable qui est supprimée dans l'étape suivante de sélection descendante. Meilleure sélection de sous-ensemble La méthode de meilleure sélection de sous-groupe classe les modèles par statistique de score de probabilité (Chi-deux). Elle identifie un nombre donné de modèles ayant la statistique du chi-deux avec le plus grand score pour toutes les tailles de modèles possibles (allant d'une variable explicative jusqu'à un modèle unique avec toutes les variables explicatives). S'il n'y a pas plus de 10 variables explicatives, tous les modèles possibles sont répertoriés pour chaque taille de modèle. S'il y a plus de 10 variables explicatives, le nombre de modèles sélectionnés pour chaque taille de modèle est, au maximum, égal au nombre total de variables explicatives. Les plus petits modèles contiennent n variables explicatives, n allant du nombre de variables explicatives dans la liste Variables à inclure jusqu'au nombre total de variables explicatives dans le modèle. Si la liste Variables à inclure est vide, la valeur de n est 1. Remarque : les paramètres de l'intercept restent toujours dans le modèle.
Spécifier les niveaux de significativité	Pour la méthode de sélection ascendante ou pas à pas, le niveau de significativité que vous indiquez dans la zone Pour entrer dans le modèle détermine si une variable explicative doit être incluse dans le modèle. Les variables qui ne sont pas incluses dans le modèle sont contrôlées et la variable ayant la plus petite p-value est incluse si la p-value est inférieure ou égale au niveau de significativité indiqué. Le niveau de significativité par défaut est 0,05. Pour la méthode de sélection descendante ou pas à pas, le niveau de significativité que vous indiquez dans la zone Pour rester dans le modèle détermine si une variable explicative doit être conservée dans le modèle. Les variables du modèle sont contrôlées et la variable avec la plus grande p-value est

Nom d'option	Description
	supprimée si la p -value dépasse le niveau de significativité indiqué. Le niveau de significativité par défaut est 0,05.
Sélectionner les variables à garder dans le modèle	La case Variables à inclure répertorie les variables explicatives qui sont incluses ou supprimées dans le modèle durant le processus de sélection des variables. Cochez les cases correspondant aux variables que vous souhaitez toujours inclure dans le modèle. Si vous ne souhaitez plus inclure systématiquement une variable donnée dans le modèle, désélectionnez la case à cocher affichée en regard du nom de cette variable.

Hasards proportionnels : Définir les options de méthode

Dans la liste, cliquez sur **Méthodes** pour accéder à ces options.

Nom d'option	Description
Calculer l'intervalle de confiance du rapport de hasard	affiche l'intervalle de confiance pour chaque variable explicative. Vous pouvez soit saisir une valeur pour le niveau de confiance soit en sélectionner une dans la liste Niveau de confiance . La valeur par défaut est 95 pour cent.
Corrélations des résultats estimés des paramètres	affiche les matrices de corrélations estimées pour les paramètres estimés.
Covariances des résultats estimés des paramètres	affiche la matrice de covariance estimée des paramètres estimés.
Méthode en présence d'ex-aequos	spécifie le lissage des valeurs manquantes. Vous avez le choix entre ces options : ■ Probabilité approximative de Breslow utilise la méthode de probabilité approximative de Breslow. Si les valeurs manquantes ne sont pas nombreuses, cette méthode offre des approximations satisfaisantes pour calculer les probabilités conditionnelles exactes pour le modèle d'échelle de temps continue. Il s'agit de la méthode par défaut.

Nom d'option	Description
	■ Modèle logistique discret remplace le modèle à risques proportionnels par le modèle logistique discret. Si l'échelle de temps est réellement discrète, vous devez utiliser cette méthode. Si l'analyse des études cas-témoins présente plusieurs cas dans un ensemble correspondant, cette méthode est obligatoire. ■ Probabilité approximative d'Efron utilise la méthode de probabilité approximative d'Efron. Si les valeurs manquantes ne sont pas nombreuses, cette méthode offre des approximations satisfaisantes pour calculer les probabilités conditionnelles exactes pour le modèle d'échelle de temps continue. ■ Calculer probabilités conditionnelles exactes calcule la probabilité conditionnelle exacte dans le cadre de la supposition des hasards proportionnels selon laquelle tous les événements liés se produisent avant les heures coupées de même valeur ou avant les valeurs plus élevées. Remarque : cette méthode peut nécessiter une quantité considérable de ressources informatiques. S'il n'y a pas de liens, les quatre méthodes donnent la même probabilité et obtiennent les mêmes estimations.

Hasards proportionnels : Générer des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Remarque : avant de créer un graphique d'évaluation du modèle, vous devez attribuer une variable à la fonction **Variables explicatives**.

Nom d'option	Description
Graphique du cumul des résidus d'une loi de martingale	Pour chaque variable explicative, cette option trace les résidus cumulés observés d'une loi de martingale par rapport aux valeurs de la variable explicative, ainsi que 20 simulations de structures résiduelles. Si vous sélectionnez l'option Panel graphique du cumul des résidus d'une loi de martingale , vous pouvez créer un graphique comportant quatre panels. Chaque

Nom d'option	Description
	panel contient les résidus cumulés observés d'une loi de martingale et deux simulations de structures résiduelles.
Graphique (d'un processus) de scores normalisés	Pour chaque variable explicative, cette option représente la composante d'un processus de score observée par rapport au temps de suivi, ainsi que 20 simulations de représentations.
Graphique de la fonction de survie	Cette option représente les fonctions de survie estimées pour un jeu de référence des variables explicatives. Ce jeu est constitué des moyennes d'échantillon correspondantes.
Graphique de la fonction de risque cumulée	Cette option représente la fonction de hasard cumulée estimée pour un jeu de référence des variables explicatives. Ce jeu est constitué des moyennes d'échantillon correspondantes. S'il n'y a pas de liens, les quatre méthodes donnent la même probabilité et obtiennent les mêmes estimations.

Hasards proportionnels : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Enregistrer les données de sortie	enregistre certains ou tous vos résultats dans une table de sortie. Cochez les cases des statistiques que vous voulez inclure dans la sortie. ■ L'option Estimations de survie crée une table de sortie contenant les estimations de survie, les diagnostics, ainsi que les estimateurs linéaires. ■ L'option Estimations du coefficient de régression crée une table de sortie contenant les estimations du coefficient de régression. Si vous voulez inclure la matrice de covariance dans la table, activez la case à cocher Inclure matrice de covariance estimée. ■ L'option Estimations de la fonction de survie crée une table de sortie contenant la fonction de survie qui correspond aux moyennes des variables explicatives pour chaque niveau. La table inclut également les estimations de la fonction de survie et l'estimateur linéaire, les erreurs type, les intervalles de

Nom d'option	Description
	confiance, le log. de survie et le log. du log. négatif de la survie. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Supprimer l'affichage des résultats et des graphes	N'affiche pas les résultats. Si vous activez cette case à cocher, les options du volet Graphes ne sont pas disponibles. Par défaut, les résultats sont affichés.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Courbes Q-Q

<i>À propos de la tâche Courbes Q-Q</i>	539
<i>Analyse de capacité : Attribuer des variables aux fonctions d'analyse</i>	540
<i>Courbes Q-Q : Sélectionner la distribution</i>	542
<i>Courbes Q-Q : Personnaliser la distribution</i>	543
<i>Analyse de capacité : Définir les options d'apparence du graphe</i>	543
Définir les options d'un axe	543
Ajouter des lignes de référence	544
Ajouter un encart	544
Définir d'autres options de tracé	544
Définir les options des limites de spécification	547
<i>Analyse de capacité : Spécifier les types d'analyse</i>	547
<i>Indiquer des titres et des notes de bas de page</i>	548
<i>Afficher les propriétés</i>	548

À propos de la tâche Courbes Q-Q

Une analyse de capacité de processus compare la distribution des sorties d'un processus sous contrôle statistique à ses limites de spécification afin de déterminer la cohérence avec laquelle ces spécifications sont respectées. Une courbe quantile-quantile (courbe Q-Q) est utilisée dans les analyses de capacité de processus pour comparer les valeurs triées d'une variable aux quantiles d'une distribution théorique spécifiée comme normale. Si la distribution des données coïncide avec la distribution théorique, les points de la courbe forment un motif linéaire. Ainsi, vous pouvez utiliser une courbe Q-Q pour déterminer avec quelle précision une distribution théorique modélise un ensemble de mesures.

Remarque : les courbes Q-Q sont similaires aux courbes de probabilité. Les courbes Q-Q sont préférables pour réaliser une évaluation graphique de paramètres

de distribution et d'indices de capabilité, alors que les courbes de probabilité sont mieux adaptées à une estimation graphique des centiles.

Vous pouvez utiliser une courbe Q-Q pour analyser un processus de fabrication. Supposons par exemple que la distance entre deux trous percés dans 50 feuilles de tôle est mesurée et enregistrée dans une table. Le processus de découpe est sous contrôle statistique et vous décidez de vérifier si la distribution du processus est normale. La courbe Q-Q compare les observations de la table avec les quantiles de la distribution normale. La linéarité du motif de point indique si les mesures sont normalement distribuées.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	CAPABILITY
Produits SAS requis	SAS Base, SAS/QC, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Analyse de capabilité : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables d'analyse	indique les variables numériques faisant l'objet de l'analyse. Vous devez attribuer au moins une variable à cette fonction. Pour toutes les courbes (sauf la courbe P-P, de probabilités théoriques par rapport aux probabilités calculées) et pour chaque variable d'analyse, vous pouvez indiquer une limite supérieure, une valeur cible et une limite inférieure dans la zone Limites de spécification pour <i>nom_variable</i>. Lorsque vous

Nom du rôle	Description
	indiquez ces limites, la sortie inclut automatiquement une table contenant les tests de la normalité, une table contenant les limites de spécification et une table contenant les indices de capacité du processus.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Poids relatif	spécifie la variable à utiliser comme variable de poids relatif. Lorsque vous attribuez une variable à cette fonction, la valeur de la variable de chaque observation est utilisée pour calculer les statistiques pondérées. Vous ne pouvez attribuer qu'une seule variable à cette fonction. Remarque : lorsque vous créez un histogramme, une courbe de probabilité, une courbe P-P ou une courbe CDF, il est déconseillé d'attribuer une variable à la fonction Poids relatif. Si vous attribuez malgré tout une variable à cette fonction, le résultat doit être interprété avec précaution.
Variables de classification (disponible pour la tâche d'histogramme uniquement)	indique une ou deux variables à utiliser en tant que variables de classification. Lorsque vous attribuez une variable à cette fonction, la tâche crée un histogramme comparatif vous permettant de comparer la distribution d'une variable d'analyse sur des niveaux de classification. Remarque : pour créer des courbes de probabilité comparatives ou des courbes Q-Q comparatifs, utilisez la tâche Analyse de distribution.

Courbes Q-Q : Sélectionner la distribution

Dans la liste, sous l'en-tête Distributions, cliquez sur **Récapitulatif** pour accéder à ces options.

Vous pouvez ajuster une distribution théorique sur un courbe Q-Q. Après avoir sélectionné la distribution, sélectionnez son nom dans la liste pour en spécifier les valeurs de paramètres et les propriétés de courbe.

Nom d'option	Description
Distributions	Utilisez les cases à cocher pour sélectionner une ou plusieurs de ces courbes de densité : ■ Normal ajuste une densité normale avec une moyenne (μ) et un écart-type (σ). ■ Lognormal ajuste une densité lognormale avec un paramètre d'échelle (ζ), un paramètre de seuil (θ) et un paramètre de forme (σ). ■ Exponentiel ajuste une densité exponentielle avec un paramètre d'échelle (σ) et un paramètre de seuil (θ). ■ Bêta ajuste une densité bêta avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et des paramètres de forme (α et β). ■ Gamma ajuste une densité gamma avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et un paramètre de forme (α). ■ Weibull (3 paramètres) ajuste une densité Weibull à trois paramètres avec un paramètre d'échelle (σ), un paramètre de seuil (θ) et un paramètre de forme (c). ■ Weibull (2 paramètres) ajuste une densité de Weibull à deux paramètres avec un paramètre d'échelle (σ) et un paramètre de forme (c). Dans la densité de Weibull à deux paramètres, vos données sont supposées contenir le paramètre de seuil inférieur connu $\theta=0$.
Style des graphiques	Sélectionnez le style des graphiques à utiliser dans la sortie. Vous avez le choix entre les graphiques SAS traditionnels et les graphiques statistiques ODS. Pour plus d'informations sur les graphiques ODS, voir SAS Output Delivery System: User's Guide.

Courbes Q-Q : Personnaliser la distribution

Dans la liste, sous l'en-tête Distributions, cliquez sur le nom de la distribution que vous voulez personnaliser pour pouvoir accéder à ces options.

Remarque : ces options de distribution ne sont disponibles que si vous avez sélectionné la distribution au préalable sur la page **Récapitulatif**, ou bien si vous activez la case à cocher **Nom de la distribution** sur la page de cette distribution.

Pour les distributions normales, lognormales, exponentielles, Weibull, bêta et gamma, vous pouvez spécifier les paramètres pour toutes les variables d'analyse ou pour chacune d'entre elles. Par défaut, la tâche estime les valeurs de paramètre de toutes les variables d'analyse.

Pour définir les paramètres de chaque variable d'analyse :

- 1 Désactivez la case à cocher **Appliquer la distribution à toutes les variables**.
- 2 Activez la case à cocher de la variable dans la zone **Variables d'analyse**.
- 3 Spécifiez les valeurs des paramètres.

Remarque : pour revenir aux valeurs des paramètres estimés, cliquez sur **Utiliser les estimations**.

Analyse de capacité : Définir les options d'apparence du graphe

Définir les options d'un axe

Dans la liste, sous l'en-tête **Apparence** ⇒ **Axes**, cliquez sur **Axes** pour accéder à ces options.

Vous pouvez définir ces options :

- la couleur des axes
- la couleur des graduations
- un libellé pour les axes horizontaux et verticaux

Ajouter des lignes de référence

Dans la liste, sous l'en-tête **Apparence** ⇒ **Axes**, cliquez sur **Horizontal** ou **Vertical** pour accéder à ces options.

Pour utiliser les lignes de référence, sélectionnez **Utiliser les lignes de référence**. Vous pouvez formater la ligne de référence et spécifier l'emplacement de chaque ligne de référence.

Ajouter un encart

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Encart** pour accéder à ces options.

Vous pouvez améliorer un graphe en ajoutant une case ou une table (appelée encart) contenant les statistiques descriptives directement sur le graphe. Si vous créez un histogramme comparatif, un encart apparaît dans chaque histogramme.

Sélectionnez **Inclure l'encart** pour ajouter des statistiques descriptives au graphe. Sélectionnez ensuite les statistiques que vous voulez inclure dans l'encart.

Remarque : les statistiques permettant de calculer les indices de capabilité ne sont disponibles que si vous avez créé des limites de spécification pour la variable attribuée à la fonction **Variables d'analyse**. Vous pouvez créer des limites de spécification pour les histogrammes, les courbes CDF, les courbes de probabilité et les diagrammes Q-Q. Pour plus d'informations sur le calcul d'indices de capabilité, consultez l'aide relative à la procédure CAPABILITY.

Définir d'autres options de tracé

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Options** pour accéder à ces options.

Les options d'apparence disponibles dépendent du type de graphe créé par la tâche.

Nom du rôle	Description
Arrière-plan	spécifie la couleur d'arrière-plan du graphe.
Contour de la barre et remplissage de la barre	spécifie les couleurs pour les barres d'un histogramme. Remarque : ces options ne sont disponibles que pour les histogrammes.
Inverser les deux axes en plaçant les centiles sur l'axe opposé	inverse les axes du graphe. Si vous sélectionnez cette option, les options et les lignes de référence que vous avez sélectionnées pour les axes risquent d'être incorrectes en raison du changement d'orientation du graphe. Remarque : cette option n'est disponible que pour les courbes de probabilité et les courbes Q-Q.
Supprimer légendes par défaut	empêche l'affichage de la légende par défaut avec le graphe.
Utiliser des milieux	spécifie les milieux dans l'histogramme. Les options disponibles dans la liste déroulante varient selon que vous créez un histogramme ou un histogramme comparatif. Si vous attribuez une variable à la fonction Variables de classification, vous créez un histogramme comparatif. Pour les histogrammes comparatifs, ces options sont disponibles : ■ Spécifier les milieux permet d'indiquer les milieux de début et de fin. Vous pouvez également spécifier l'intervalle d'escalier à utiliser pour calculer les autres milieux. Le milieu correspond à la valeur située au milieu de la plage de valeurs que représente chaque barre. ■ Clé détermine les milieux des données de la cellule clé. Par défaut, les niveaux de la variable de classification s'affichent de haut en bas (de gauche à droite) dans l'ordre croissant des valeurs internes (non formatées) de la première variable de classification. Si vous n'avez spécifié qu'une seule de ces variables, la cellule clé correspond au niveau qui apparaît le premier dans cet ordre. Si vous en avez spécifié deux, la cellule clé correspond à une combinaison des niveaux des variables 1 et 2 qui apparaissent en premier dans cet ordre. Ainsi, le choix de la cellule clé détermine l'axe horizontal uniforme utilisé pour les autres cellules. La liste des milieux de la cellule clé est donc étendue dans l'une ou l'autre direction si besoin jusqu'à ce qu'elle couvre les données dans les cellules restantes. ■ Uniforme détermine le nombre de milieux en fonction de la taille de l'échantillon complet.

Nom du rôle	Description
	Si vous n'attribuez pas de variable à la fonction Variables de classification, vous créez un histogramme. Les options disponibles sont alors les suivantes : ■ Spécifier les milieux permet d'indiquer les milieux de début et de fin. Vous pouvez également spécifier l'intervalle d'escalier à utiliser pour calculer les autres milieux. Le milieu correspond à la valeur située au milieu de la plage de valeurs que représente chaque barre. ■ Indiquer le nombre de classes permet d'indiquer le nombre de classes (également appelées intervalles d'histogrammes) des données. Par défaut, le nombre de classes affichées dépend de l'étendue des données. Cependant, vous pouvez indiquer un maximum. Remarque : vous ne pouvez indiquer des milieux que si vous utilisez la tâche Histogramme.
Indiquer taille du graphique personnalisé (en pixels)	spécifie la taille de l'image. À titre d'indication, la taille totale de l'écran est affichée dans les zones Largeur et Hauteur .

Remarque : des options de comparaison sont disponibles pour les histogrammes comparatifs. Vous pouvez créer un histogramme comparatif en attribuant une variable à la fonction **Variables de classification** dans la tâche Histogramme.

Nom du rôle	Description
Lignes et Colonnes	Si vous attribuez au moins une variable de classification à la fonction d'analyse, vous pouvez utiliser les cases Lignes par page et Colonnes par page pour modifier la disposition des colonnes et des lignes dans l'histogramme comparatif. Si vous attribuez une variable de classification, la disposition par défaut est de deux lignes et d'une colonne par page. Si vous attribuez deux variables de classification, la disposition par défaut est de deux lignes et de deux colonnes par page. Vous pouvez également utiliser les listes déroulantes Côté du cadre et Haut du cadre afin d'indiquer les couleurs utilisées pour remplir les zones de libellés de ligne et libellés de colonne.
Nombre de classes	Par défaut, le nombre de classes affichées dépend de l'étendue des données. Cependant, vous pouvez indiquer un maximum. Vous pouvez également spécifier un écart-type (par exemple, le nombre de classes affichées), sur la base d'une valeur de

Nom du rôle	Description
	l'écart-type comprise au-dessus et en dessous de la moyenne des données.

Définir les options des limites de spécification

Remarque : ces options sont disponibles pour les histogrammes, les courbes de probabilité, les graphiques Q-Q et les courbes CDF.

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Limites de spécification** pour indiquer la limite de spécification inférieure, la cible, et la limite de spécification supérieure.

Remarque : les options **Limites de spécification** ne sont disponibles que si vous indiquez une ou plusieurs limites pour au moins une variable d'analyse. Pour indiquer une limite de spécification pour une variable d'analyse, cliquez sur **Données** dans la liste, puis sélectionnez la variable d'analyse.

Analyse de capacité : Spécifier les types d'analyse

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Par défaut, le rapport Analyse de capacité se compose des tables suivantes : une table d'intervalles de confiance de base, une table de mesures de base, une table de tests de position et une table de moments.

Pour ajouter une table au rapport, sélectionnez la table dans la zone **Tables**.

Pour spécifier les options d'analyse, sélectionnez la table dans la zone **Tables**. Si les options sont disponibles, une case s'affiche sur la droite avec les options à spécifier.

Pour supprimer toutes les tables de statistiques descriptives du rapport, activez la case à cocher **Supprimer stat. descriptives et indices de capacité**. La sélection de cette option ne supprime pas ces tables :

- tables créées par la distribution sélectionnée. Pour supprimer ces tables, activez la case à cocher **Supprimer les tables de distribution** dans le panneau correspondant à cette distribution.

- Tables créées par l'instruction INTERVALS. Pour supprimer ces tables, vous pouvez spécifier l'option NOPRINT. Pour plus d'informations sur l'instruction INTERVALS, consultez l'aide relative à la procédure CAPABILITY.

Pour enregistrer vos résultats dans une table de sortie, activez la case à cocher **Enregistrer les statistiques de sortie dans une table**. La table inclura les statistiques à une variable et les indices de capabilité.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Carte radar

<i>À propos de la tâche Carte radar</i>	551
<i>Carte radar : Attribuer des variables aux fonctions d'analyse</i>	552
<i>Carte radar : Définir les options d'apparence</i>	553
Définir les options de la carte Radar	553
Options d'apparence supplémentaires	553
<i>Indiquer des titres et des notes de bas de page</i>	554
<i>Afficher les propriétés</i>	554

À propos de la tâche Carte radar

La tâche Radar crée des graphiques radars (ou étoile) affichant la fréquence relative des mesures de données dans les problèmes de contrôle qualité ou d'étude de marché. Les statistiques s'affichent le long des rayons qui émanent du centre du graphique.

Par exemple, une carte radar peut présenter les ventes totales de trois sites de fabrication. Les ventes d'un site sont représentées par la longueur de la branche correspondante.

Vous ne pouvez pas créer une Carte radar si vous avez sélectionné les formats de sortie Applet Java ou Image Java.

Pour plus d'informations, voir "[Sélectionner un format de sortie pour les graphiques](#)" page 6.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GRADAR
Produits SAS requis	SAS Base, SAS/GRAPH

Nom du prérequis	Noms des procédures et des produits
Autres produits SAS recommandés	Néant

Carte radar : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Colonne à représenter	Les valeurs de la colonne que vous attribuez à cette fonction déterminent les rayons du graphique. La colonne doit avoir au moins trois observations car il faut trois points pour définir un plan.
Grouper les graphiques par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de graphes créés. Un graphique distinct est créé pour chaque valeur unique.
Nombre d'occurrences	Les valeurs de la colonne que vous attribuez à cette fonction indiquent la fréquence (nombre d'occurrences) des valeurs de la fonction Colonnes à représenter . La colonne doit être numérique. Si vous omettez cette fonction, chaque valeur de la fonction Colonnes à représenter est comptée une fois.
Superposition	Les valeurs de la colonne que vous attribuez à cette fonction génèrent un graphique radar comparatif en utilisant les niveaux de la colonne spécifiée. Tous les graphiques sont affichés dans le même ensemble de rayons. Vous ne pouvez pas utiliser cette fonction si vous attribuez des colonnes à la fonction Horizontale ou Verticale .
Horiz.	La colonne que vous attribuez à cette fonction génère un radar pour chaque valeur de la colonne spécifiée. Les

Nom du rôle	Description
	graphiques sont tracés de gauche à droite. Si vous attribuez aussi une colonne à la fonction Verticale , les graphiques sont tracés de gauche à droite et de haut en bas.
En bas	La colonne que vous attribuez à cette fonction génère un radar pour chaque valeur de la colonne spécifiée. Les graphiques sont tracés de haut en bas. Si vous attribuez aussi une colonne à la fonction Horizontale , les graphiques sont tracés de gauche à droite et de haut en bas.

Carte radar : Définir les options d'apparence

Définir les options de la carte Radar

Dans la liste, sous l'en-tête Apparence, cliquez sur **Radar** pour accéder à ces options.

Sélectionnez le style de la carte.

- L'option **Couronne** crée un polygone avec des sommets en étoile émanant du cercle intérieur.
- L'option **Polygone** crée un polygone fermé.
- L'option **Radial** crée des rayons qui émanent du centre.
- L'option **Rayon** crée des rayons qui émanent du cercle intérieur.
- L'option **Part** crée un polygone fermé avec des rayons émanant du centre jusqu'à la longueur complète du rayon (valeur par défaut).

Options d'apparence supplémentaires

Certaines options sont communes aux tâches graphiques. Pour plus d'informations sur ces options, voir "[Personnaliser l'aire](#)" page 734.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Échantillon aléatoire

<i>À propos de la tâche Échantillon aléatoire</i>	557
<i>Échantillon aléatoire : Créer un échantillon aléatoire</i>	558

À propos de la tâche Échantillon aléatoire

La tâche Échantillon aléatoire crée une table de sortie contenant un échantillon aléatoire des lignes de la table d'entrée.

Vous pouvez utiliser cette tâche si vous avez besoin d'un sous-ensemble des données. Par exemple, supposons que vous vouliez contrôler les frais de déplacement de votre personnel dans le but d'en améliorer la procédure de décompte et de réduire éventuellement les dépenses. Comme vous ne disposez pas des ressources nécessaires pour examiner toutes les notes de frais, vous voulez utiliser un échantillon statistique pour sélectionner objectivement celles à contrôler.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	SURVEYSELECT
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	Néant

Échantillon aléatoire : Créer un échantillon aléatoire

- 1 Spécifiez un nom pour la tâche.

Remarque : Cette option est disponible dans SAS Enterprise Guide uniquement.

- 2 Sélectionnez la source de données d'entrée pour la tâche.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

- 3 Spécifie les variables à inclure dans la table de sortie. Par défaut, toutes les variables sont incluses dans la table de sortie.

Remarque : lorsque vous attribuez une variable à la fonction **Variables de niveau (de discrétisation)**, celle-ci est aussi attribuée automatiquement à la fonction **Variables de sortie**. Si vous supprimez une variable dans la sortie, elle est également supprimée en tant que variable de niveau.

- 4 Spécifiez la taille d'échantillon dans le nombre voulu de lignes ou dans le pourcentage voulu des lignes d'entrée. Par exemple, si vous spécifiez 3% de lignes et qu'il y a 400 lignes d'entrée, l'échantillon aura 12 lignes.

Remarque : si vous attribuez des variables à la fonction **Variables de niveau (de discrétisation)**, la taille d'échantillon indiquée ici s'applique à chaque niveau plutôt qu'à la table d'entrée dans sa totalité.

- 5 Sélectionner une méthode d'échantillonnage. Vous avez le choix entre ces options :

Simple (pas de valeurs en double)

Lorsqu'une ligne est sélectionnée, elle n'est plus disponible pour une sélection ultérieure. Cela rend impossible la sélection multiple d'une même ligne.

Non limité (valeurs en double autorisées)

Lorsqu'une ligne est sélectionnée, elle reste disponible pour une sélection ultérieure. Cela rend possible la sélection multiple d'une même ligne. Vous pouvez spécifier comment plusieurs sélections d'une même ligne seront enregistrées dans la table de sortie. Vous avez le choix entre les options suivantes :

- **Afficher chaque observation une fois dans la sortie (exclure les valeurs en double)** - une ligne qui est sélectionnée n fois apparaîtra une fois dans l'échantillon. Dans la sortie, la variable NumberHits (qui est calculée automatiquement par la tâche Échantillon aléatoire) répertorie le nombre d'occurrences de l'observation dans la table d'entrée.
 - **Afficher toutes les observations dans la sortie (inclure les valeurs en double)** - une ligne qui est sélectionnée n fois apparaîtra n fois dans l'échantillon.
- 6 Spécifie les variables à utiliser pour partitionner la table d'entrée en sous-groupes mutuellement exclusifs, qui ne se chevauchent pas et qui sont connus en tant que niveaux. Chaque niveau est défini par un ensemble de valeurs des variables de niveau (de discrétisation), et est échantillonné séparément. L'échantillon complet est l'union des échantillons venant de tous les niveaux.

Remarque : si vous n'attribuez pas de variables à cette fonction, la table d'entrée entière est traitée comme un seul niveau.

Vous pouvez allouer la taille d'échantillon totale parmi les niveaux proportionnellement à la taille du niveau. Par exemple, la variable GENDER a les valeurs possibles de M et F, et la variable VOTED a les valeurs possibles de Y et N. Si vous attribuez GENDER et VOTED à la fonction **Variable de niveau (de discrétisation)**, la table d'entrée sera divisée en quatre niveaux : les hommes qui ont voté, les hommes qui n'ont pas voté, les femmes qui ont voté et les femmes qui n'ont pas voté.

La table d'entrée contient 20 000 lignes. Voici la distribution des données :

- 7 000 hommes qui ont voté
- 4 000 hommes qui n'ont pas voté
- 5 000 femmes qui ont voté
- 4 000 femmes qui n'ont pas voté

Le proportion d'hommes qui ont voté est donc de $7,000/20,000=0.35$ ou 35%. Les proportions dans l'échantillon doivent refléter les proportions des niveaux dans la table d'entrée. Par exemple, si la table d'exemple contient 100 observations, 35% des valeurs de l'échantillon doivent être sélectionnées dans le niveau des hommes qui ont voté pour refléter les proportions de la table d'entrée.

Si vous cochez **S'il n'y a pas assez d'observations pour remplir un niveau, les sélectionner toutes**, la tâche sélectionne toutes les observations d'un niveau lorsque sa taille est inférieure à la taille d'échantillon indiquée.

- 7 Indique le nom et l'emplacement de la table de sortie. Par défaut, la table est enregistrée dans la bibliothèque Work.
- 8 Indique la valeur initiale pour la génération de nombres aléatoires. Si vous n'indiquez pas de nombre de valeurs initiales aléatoires, une valeur initiale basée sur l'horloge système sera utilisée pour créer l'échantillon.

- 9 Indique s'il faut générer une table de synthèse qui inclue la valeur initiale utilisée pour créer l'échantillon. En indiquant cette même valeur initiale plus tard avec la même table d'entrée, vous pouvez reproduire le même échantillon.

Classer les données

<i>À propos de la tâche Classer les données</i>	561
<i>Classer les données : Attribuer des variables aux fonctions d'analyse</i>	562
<i>Classer les données : Définir les options de l'analyse</i>	562
<i>Rang : Définition des options des résultats</i>	564
<i>Afficher les propriétés</i>	565

À propos de la tâche Classer les données

La tâche Classer les données calcule le rang d'une ou de plusieurs variables numériques à travers les observations d'une table SAS et renvoie les rangs comme sortie vers une nouvelle table SAS.

Par exemple, vous pouvez classer les ventes de chaque produit commercialisé par votre entreprise. Dans ce cas, la variable de classement pourrait être l'ordre des ventes des produits. Le produit présentant le chiffre de ventes le plus élevé sera alors classé en premier.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	RANK
Produits SAS requis	Base SAS
Autres produits SAS recommandés	SAS/STAT

Classer les données : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Colonnes à classer	Chaque colonne attribuée à cette fonction sera classée. Vous devez attribuer au moins une variable à cette fonction. Par défaut, la colonne des classements prend le nom <code>rang_nom_colonne</code> , où <code>nom_colonne</code> est le nom de la colonne d'origine. Pour spécifier un nouveau nom, saisissez-le dans le champ de saisie Nom de la colonne de rang .
Classer par	Si vous attribuez une ou plusieurs colonnes à cette fonction, la table d'entrée sera triée par la ou les colonnes sélectionnées et les rangs seront calculés à l'intérieur de chaque groupe.

Classer les données : Définir les options de l'analyse

Dans la liste, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Méthode de classement	spécifie la méthode de classement. Vous avez le choix entre ces options :

Nom d'option	Description
	■ L'option Du plus petit au plus grand classe les valeurs d'origine de la plus petite à la plus grande. ■ Un centile divise les valeurs d'origine en 100 groupes dans lesquels les valeurs les plus faibles reçoivent une valeur de centile de 0 et les valeurs les plus fortes une valeur de centile de 99. ■ L'option Déciles divise les valeurs d'origine en 10 groupes, dans lesquels les valeurs les plus faibles reçoivent une valeur de décile de 0 et les valeurs les plus fortes une valeur de décile de 9. ■ L'option Quartiles divise les valeurs d'origine en 4 groupes, dans lesquels les valeurs les plus faibles reçoivent une valeur de quartile de 0 et les valeurs les plus fortes une valeur de quartile de 3. ■ Group = n (NTILES) divise les valeurs d'origine en n groupes, dans lesquels les valeurs les plus faibles reçoivent la valeur 0 et les valeurs les plus élevées, la valeur $n-1$. Indiquez la valeur de n dans le champ Nombre de groupes. ■ L'option Rangs distribués sur n fractions calcule le rang en divisant chaque rang par le nombre d'observations qui comportent des valeurs non-manquantes de la variable de classement. ■ L'option Rangs distribués sur n+1 fractions calcule le rang en divisant chaque rang par le dénominateur $n+1$, où n est le nombre d'observations qui comportent des valeurs non-manquantes de la variable de classement. ■ L'option Pourcentages divise chaque rang par le nombre d'observations comportant des valeurs non-manquantes de la variable et multiplie le résultat par 100 pour obtenir un pourcentage. ■ L'option Estimations normales (Formule de Blom) calcule les estimations normales des rangs selon la formule de Blom. ■ L'option Estimations normales (formule de Tukey) calcule les estimations normales des rangs selon la formule de Tukey. ■ Estimation normale (formule de van de Waerden) calcule l'estimation normale à partir des rangs, en utilisant la formule de van der Waerden. ■ L'option Calcul des scores selon la formule de Savage (exponentiel) calcule les scores des rangs selon la formule de Savage (exponentiel).
Si valeurs liées, utiliser	spécifie la méthode de résolution des valeurs liées. Si vous ne faites pas de sélection, la tâche utilise la méthode Moyenne (Midrank).

Nom d'option	Description
	■ Moyenne (Midrank) nécessite que les valeurs liées reçoivent la moyenne du rang correspondant. ■ Rang élevé nécessite que les rangs correspondants les plus élevés soient utilisés. ■ Rang faible nécessite que les plus faibles des rangs correspondants soient utilisés.
Inverser la méthode pour aller du plus grand au plus petit	indique que l'ordre du classement doit être inversé.

Rang : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Par défaut, la table de sortie contient les colonnes d'origine ainsi que les colonnes classées. Si vous souhaitez remplacer la colonne d'origine par les colonnes classées, désactivez la case à cocher **Inclure les valeurs de classement**.

Le nom de la colonne des classements est spécifié lorsque la colonne d'origine est attribuée à la fonction **Colonnes à classer**. Par défaut, la colonne des classements prend le nom `rang_nom_colonne`, où `nom_colonne` est le nom de la colonne d'origine. Pour modifier le nom de la colonne, cliquez sur **Données** dans la liste.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Analyse RFM (Récence, Fréquence et Montant)

<i>À propos de la tâche Analyse RFM (Récence, Fréquence et Montant)</i>	567
<i>Analyse RFM (Récence, Fréquence et Montant) : Sélectionner le type des données à analyser</i>	568
<i>Analyse RFM (Récence, Fréquence et Montant) : Attribuer des variables aux fonctions d'analyse</i>	569
<i>Analyse RFM (Récence, Fréquence et Montant) : Regrouper les données dans des classes</i>	570
<i>Analyse RFM (Récence, Fréquence et Montant) : Enregistrer les résultats</i>	571
<i>Analyse RFM (Récence, Fréquence et Montant) : Sélectionner des graphes</i>	572

À propos de la tâche Analyse RFM (Récence, Fréquence et Montant)

La tâche Analyse RFM (Récence, Fréquence et Montant) est une technique utilisée pour identifier les clients existants qui sont le plus susceptibles de répondre à une nouvelle campagne ou offre de produit. Cette technique est communément utilisée dans le marketing direct.

Les données en entrée de cette tâche peuvent être des données transactionnelles ou données du client. La tâche Analyse RFM (Récence, Fréquence et Montant) agrège les données transactionnelles dans les données du client. Dans les données transactionnelles, chaque enregistrement représente une transaction.

Voici comment se déterminent les scores de récence, de fréquence et monétaires :

- Le score de récence est déterminé par le tri, par ordre croissant, des valeurs de la date de transaction la plus récente, puis par le regroupement de ces valeurs

dans des classes. La classe avec les dates les plus anciennes se voit attribuer un score de récence de 1, etc.

- Le score de fréquence est déterminé par le tri, par ordre croissant, des valeurs du nombre de transactions, puis par le regroupement de ces valeurs dans des classes. La classe avec le plus petit nombre de transactions se voit attribuer un score de fréquence de 1, etc.
- Le score monétaire est déterminé par le tri, par ordre croissant, des valeurs du montant total de la transaction puis par le regroupement de ces valeurs dans des classes. La classe avec le plus petit montant se voit attribuer un score monétaire de 1, etc.

Le score RFM est calculé à l'aide de la formule suivante : $\text{Score RFM} = 100 * \text{score de récence} + 10 * \text{score de fréquence} + \text{score monétaire}$. Le segment client le moins favorable a le score RFM le plus faible. Le segment client le plus favorable a le score RFM le plus élevé.

Remarque : la tâche Analyse RFM (Récence, Fréquence et Montant) requiert SAS 9.4 ou version ultérieure.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	
Produits SAS requis	Base SAS
Autres produits SAS recommandés	

Analyse RFM (Récence, Fréquence et Montant) : Sélectionner le type des données à analyser

Dans la liste, cliquez sur **Type RFM** pour accéder à ces options.

Nom d'option	Description
Données au niveau des transactions	Chaque enregistrement des données représente une transaction. La tâche Analyse RFM (Récence, Fréquence et Montant) agrège les transactions individuelles par ID client.

Nom d'option	Description
Données du client	Chaque enregistrement représente un client. Les transactions individuelles sont combinées avant l'ouverture des données dans la tâche Analyse RFM (Récence, Fréquence et Montant).

Analyse RFM (Récence, Fréquence et Montant) : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Les fonctions d'analyse disponibles varient selon le type de données que vous voulez analyser.

Tableau 72.1 Options pour les données au niveau transaction

Nom du rôle	Description
Date de la transaction	la variable qui spécifie la date de la transaction. Ces informations déterminent la date de transaction la plus récente.
Montant de la transaction	la variable montant qui spécifie le montant de la transaction. Ces données sont utilisées pour calculer le montant total des transactions, qui peut être la somme, la moyenne, la médiane ou le maximum du montant des transactions.
ID client	la variable qui contient un nombre ou une chaîne pouvant identifier un client de manière unique.

Du fait que la tâche RFM (Récence, Fréquence et Montant) agrège les données transactionnelles dans les données du client, vous devez sélectionner une méthode

d'agrégation pour les transactions. Par défaut, la méthode d'agrégation est la Somme.

Vous pouvez également indiquer s'il faut inclure dans l'agrégation les clients ayant des enregistrements de transactions incomplets (ou avec valeurs manquantes).

- Si vous sélectionnez cette option, la table de sortie contient les scores des clients ayant des enregistrements de transactions incomplets. Pour chaque client, seuls les enregistrements de transaction ayant des valeurs non manquantes pour la date de transaction et le montant de la transaction sont agrégés. Les enregistrements avec valeurs manquantes sont ignorés.
- Si vous ne sélectionnez pas cette option, la table de sortie ne contient pas les scores des clients ayant des enregistrements de transactions incomplets. Les enregistrements de transactions ne sont agrégés que pour les clients ayant des enregistrements complets.

Tableau 72.2 Options pour les données du client

Nom du rôle	Description
Transaction la plus récente	la variable qui spécifie la date de la transaction la plus récente de chaque client avec la société. Ces données sont utilisées pour calculer le score de récence.
Nombre de transactions	la variable qui spécifie le nombre de transactions pour chaque client. Ces données sont utilisées pour calculer le score de fréquence.
Montant des transactions	la variable qui spécifie le total du montant monétaire dépensé par chaque client. Ces données sont utilisées pour calculer le score monétaire.
ID client	la variable qui contient un nombre ou une chaîne pouvant identifier un client de manière unique.

Analyse RFM (Récence, Fréquence et Montant) : Regrouper les données dans des classes

Dans la liste, cliquez sur **Discretisation** pour accéder à ces options.

Nom d'option	Description
Méthodes de discrétisation	spécifie la méthode de catégorisation des données. Vous avez le choix entre ces options : ■ La méthode Imbriqué affecte un simple rang à chaque valeur de récence. Dans chaque rang de récence, un rang de fréquence est affecté aux clients ; et dans chaque rang de fréquence, un rang monétaire est affecté aux clients. L'avantage de cette méthode est qu'elle rend une distribution équilibrée des scores RFM combinés et prend en considération que ce comportement d'achat (par exemple, de fréquence et monétaire) peut changer dans le temps. L'inconvénient de cette méthode est qu'il prend plus de temps pour interpréter les scores de fréquences et monétaires car tous les scores dépendent les uns des autres. ■ La méthode Indépendant classe indépendamment les valeurs de récence, fréquence et monétaires. L'avantage de cette méthode est que l'interprétation des scores est sans ambiguïté. Par exemple, un score de fréquence de 5 pour un client a la même signification qu'un score de fréquence de 5 pour un autre client. Pour de plus petits échantillons, l'inconvénient de cette méthode c'est une distribution moins équilibrée des scores RFM combinés, et la possibilité que certains segments n'aient pas de clients.
Nombre de classes	Vous pouvez spécifier le nombre de classes pour les valeurs de récence, fréquence et monétaires. Les valeurs indiquées sont multipliées pour créer le nombre total de classes. Par exemple, si vous spécifiez 5 classes pour la récence, 5 classes pour la fréquence, et 5 pour la valeur monétaire, le nombre total de classes est 125.

Analyse RFM (Récence, Fréquence et Montant) : Enregistrer les résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Emplacement pour l'enregistreme	Par défaut, les scores de récence, de fréquence et monétaires (RFM) sont enregistrés dans une table de la bibliothèque Work. Pour changer cet emplacement, cliquez sur Modifier.

Nom d'option	Description
nt des scores RFM	
Variables de la table de sortie	Vous pouvez personnaliser les noms des variables de la table de sortie.

Analyse RFM (Récence, Fréquence et Montant) : Sélectionner des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Nom d'option	Description
Carte de monétisation	crée une carte thermique qui totalise les valeurs monétaires pour chaque combinaison de valeurs de classe de récence et de fréquence. Ce graphe s'affiche par défaut.
Graphiques de distribution RFM	crée un histogramme que vous pouvez utiliser pour vérifier la régularité de la distribution des valeurs de classes. Ce graphe s'affiche par défaut.
Synthèse des segments RFM	crée un tableau que vous pouvez utiliser pour vérifier la régularité de la distribution des valeurs de classes.

Analyse de la régression avec erreurs autorégressives

<i>À propos de la tâche Analyse de la régression avec erreurs autorégressives</i>	573
<i>Analyse de la régression avec erreurs autorégressives : Attribuer des variables aux fonctions d'analyse</i>	574
<i>Analyse de la régression avec erreurs autorégressives : Définir les options du modèle</i>	575
<i>Analyse de la régression avec erreurs autorégressives : Définir les options statistiques</i>	576
<i>Analyse de la régression avec erreurs autorégressives : Générer des graphes</i>	579
<i>Analyse de la régression avec erreurs autorégressives : Définition des options des résultats</i>	580
<i>Indiquer des titres et des notes de bas de page</i>	582
<i>Afficher les propriétés</i>	582

À propos de la tâche Analyse de la régression avec erreurs autorégressives

La tâche Analyse de la régression avec erreurs autorégressives évalue et prévoit des modèles de régression linéaire pour des données de séries chronologiques lorsque les erreurs ne sont pas indépendantes sur une période donnée (c'est-à-dire autocorrélées) ou lorsque la variance de l'erreur n'est pas constante (c'est-à-dire hétéroscédastique). Cette tâche génère des prévisions et l'intervalle de confiance de prévision lorsque la table de sortie comprend les valeurs futures des variables indépendantes. Elle est utile car elle exploite la partie chronologique et la partie systématique du modèle pour générer des valeurs prédites. Ce modèle tient compte des départs récents observés dans la tendance pour produire des prévisions.

Vous pouvez utiliser cette tâche pour effectuer ces tâches :

- Comparer des estimations de différents modèles et méthodes d'estimation dans une série chronologique. Par exemple, vous pouvez utiliser cette tâche pour comparer des modèles d'investissement sur une période donnée.
- Créer une prédiction de prévisions à court terme. De nombreuses séries chronologiques présentent une autocorrélation positive élevée car elles possèdent l'apparence lissée d'une progression aléatoire. Cette tâche améliore sensiblement l'adéquation des modèles à vos besoins, non seulement en introduisant des paramètres supplémentaires mais aussi en capturant les tendances aléatoires.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	AUTOREG
Produits SAS requis	SAS Base, SAS/ETS
Autres produits SAS recommandés	Néant

Analyse de la régression avec erreurs autorégressives : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variable dépendante	spécifie la variable dépendante pour l'analyse de régression. Vous devez attribuer une seule variable à cette fonction.

Nom du rôle	Description
Variables explicatives	servent de variables de régression indépendantes pour le modèle de régression. Si aucune variable n'est attribuée à cette fonction, seule la moyenne est ajustée. (Cette méthode permet d'obtenir les autocorrélations d'une série).
Variable ID temps	indique la variable utilisée pour attribuer des dates aux lignes. Vous devez attribuer une seule variable à cette fonction. La première variable ayant un format date, temps ou date temps SAS est automatiquement attribuée à cette fonction. Vous pouvez aussi créer une variable ID temps. Vous devez attribuer une variable à cette fonction si vous souhaitez générer des représentations graphiques à partir des analyses.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables . Remarque : vous ne pouvez pas grouper en fonction d'une variable que vous avez déjà sélectionnée comme variable dépendante ou variable indépendante.

Analyse de la régression avec erreurs autorégressives : Définir les options du modèle

Dans la liste, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Méthode	spécifie les méthodes du modèle. Vous avez le choix entre ces options : ■ Estimations du maximum de vraisemblance ■ Estimations inconditionnelles des moindres carrés ■ Estimations de Yule-Walker ■ Estimations itératives de Yule-Walker

Nom d'option	Description
Nombre maximum d'itérations	précise le nombre maximum d'itérations. La valeur par défaut est 50.
Utiliser la première séquence des données non manquantes uniquement	utilise la première séquence contiguë de données sans valeurs manquantes. Si vous ne sélectionnez pas cette option, toutes les observations complètes sont utilisées.
Inclure le paramètre de l'intercept dans le modèle	indique que le paramètre de la constante doit être inclus dans le modèle. Par défaut, il est inclus.
Ajuster un modèle autorégressif	spécifie l'ajustement du modèle autorégressif. Vous pouvez sélectionner une ou plusieurs de ces options : ■ Activez la case à cocher Ordre du traitement autorégressif, puis saisissez une valeur dans la zone ou bien sélectionnez Indiquer les décalages du traitement autorégressif et saisissez le sous-ensemble de décalages des erreurs autorégressives. ■ Activez la case à cocher Supprimer les paramètres autorégressifs non significatifs pour supprimer ces paramètres dans le modèle en commençant par les moins significatifs (sélection descendante). La suppression est effectuée une seule fois sur les estimations de Yule-Walker calculées après l'estimation ordinaire initiale des moindres carrés. ■ Saisissez une valeur dans la zone Supprimer si le niveau de significativité est supérieur pour spécifier le critère de niveau de significativité utilisé pour la sélection descendante.

Analyse de la régression avec erreurs autorégressives : Définir les options statistiques

Dans la liste, cliquez sur **Statistiques** pour accéder à ces options.

Vous pouvez calculer une ou plusieurs de ces statistiques :

Nom d'option	Description
Statistiques de Durbin-Watson	affiche les statistiques de Durbin-Watson jusqu'à l'ordre n . Utilisez la zone Inclure jusqu'à l'ordre pour spécifier n . Les valeurs possibles vont de 1 à 32.767. La valeur par défaut est 1.
Probabilité marginale du test de Durbin-Watson	affiche les p -values des statistiques du test de Durbin-Watson généralisées pour des échantillons de grande taille. Sélectionnez l'option Utiliser approx. linéarisée de la matrice de conception pour calculer les p -values qui utilisent une approximation linéarisée de la matrice de conception lorsque le modèle est non linéaire en raison de la présence d'un traitement des erreurs.
Statistique t de Durbin	affiche la statistique t de Durbin qui détecte l'autocorrélation résiduelle en présence de variables dépendantes retardées. Remarque : si vous n'avez pas affecté une variable au rôle Variables explicatives , vous ne pouvez calculer que la statistique t de Durbin.
Statistiques h et t de Durbin	permet de choisir d'afficher la statistique t ou h de Durbin. Dans la liste déroulante Type , sélectionnez la statistique à calculer. La statistique h de Durbin teste la présence d'une autocorrélation de premier ordre lorsque les régresseurs contiennent la variable dépendante retardée. Vous pouvez sélectionner cette variable dans la liste déroulante Variable dépendante retardée . (Les variables dans cette liste déroulante sont les variables auxquelles vous avez attribué la fonction Variables explicatives). Si la statistique h de Durbin ne peut pas être calculée, la statistique t asymptotiquement équivalente s'affichera à la place. Remarque : si vous avez attribué une variable à la fonction Variables explicatives , vous pouvez calculer soit la statistique h , soit la statistique t de Durbin.
Test général de Godfrey du multiplicateur de Lagrange par rapport à ARMA	affiche le test du multiplicateur de Lagrange par rapport aux erreurs ARMA jusqu'à l'ordre r . Utilisez la zone Inclure jusqu'à l'ordre pour spécifier r . Les valeurs possibles vont de 1 à 32.767. La valeur par défaut est 4.
Test de normalité de Jarque-Bera	affiche les statistiques de test de normalité de Jarque-Bera pour les résidus de régression.
Test de la racine unitaire de Phillips-Perron	affiche le test de la racine unitaire de Phillips-Perron lorsque le modèle ne comporte aucun régresseur. Pour spécifier des points de troncature pour les estimateurs de variance, activez la case à cocher Inclure retards de troncatures ? et saisissez les valeurs de retard dans la zone Retards de troncature . Trois types de test de Phillips-

Nom d'option	Description
	Perron sont signalés : moyenne zéro, moyenne simple et tendance déterministe. Remarque : si vous avez attribué une variable à la fonction Variables explicatives, les régresseurs sont inclus dans le modèle et le test de cointégration de Phillips-Ouliaris est affiché à la place.
Statistiques Q et LM pour l'absence d'effets ARCH	affiche une table des statistiques permettant de tester l'hétéroscédasticité.
Covariances estimées des résultats estimés des paramètres	affiche les covariances estimées pour les paramètres estimés.
Corrélations estimées des résultats estimés des paramètres	affiche les corrélations estimées pour les paramètres estimés.
Autocorrélations partielles	affiche les autocorrélations partielles.
Coefficients de transformation	affiche les coefficients de transformation des premières observations p . Ces coefficients sont formés par une scalaire multipliée par la racine inverse de Cholesky de la matrice Toeplitz d'autocovariances.
Inverse de la matrice Toeplitz des autocovariances	affiche l'inverse de la matrice Toeplitz des autocovariances pour la solution de Yule-Walker.
Log. de la vraisemblance du modèle de régression	affiche la valeur du logarithme de la vraisemblance du modèle de régression, sous l'hypothèse de normalité des erreurs.
Coef. de détermination de la régression	imprime le coefficient de détermination de la régression R^2 non centré. Ce coefficient R^2 est utile pour le calcul des statistiques de test du multiplicateur Lagrange, car la plupart de ces statistiques sont calculées en tant que T^*URSQ , où T correspond au nombre d'observations utilisées dans l'estimation.

Nom d'option	Description
Historique d'itérations de la fonction objective et des paramètres estimés	affiche la fonction objective et les estimations de paramètres à chaque itération. La fonction objective est la fonction de log-vraisemblance complète du journal pour la méthode du maximum de vraisemblance. La somme des carrés des erreurs est produite en tant que fonction objective des moindres carrés inconditionnels. Pour la méthode des estimations du maximum de vraisemblance, cette option affiche la valeur de la fonction du logarithme de la vraisemblance complète et non la vraisemblance concentrée.
Test RESET de Ramsey	affiche les statistiques du test RESET de Ramsey. Remarque : cette option n'est disponible que si vous avez attribué une variable à la fonction Variables explicatives .

Analyse de la régression avec erreurs autorégressives : Générer des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Par défaut, la sortie comprend tous les graphes appropriés pour la sélection de données actuelle. Vous pouvez néanmoins choisir les graphes à inclure dans la sortie en sélectionnant l'option **Liste personnalisée de graphes**. Vous avez le choix entre ces options :

Nom d'option	Description
Courbe de fonction d'autocorrélation	crée un graphique de la fonction d'autocorrélation.
Graphe de la fonction d'autocorrélation inversée des résidus	crée un graphique de la autocorrélation inverse.
Graphe de la fonction d'autocorrélation	crée un graphique de la fonction d'autocorrélation partielle.

Nom d'option	Description
ion partielle des résidus	
Courbe des valeurs prédites et réelles	crée un graphique des valeurs prédites par rapport aux valeurs observées.
Tracer la statistique D de Cook	crée un graphique de la statistique D de Cook par rapport au nombre d'observations. Remarque : cette option n'est pas disponible si vous avez activé la case à cocher Ajuster un modèle autorégressif dans le panneau Options.
Graphique Q-Q des résidus	crée une courbe de quantile normal des résidus.
Trace les résidus	crée un graphique des résidus comparés aux valeurs prédites. Un résidu correspond à la différence entre la valeur observée et la valeur prédite.
Trace les résidus studentisés	crée un graphique des résidus comparés aux valeurs prédites. Remarque : cette option n'est pas disponible si vous avez activé la case à cocher Ajuster un modèle autorégressif dans le panneau Options.
Trace les probabilités de bruit blanc	crée un graphique des probabilités de bruit blanc.
Trace l'histogramme des résidus	crée un histogramme des résidus d'ajustement.

Analyse de la régression avec erreurs autorégressives : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Inclure les statistiques	enregistre les statistiques au niveau observation. Vous avez le choix entre ces options : ■ Valeurs prédites, résidus et intervalle de confiance pour la moyenne inconditionnelle enregistre les valeurs prédites, les résidus et les bornes supérieure et inférieure de l'intervalle de confiance formés uniquement par la partie structurelle du modèle. Utilisez la liste déroulante Niveaux de confiance afin de définir la taille de l'intervalle de confiance pour les estimations de la partie structurelle du modèle. ■ Valeurs prédites, résidus et intervalle de confiance pour les futures réalisations enregistre les valeurs prédites, les résidus et les bornes supérieure et inférieure de l'intervalle de confiance formés par les parties structurelles et autorégressives du modèle. Utilisez la liste déroulante Niveaux de confiance pour définir la taille de l'intervalle de confiance pour les estimations des valeurs futures des séries chronologiques de réponse. ■ Niveaux de confiance et statistiques de somme cumulée (CUSUM) enregistre les statistiques CUSUM et CUSUMSQ, ainsi que les bornes supérieure et inférieure de leur intervalle de confiance. Utilisez la liste déroulante Niveaux de confiance pour définir le niveau de significativité pour les limites supérieure et inférieure. ■ Résidus récurrents et BLUS enregistre les valeurs des résidus BLUS de Theil, les résidus récurrents utilisés pour calculer les statistiques CUSUM et CUSUMSQ, et la partie de la variance de l'erreur prédite (vt) utilisée pour calculer les résidus récurrents.
Inclure les paramètres estimés	crée une table contenant toutes les variables utilisées dans le modèle. Chaque variable de régression contient l'estimation du paramètre de régression correspondant. Pour inclure la matrice de covariance pour les résultats estimés des paramètres, activez la case à cocher Inclure la matrice de covariance.
Supprimer l'affichage des résultats	n'affiche pas les résultats. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Analyse de régression des données du panel

<i>À propos de la tâche Analyse de régression des données du panel</i>	585
<i>Analyse de régression des données du panel : Attribuer des variables aux fonctions d'analyse</i>	586
<i>Analyse de régression des données du panel : Définir les options du modèle</i>	587
<i>Analyse de régression des données du panel : Définition des options des résultats</i> ...	588
<i>Indiquer des titres et des notes de bas de page</i>	589
<i>Afficher les propriétés</i>	589

À propos de la tâche Analyse de régression des données du panel

La tâche Analyse de régression des données du panel analyse une classe de modèles économétriques linéaires généralement obtenus en combinant des séries chronologiques et des données transversales. Cette tâche traite des tables du panel qui contiennent les observations de séries chronologiques sur chacune des unités transversales.

Pour exécuter cette tâche, les prérequis sont :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	TSCSREG
Produits SAS requis	SAS Base, SAS/ETS
Autres produits SAS recommandés	Néant

Analyse de régression des données du panel : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variable dépendante	indique la variable qui fera l'objet de la régression. Vous devez attribuer une seule variable à cette fonction.
Variable explicative	indique les variables indépendantes pour régresser la variable de réponse. Vous devez attribuer au moins une variable à cette fonction.
Variable ID des sections croisées	indique la variable qui définit les sections croisées. Vous devez attribuer une seule variable à cette fonction. Les données sont automatiquement triées par les valeurs de cette variable.
Variable ID temps	indique la variable utilisée pour attribuer des périodes aux observations de chaque section croisée. Vous devez attribuer une seule variable à cette fonction. Les données sont automatiquement triées par les valeurs de cette variable.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe.

Analyse de régression des données du panel : Définir les options du modèle

Dans la liste, ces options apparaissent sous l'en-tête **Options**.

Nom du graphe	Description
Méthode d'estimation à utiliser	Pour indiquer les méthodes d'estimation, sélectionnez les boutons d'option correspondant aux méthodes voulues. Les performances des méthodes d'estimation pour les paramètres de régression du modèle dépendent des caractéristiques statistiques des composantes de l'erreur dans le modèle. La méthode par défaut est Effets aléatoires à deux critères. ■ Un modèle est appelé modèle d'effets fixes à un facteur lorsque la spécification dépend uniquement de la section croisée dont l'observation fait partie. ■ Un modèle est appelé modèle d'effets fixes à deux critères lorsque la spécification dépend à la fois de la section croisée et de la série temporelle dont l'observation fait partie. ■ Un modèle est appelé modèle d'effets aléatoires à un facteur lorsque la spécification dépend uniquement de la section croisée dont l'observation fait partie. Dans ce cas, les effets sont aléatoires. Pour un panel équilibré, la méthode Fuller et Battese est utilisée par défaut. Pour un panel non équilibré, la méthode Wansbeek et Kapteyn est utilisée par défaut. ■ Un modèle est appelé Modèle d'effets aléatoires (composants de variance) à deux critères lorsque la spécification dépend uniquement de la section croisée dont l'observation fait partie. Dans ce cas, les effets sont aléatoires. Pour un panel équilibré, la méthode Fuller et Battese est utilisée par défaut. Pour un panel non équilibré, la méthode Wansbeek et Kapteyn est utilisée par défaut. ■ Un modèle autorégressif (méthode de Parks) suppose une structure d'erreur autorégressive de premier ordre avec corrélation contemporaine entre sections croisées. Si vous sélectionnez la méthode de Parks, vous pouvez utiliser une courbe spline cubique pour interpoler les valeurs manquantes de la série. Une spline cubique est une fonction segmentée composée de fonctions polynomiales de troisième degré (cubiques) jointes afin que la courbe tout entière et ses première et seconde valeurs dérivées soient continues. La courbe spline s'ajuste aux valeurs non manquantes des

Nom du graphe	Description
	variables. Cette courbe est utilisée pour établir les approximations de durée continue des séries d'entrée, tandis que les séries chronologiques de sortie sont générées à partir des approximations de spline. Par défaut, la case à cocher Interpolation de spline cubique des valeurs manquantes est sélectionnée. Remarque : l'option Interpolation de spline cubique des valeurs manquantes doit être utilisée avec prudence. En effet, l'interpolation introduit souvent des distorsions dans le modèle, ce qui peut entraîner des estimations de paramètre et des tests d'hypothèses de mauvaise qualité. Pour plus d'informations sur ces méthodes, voir <i>SAS/ETS: User's Guide</i>.
Inclure le paramètre de l'intercept dans le modèle	Pour exclure le paramètre de l'intercept, désactivez la case à cocher de cette option. Par défaut, la case à cocher est activée.

Analyse de régression des données du panel : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Pour enregistrer les résultats estimés des paramètres dans une table de sortie, activez la case à cocher **Enregistrer résultats estimés des paramètres**. Si vous le souhaitez, activez les cases à cocher permettant d'inclure la **Matrice de covariance**, la **Matrice de corrélation** ou les deux.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Pour afficher les résultats, activez la case à cocher **Afficher analyse**. Pour supprimer les résultats, désactivez cette case à cocher. Par défaut, les résultats sont affichés.

Si vous choisissez d'afficher vos résultats, activez les cases à cocher pour les éléments que vous voulez faire apparaître. Certains éléments ne sont disponibles qu'avec la méthode d'estimation de Parks.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

.....

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,

- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Graphiques statistiques SAS ODS

<i>À propos des Graphiques statistiques SAS ODS</i>	591
<i>Utiliser ODS Graphics Designer</i>	592
<i>Afficher le graphique statistique ODS</i>	592

À propos des Graphiques statistiques SAS ODS

SAS ODS Statistical Graphics, plus communément appelé SAS ODS Graphics, est une extension de SAS Output Delivery System (ODS). ODS gère toutes les sorties créées par des procédures et vous permet d'afficher la sortie dans une variété de formats, notamment HTML et PDF.

De nombreuses procédures analytiques SAS utilisent les fonctionnalités ODS Graphics pour générer des graphiques de façon aussi automatique que ces procédures génèrent des tables. ODS Graphics utilise la syntaxe du langage GTL (Graph Template Language) qui fournit la puissance et la flexibilité permettant de créer de nombreux graphiques complexes. GTL est un langage complet qui permet de définir des graphiques statistiques.

Dans SAS Enterprise Guide, vous pouvez utiliser l'ODS Graphics Designer pour définir des graphiques statistiques sans connaître le GTL. Lorsqu'une définition graphique est créée, vous pouvez l'utiliser pour créer un graphique statistique ODS dans SAS Enterprise Guide ou SAS Add-In for Microsoft Office.

Utiliser ODS Graphics Designer

Remarque : le SAS ODS Graphics Designer n'est pas disponible dans SAS Add-In for Microsoft Office, mais vous pouvez ouvrir les fichiers SGD qui ont été créés à l'aide de SAS ODS Graphics Designer dans l'Add-in SAS.

SAS ODS Graphics Designer est une application graphique interactive que vous pouvez utiliser pour créer des graphiques personnalisés. Cette application crée des graphiques qui sont basés sur le Graph Template Language (GTL), qui est aussi le langage utilisé par les procédures analytiques SAS et les procédures graphiques SAS ODS. ODS Graphics Designer fournit une interface utilisateur graphique vous permettant de créer facilement des graphiques sans avoir besoin de connaître les détails des modèles et du GTL.

Remarque : à partir de la première version de maintenance de SAS 9.4, vous pouvez utiliser le SAS ODS Graphics Designer dans SAS Enterprise Guide sans l'installation de Base SAS sur votre machine locale.

Grâce à une interface de type pointer-cliquer, vous pouvez créer des vues graphiques simples ou complexes des données à analyser. L'ODS Graphics Designer vous permet de créer des graphiques sophistiqués en utilisant un large éventail de types de graphes. Vous pouvez créer des graphiques multi-cellules, des volets de classification et des matrices du nuage de points. Vos graphiques peuvent comporter des titres, des notes de bas de pages, des légendes et autres éléments graphiques. Vous pouvez enregistrer les résultats sous forme d'image à inclure dans un rapport ou sous forme de fichier ODS Graphics Designer (SGD) modifiable. Pour plus d'informations, voir [“Afficher le graphique statistique ODS” page 592](#).

Pour ouvrir l'ODS Graphics Designer dans SAS Add-In for Microsoft Office ou SAS Enterprise Guide, faites un double clic sur **Afficher le graphique statistique ODS** dans le volet **Tâches**. Pour plus d'informations, voir [SAS Output Delivery System: User's Guide](#).

Afficher le graphique statistique ODS

La tâche Graphique statistique ODS est disponible à la fois dans SAS Enterprise Guide et SAS Add-In for Microsoft Office. Dans cette tâche, vous sélectionnez la définition de graphique (fichier SGD) à utiliser pour une source de données sélectionnée. Vous pouvez créer une définition de graphique dans SAS Enterprise Guide. Pour plus d'informations, voir [“Utiliser ODS Graphics Designer” page 592](#).

Lorsqu'une définition de graphique est créée, elle peut être utilisée dans des sessions SAS locales et distantes. Vous pouvez même restituer ces graphiques dans des applications stockées SAS.

Pour utiliser une définition de graphique (fichier SGD) pour utiliser un graphique statistique ODS :

- 1 Dans le volet **Tâches** dans SAS Add-In for Microsoft Office ou SAS Enterprise Guide, faites un double clic sur **Afficher le graphique statistique ODS**.

La boîte de dialogue **Afficher le graphique statistique ODS** apparaît.

- 2 Sélectionnez le fichier SGD contenant la définition de graphique à utiliser.
- 3 Sélectionnez la source de données pour le graphique.
- 4 (facultatif) Indiquez la largeur et la hauteur du graphique.
- 5 (facultatif) Indiquez un titre et une note de bas de page.
- 6 Cliquez sur **Exécuter**.

Pour plus d'informations sur les définitions de graphique, voir [SAS Output Delivery System: User's Guide](#).

SAS Rapid Predictive Modeler

<i>À propos de SAS Rapid Predictive Modeler</i>	595
<i>Stratégies d'échantillonnage de SAS Rapid Predictive Modeler</i>	596
<i>SAS Rapid Predictive Modeler : Organiser les données</i>	597
<i>SAS Rapid Predictive Modeler : Attribuer des variables aux fonctions d'analyse</i>	599
<i>SAS Rapid Predictive Modeler : Sélectionner un modèle</i>	600
Définir les options du modèle	600
Comprendre les modèles de SAS Rapid Predictive Modeler	602
<i>SAS Rapid Predictive Modeler : Définir les options du rapport</i>	605
Options du rapport	605
À propos des rapports standard de SAS Rapid Predictive Modeler	606
<i>SAS Rapid Predictive Modeler : Définir les options de sortie</i>	606
<i>SAS Rapid Predictive Modeler : Enregistrer un modèle</i>	607
À propos de l'inscription d'un modèle	608
Inscrire un modèle dans le référentiel de métadonnées SAS	608

À propos de SAS Rapid Predictive Modeler

Le logiciel SAS Rapid Predictive Modeler permet de créer des modèles pour les problèmes de classification et de régression du data mining suivants :

- modèles de classification prédisant la valeur d'une variable discrète. Il s'agit notamment des modèles de classification capables de prédire la valeur d'une variable, comme Vrai ou Faux, Achat ou Refus, Elevé, Moyen ou Faible, et Attrition ou Rétention.
- modèles de régression prédisant la valeur d'une variable à nombre réel. Ces modèles peuvent notamment prédire des montants tels que le chiffre d'affaires, les ventes ou le taux de réussite d'après des valeurs réelles.

Pour créer un modèle à l'aide de SAS Rapid Predictive Modeler, vous devez fournir une table dont chaque ligne contient un ensemble de variables de prédiction indépendantes (ou variables explicatives) et au moins une variable dépendante (ou variable à expliquer). Le logiciel SAS Rapid Predictive Modeler détermine si les variables sont continues ou catégorielles et choisit les variables explicatives à inclure dans le modèle. Pour plus d'informations sur la façon dont la tâche échantillonne les données, voir *Stratégies d'échantillonnage de SAS Rapid Predictive Modeler*.

Vous pouvez enregistrer un modèle sous forme de code SAS, puis le déployer dans un environnement SAS. Vous pouvez ensuite utiliser le code du modèle SAS pour procéder au scoring des nouvelles données, puis vous servir des résultats obtenus pour prendre des décisions commerciales en toute connaissance de cause. C'est ce que l'on appelle le scoring du modèle. Les données ayant subi un scoring permettent d'identifier les clients à sélectionner pour certaines offres, d'anticiper les besoins de production, d'entreposage ou logistiques, de gérer le taux d'attrition ou de détecter les éventuelles transactions frauduleuses.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	
Produits SAS requis	SAS Enterprise Miner
Autres produits SAS recommandés	

Stratégies d'échantillonnage de SAS Rapid Predictive Modeler

L'application SAS Rapid Predictive Modeler utilise une approche d'échantillonnage composite. Le nombre d'observations incluses dans l'échantillon de données dépend de ces facteurs :

- Nombre de variables explicatives
- dans la source de données
- Présence ou absence dans les données de variables à expliquer comportant un événement rare
- Nombre d'événements dans les données

Voici les instructions utilisées par la tâche SAS Rapid Predictive Model pour déterminer le nombre d'observations traitées :

Nombre de variables explicatives	Nombre d'observations traitées
< 100	80 000
100 – 200	40 000
> 200	20 000

Condition	Événement rare : Oui	Événement rare : Non
< nombre d'observations en cours de traitement OU nombre total d'événements < $(0,10 \times \text{nombre d'observations en cours de traitement})$	Les données sont échantillonnées afin d'obtenir un ratio de 10 non-événements pour 1 événement	Pas d'échantillonnage
nombre total d'événements > $(0,10 \times \text{nombre d'observations traitées})$	Échantillonnez la proportion suivante d'événements rares : $10 \times [(0,10 \times (\text{nombre d'observations en cours de traitement})) / \text{nombre d'événements}]$	Échantillonnage stratifié

SAS Rapid Predictive Modeler : Organiser les données

Avant de pouvoir créer un modèle, vous devez avoir des données en entrée qui représentent les événements historiques et les caractéristiques pouvant être utilisés pour la prédiction. Il vous faut également des données cibles qui représentent l'événement ou la valeur à prédire. Dans de nombreux cas, les données en entrée proviennent d'une période de temps et les données cibles proviennent d'une période ultérieure. Les données en entrée et les données cibles combinées, utilisées pour développer votre modèle s'appellent les données d'apprentissage.

Par exemple, vous pouvez exploiter les recettes de l'exercice précédent pour prévoir le chiffre d'affaires de l'exercice suivant ou les clients qui profiteront d'une offre spéciale. L'utilisation des données historiques issues d'événements passés pour prédire la performance sur les événements futurs s'appelle l'apprentissage du modèle.

Pour obtenir de meilleurs résultats, vos données d'apprentissage doivent contenir un grand nombre d'observations stockées en tant que lignes des données. Par exemple, la plupart des modèles de vente au détail utilisent des données en entrée contenant des dizaines de milliers d'observations.

Si la variable à expliquer contient un événement rare (une offre intéressant seulement 1 % des clients, par exemple), vous devez vous assurer que les données d'apprentissage de votre table contiennent un nombre significatif de ces clients. Vous pouvez suréchantillonner les données d'apprentissage pour être certain de sélectionner tous les clients ayant accepté l'offre, et fournir un nombre égal de clients l'ayant refusée. Grâce au suréchantillonnage, un modèle qui comporte une variable à expliquer avec un événement rare identifie plus facilement une solution stable.

Lorsque vous réalisez un suréchantillonnage pour augmenter le nombre d'occurrences d'un événement rare dans les données d'apprentissage, vous gonflez artificiellement le nombre d'événements ciblés dans les données d'apprentissage par rapport à la population naturelle. Pour compenser la différence entre données d'apprentissage et données de population, SAS Rapid Predictive Modeler propose des paramètres de probabilité a priori. Ces paramètres indiquent la fréquence proportionnelle réelle de l'événement ciblé dans les données de population.

Les données explorées avec SAS Rapid Predictive Modeler doivent être organisées en lignes (observations) et en colonnes (variables). L'une des colonnes doit représenter une variable à expliquer. Prenons l'exemple de table suivant :

Nom	Age	Sexe	Revenu	Traitement	Achat
Ricardo	29	M	33000	O	O
Susan	35	F	51000	O	N
Jeremy	49	M	110000	N	O

Nom

colonne contenant les valeurs ID pour chaque observation. SAS Rapid Predictive Modeler ne traite pas les colonnes de variables ID pour le contenu analytique.

Age, Sexe, Revenu, et Traitement

colonnes en entrée (variables explicatives) utilisées par SAS Rapid Predictive Modeler.

Achat

colonne cible.

Lorsque vous configurez votre table de données en entrée, vous pouvez également désigner une colonne de fréquence. Les variables de fréquence peuvent représenter le poids relatif à attribuer à une ligne (ou observation). Ainsi, dans certaines tables, une même ligne peut représenter plusieurs observations. Vous pouvez également sélectionner des colonnes afin que SAS Rapid Predictive Modeler les ignore pendant une analyse.

Les données d'apprentissage requièrent toujours des valeurs de variables explicatives et à expliquer. En revanche, les données utilisées pour le scoring

requièrent uniquement des valeurs de variables explicatives : la colonne cible est donc facultative. Si le modèle sert à établir des prévisions à partir de nouvelles données, la colonne cible n'est pas nécessaire. S'il est utilisé pour contrôler l'efficacité, elle est en revanche présente. D'autre part, les données utilisées pour le scoring incluent généralement une colonne d'identification.

SAS Rapid Predictive Modeler : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Les fonctions d'analyse disponibles varient selon le type de données que vous voulez analyser.

Nom du rôle	Description
Variable dépendante	spécifie la valeur que vous voulez prédire ou classifier. La variable dépendante est également connue sous le nom de variable à expliquer. Vous devez attribuer au moins une variable à cette fonction.
Nombre d'occurrences	spécifie la variable à utiliser pour représenter les poids que vous voulez appliquer aux lignes de données. Vous ne pouvez attribuer qu'une seule variable à cette fonction.
ID	spécifie les variables qui sont utiles pour les fonctions de reporting et de sélection de scoring. Ces variables ne sont pas incluses dans l'analyse analytique.
Exclu	spécifie les variables que vous ne voulez pas inclure dans l'analyse analytique.

SAS Rapid Predictive Modeler : Sélectionner un modèle

Définir les options du modèle

Dans la liste, cliquez sur **Modèle** pour accéder à ces options.

Avec ces options, vous pouvez spécifier le niveau de complexité du modèle à créer. Les méthodes de modélisation sont hiérarchisées : la méthode intermédiaire comprend les modèles classiques et intermédiaires, tandis que la méthode avancée inclut les modèles classiques, intermédiaires et avancés.

Les modèles créés selon la méthode classique sont souvent plus rapides à exécuter que ceux élaborés avec la méthode intermédiaire, mais ils risquent d'être moins précis. Cela vaut également pour la comparaison des modèles créés selon les méthodes intermédiaire et avancée. Pour plus d'informations sur les opérations de data mining réalisées pour chaque modèle, voir ["Comprendre les modèles de SAS Rapid Predictive Modeler" page 602](#).

Les fonctions de modélisation de SAS Enterprise Miner sont exécutées lorsque vous utilisez SAS Rapid Predictive Modeler, selon la méthode de modélisation sélectionnée.

Pour la variable dépendante, cliquez sur **Décision et Valeurs a priori** pour spécifier les informations suivantes :

Tableau 76.1 Options de décision et de valeurs a priori pour la variable dépendante

Nom d'option	Description
Niveau de l'événement	spécifie la valeur de variable à expliquer qualitative que vous voulez modéliser. SAS Rapid Predictive Modeler crée un modèle indiquant la probabilité de chaque événement cible, mais sa fonction de reporting est améliorée lorsque le niveau souhaité est connu.
Probabilités a priori	affiche le nombre et la proportion de niveaux de variable à expliquer présents dans les données d'apprentissage du modèle. Vous pouvez ajuster ces valeurs lorsque votre variable à expliquer est une variable catégorielle, et que les données d'apprentissage et de population comportent différentes répartitions. Prenons l'exemple d'un modèle soumis à un apprentissage à partir de données suréchantillonnées où 50 % des observations correspondent aux répondants et 50 %

Nom d'option	Description
	aux personnes n'ayant pas répondu. Néanmoins, les données de population que le score du modèle soumettra à un scoring historique ne contiennent que 10 % de personnes ayant répondu et 90 % de personnes n'ayant pas répondu. Dans ce cas, vous pouvez utiliser les paramètres de probabilité a priori pour informer le modèle des proportions historiques attendues de répondants et de non-répondants. ■ Si vous ne souhaitez pas définir de probabilités a priori, conservez la valeur par défaut Néant. ■ Pour affecter des probabilités identiques à tous les niveaux de la variable à expliquer, sélectionnez Égal. ■ Pour définir vos propres probabilités a priori pour les niveaux de variable à expliquer des données ayant subi un scoring, sélectionnez Défini par l'utilisateur et saisissez les probabilités. Les probabilités a priori spécifiées ne doivent pas dépasser 1.
Fonction de décision	spécifie les coûts, profits ou poids que vous souhaitez associer aux résultats prédits. La table contenant ces valeurs est appelée matrice de décision. Vous l'utilisez pour associer une valeur à chaque décision possible. ■ Si une matrice de décision est inutile dans un modèle, sélectionnez l'option par défaut Néant. ■ Si vous souhaitez utiliser votre modèle pour optimiser la rentabilité, sélectionnez Maximum et, le cas échéant, saisissez un poids plus élevé dans la cellule Vrais positifs de la matrice. ■ Si vous voulez vous en servir pour réduire les coûts, sélectionnez Minimum et, le cas échéant, saisissez un poids plus élevé dans la cellule Vrais négatifs de la matrice. ■ Enfin, si vous souhaitez utiliser le modèle pour prévoir des événements rares, sélectionnez la valeur Inverse afin d'identifier à la fois les valeurs vraies positives et négatives, au risque d'obtenir des prévisions fausses positives et négatives imprécises.

Tableau 76.2 Options de modélisation

Nom d'option	Description
Classique	spécifie la valeur que vous voulez prédire ou classer. La variable dépendante est également connue sous le nom de variable à expliquer. Vous devez attribuer au moins une variable à cette fonction.

Nom d'option	Description
Intermédiaire	spécifie la variable à utiliser pour représenter les poids que vous voulez appliquer aux lignes de données. Vous ne pouvez attribuer qu'une seule variable à cette fonction.
Options avancées	spécifie les variables qui sont utiles pour les fonctions de reporting et de sélection de scoring. Ces variables ne sont pas incluses dans l'analyse analytique.

Comprendre les modèles de SAS Rapid Predictive Modeler

The SAS Rapid Predictive Modeler propose les modèles Classique, Intermédiaire et Avancé. Les modèles sont de sophistication et de complexité croissantes.

- Le modèle Classique est une analyse de régression simple.
- Le modèle Intermédiaire comprend une analyse plus sophistiquée, plus l'analyse du modèle classique, et choisit le meilleur modèle.
- Le modèle Avancé comprend une analyse plus sophistiquée, plus l'analyse du modèle Classique et du modèle Intermédiaire, et choisit le meilleur modèle.

Modèle de base

Le modèle Classique effectue successivement trois opérations de data mining.

- Sélection de variables : Le modèle Classique choisit les 100 premières variables pour la modélisation.
- Transformation : Le modèle Classique soumet les 100 premières variables sélectionnées pour la modélisation à une transformation de discrétisation optimale. Cette transformation compense les valeurs de variable manquantes, et aucune imputation des valeurs manquantes n'est donc effectuée.
- Modélisation : Le modèle Classique utilise un modèle de régression ascendante. Ce modèle choisit les variables une par une dans le cadre d'un processus pas à pas, qui consiste à ajouter une variable à la fois à l'équation linéaire jusqu'à ce que la contribution des variables soit insignifiante. Le modèle de régression ascendante vise à exclure les variables dépourvues d'intérêt prédictif (ou fortement corrélées avec d'autres variables de prédiction) du processus analytique.

Modèle intermédiaire

Le modèle Intermédiaire effectue successivement sept opérations de data mining.

- Sélection de variables : Le modèle Intermédiaire choisit les 200 premières variables pour la modélisation.
- Transformation : Le modèle Intermédiaire soumet les 200 premières variables sélectionnées pour la modélisation à une transformation de puissance optimale. Ce type de transformation est un sous-ensemble de la catégorie générale de transformations de Box-Cox. La transformation de puissance optimale évalue un sous-ensemble de transformations de puissance exponentielle et choisit celle qui offre les meilleurs résultats pour le critère spécifié.
- Imputation : Le modèle Intermédiaire effectue une imputation pour remplacer les variables manquantes par les valeurs de variable moyennes. Cette opération crée également des variables indicatrices qui permettent d'identifier les observations contenant des valeurs de variable imputées.
- Sélection de variables : Le modèle Intermédiaire utilise les tests de critères Khi-2 et R-carré pour supprimer les variables qui ne sont pas associées à la variable à expliquer.
- Combinaison de techniques de sélection de variables : Le modèle Intermédiaire fusionne l'ensemble des variables sélectionnées par les tests de critères Khi-2 et R-carré.
- Modélisation : Le modèle Intermédiaire soumet les données d'apprentissage à trois algorithmes de modèles concurrents. un arbre de décision, une régression logistique et une régression pas à pas. Dans le cas du modèle de régression logistique, les données d'apprentissage sont tout d'abord soumises à un arbre de décision qui crée une variable NODE_ID transmise au modèle de régression en tant que variable explicative. La variable NODE_ID est créée pour permettre l'élaboration de modèles d'interaction de variables.
- Sélection du meilleur modèle : Le modèle Intermédiaire effectue une évaluation analytique des performances de prévision ou de classification des modèles concurrents. Le modèle offrant les meilleures performances est sélectionné pour procéder à l'analyse de modélisation. Le modèle Intermédiaire pour la sélection du meilleur modèle évalue les performances non seulement des modèles intermédiaires, mais aussi des modèles classiques.

Après avoir sélectionné le meilleur modèle intermédiaire, SAS Rapid Predictive Modeler en compare ses performances de prévision avec celles du modèle classique, puis choisit le modèle optimal.

Modèle avancé

Le modèle Avancé effectue successivement sept opérations de data mining.

- Sélection de variables : Le modèle avancé choisit les 400 premières variables pour la modélisation.

- Transformation : Le modèle Avancé applique l'algorithme de transformation multiple aux 400 variables sélectionnées pour la modélisation. L'opération de transformation multiple crée plusieurs transformations de variables destinées à être utilisées dans des sélections de variables ultérieures. Les transformations multiples entraînent une augmentation du nombre de variables explicatives. Du fait de cette augmentation, SAS Rapid Predictive Modeler sélectionne les 400 meilleures variables explicatives parmi les résultats générés par l'algorithme de transformation multiple.
- Imputation : Le modèle Avancé effectue une imputation pour remplacer les variables manquantes par les valeurs de variable moyennes. Cette opération crée également des variables indicatrices qui permettent à l'utilisateur d'identifier les observations contenant des valeurs de variable imputées.
- Sélection de variables : Le modèle Avancé utilise les tests de critères Khi-2 et r-carré pour supprimer les variables qui ne sont pas associées à la variable à expliquer. Des variables AOV16 sont créées au cours de l'analyse R-carré.
- Combinaison de techniques de sélection de variables : Le modèle Avancé fusionne l'ensemble des variables sélectionnées par les tests de critères Khi-2 et R-carré.
- Modélisation : Le modèle Avancé soumet les données d'apprentissage à quatre algorithmes de modèles concurrents. Les modèles sont : un arbre de décision, un modèle de réseau neuronal, un modèle de régression descendante et un modèle d'ensemble. Le modèle de réseau neuronal effectue des recherches restreintes afin de détecter un réseau ascendant optimal. La régression descendante est un modèle de régression linéaire qui élimine des variables en les retirant une par une à la fois jusqu'à ce que les scores du test R-carré chutent de manière significative. Le modèle d'ensemble crée de nouveaux modèles en combinant les probabilités a posteriori (pour les variables qualitatives à expliquer) ou les valeurs prédites (pour les variables continues à expliquer) de plusieurs modèles d'entrée prédécesseurs. Le nouveau modèle d'ensemble est alors utilisé pour procéder au scoring de nouvelles données. Le modèle d'ensemble utilisé dans le modèle avancé est créé à partir du résultat du modèle classique, du meilleur modèle du modèle intermédiaire, et du meilleur modèle du modèle avancé.
- Sélection du meilleur modèle : Le modèle avancé effectue une évaluation analytique des performances de prévision ou de classification des modèles concurrents d'arbre de décision, neuronaux et de régression. Le modèle qui offre les meilleures performances de prévision ou de classification est alors utilisé en entrée, avec les meilleurs modèles classique et intermédiaire, pour former un modèle d'ensemble. Ensuite, les nouveaux modèles d'ensemble, d'arbre de décision, neuronaux et de régression descendante sont soumis à une analyse comparative afin de déterminer le modèle optimal parmi les échantillons des meilleurs modèles classiques, intermédiaires et avancés.

Une fois que le SAS Rapid Predictive Modeler a sélectionné le meilleur modèle, il exécute et compare les performances de prévision du modèle Avancé avec les meilleurs modèles Intermédiaire et Classique, puis choisit le plus performant.

SAS Rapid Predictive Modeler : Définir les options du rapport

Options du rapport

Dans la liste, cliquez sur **Rapport** pour accéder à ces options.

Ils identifient les termes significatifs du modèle et génèrent des graphiques professionnels courants, tels que des diagrammes de levier (lift). Les résultats comprennent des statistiques pour les données d'apprentissage et de validation. Le processus SAS Rapid Predictive Modeler consiste en effet à diviser les données en entrée en données d'apprentissage et données de validation. Les données d'apprentissage servent à calculer les paramètres de chaque modèle et à obtenir les tests d'ajustement de l'apprentissage. Les données de validation sont ensuite soumises à un scoring avec chaque modèle et permettent d'obtenir les tests d'ajustement de la validation. Les tests d'ajustement des données de validation sont utilisés pour comparer les modèles et détecter le surapprentissage. Lorsque les statistiques d'apprentissage sont nettement meilleures que celles de validation, c'est souvent le signe d'un surapprentissage, qui survient lorsque le modèle est soumis à un apprentissage pour détecter des signaux aléatoires dans les données. Les modèles dont les statistiques de validation sont meilleures sont généralement privilégiés.

SAS Rapid Predictive Modeler génère automatiquement un ensemble succinct de rapports essentiels qui présentent une synthèse de la source de données et des variables utilisées pour la modélisation, un classement des variables explicatives importantes, plusieurs tests d'ajustement qui évaluent la précision du modèle, ainsi qu'un scorecard du modèle. Pour plus d'informations, voir À propos des rapports standard de SAS Rapid Predictive Modeler.

Vous pouvez également choisir d'inclure ces informations dans vos résultats :

- **Agrégation du modèle**
- **Classement des variables**
- **Table à double entrée**
- **Matrice de classification**
- **Tests d'ajustement**
- **Courbe d'élévation**
- **Comparaison de modèles**

Remarque : cette option n'est pas disponible pour la méthode classique, qui utilise un seul algorithme de modèle.

À propos des rapports standard de SAS Rapid Predictive Modeler

Voici les rapports standard qui sont générés automatiquement par SAS Rapid Predictive Modeler :

Courbe des gains

Ces courbes sont uniquement disponibles pour les modèles comportant des variables qualitatives à expliquer. Ce graphique affiche les centiles de données classées par valeur prédite. Le levier (lift) mesure le rapport entre le nombre d'événements cibles identifiés par le modèle et celui identifié par la sélection aléatoire.

Courbe ROC (Receiver Operating Characteristic)

La courbe ROC affiche la puissance de prévision maximale d'un modèle pour l'échantillon tout entier (et non un décile). Les données tracées comparent Sensibilité et (1 – Spécificité). L'écart entre la courbe du modèle et la diagonale (qui représente un modèle de sélection aléatoire) est appelé valeur de Kolmogorov-Smirnov (KS). Plus la valeur KS est élevée, plus le modèle est efficace.

Scorecard

Les résultats comprennent un scorecard pour que les caractéristiques du modèle puissent être interprétées à des fins professionnelles. Chaque variable continue est alors répartie en intervalles de valeurs distincts. Ensuite, chaque variable est classée selon l'importance du modèle et mise à l'échelle jusqu'à 1 000 points maximum. La valeur distincte de chaque variable reçoit alors une partie du nombre total de points de mise à l'échelle.

Informations sur le projet

Les informations du projet indiquent quel utilisateur a créé le modèle, quand le modèle a été créé, et où sont stockés les fichiers du composant du modèle.

SAS Rapid Predictive Modeler : Définir les options de sortie

Dans la liste, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Enregistrer les données du projet Enterprise Miner	indique s'il faut enregistrer les données SAS Enterprise Miner de cette tâche. Un modèle SAS Rapid Predictive Modeler est un exemple de projet SAS Enterprise Miner . Par conséquent, lorsque vous enregistrez des données SAS Enterprise Miner , vous pouvez utiliser l'interface de SAS Enterprise Miner pour ouvrir et modifier le modèle créé à l'aide de SAS Rapid Predictive Modeler. Dans SAS Enterprise Miner, vous pouvez enregistrer et exporter une analyse pour l'utiliser en dehors de SAS Enterprise Miner, et enregistrer le modèle avec un référentiel de métadonnées SAS. La case Enregistrer les données du projet Enterprise Miner est cochée par défaut. ■ Si vous avez une installation sur station de travail de SAS Enterprise Miner, le projet sera enregistré localement à l'emplacement de la station de travail spécifié dans le champ Dossier SAS App Server. Cliquez sur Parcourir pour changer cet emplacement. ■ Si vous avez une installation client/serveur de SAS Enterprise Miner, le projet est enregistré dans le dossier Racine de l'utilisateur (par exemple, C:\users\username\Documents) sur le serveur. Si vous avez utilisé SAS Management Console pour définir un emplacement racine SYSTEME (par exemple, C:\projects) sur le serveur, le champ Dossier SAS App Server est vide. Dans ce cas, cliquez sur Parcourir pour spécifier l'emplacement du stockage de votre projet. Si le champ Dossier SAS App Server reste vide et que vous n'indiquez pas un emplacement avant de tenter d'enregistrer le projet, un message d'erreur s'affiche, vous demandant de spécifier un emplacement valide.
Scorer une table	spécifie le nom de la table que vous voulez être scorée par le modèle généré par SAS Rapid Predictive Modeler.

SAS Rapid Predictive Modeler : Enregistrer un modèle

Dans la liste, cliquez sur **Inscription** pour accéder à ces options.

À propos de l'inscription d'un modèle

Les modèles sont créés à l'aide de SAS Rapid Predictive Modeler sont faciles à déployer. Ils sont enregistrés sous forme de code SAS exécutable sur toute installation SAS 9.2 (ou version ultérieure).

Remarque : vous devez exécuter ou enregistrer SAS Rapid Predictive Modeler au moins une fois avant de pouvoir inscrire votre modèle avec le référentiel de métadonnées SAS. Lors de la création du modèle, vous devez sélectionner l'option **Enregistrer les données du projet Enterprise Miner** pour que les options d'inscription soient disponibles.

Lorsque vous inscrivez un modèle, d'autres utilisateurs SAS de votre site peuvent utiliser le modèle inscrit.

- Les utilisateurs de **SAS Enterprise Guide** et de **SAS Add-In for Microsoft Office** peuvent s'en servir pour leurs propres activités de scoring. Par exemple, un modèle inscrit peut servir à scorer de nouvelles données avec la tâche Scoring du modèle.
- Les utilisateurs de SAS Enterprise Miner peuvent ouvrir ces modèles sous forme de projets et diagrammes dans SAS Enterprise Miner, et les modifier pour réaliser d'autres diagnostics ou étudier des comportements et envisager des alternatives.
- Les utilisateurs de **SAS Model Manager** peuvent importer le modèle à gérer avec d'autres ressources de modélisation, mais également de suivre la détérioration de ses performances.
- Les utilisateurs de SAS Data Integration Studio peuvent importer le modèle afin de créer des processus de scoring gérés et planifiés.

Les modèles créés à l'aide de SAS Rapid Predictive Modeler peuvent également être publiés via SAS Scoring Accelerator pour Teradata, DB2, ou Netezza.

Inscrire un modèle dans le référentiel de métadonnées SAS

- 1 Spécifiez le nom du modèle, son chemin et sa description.
- 2 Cliquez sur **Inscrire**.

Remarque : si vous n'êtes pas connecté à un SAS Workspace Server, le bouton **Inscrire** n'est pas disponible.

Nuage de points

<i>À propos de la tâche Nuage de points</i>	609
<i>Nuage de points : Sélectionner un type de graphe</i>	610
<i>Nuage de points : Attribuer des variables aux fonctions d'analyse</i>	610
<i>Nuage de points : Définir les options d'apparence</i>	612
Définir les options du nuage de points en 3D	612
Définir d'autres options de tracé	613
Options d'apparence supplémentaires	613
<i>Indiquer des titres et des notes de bas de page</i>	614
<i>Afficher les propriétés</i>	614

À propos de la tâche Nuage de points

La tâche Nuage de points crée des nuages de points en 2D, des nuages de points en 3D ou des diagrammes en bâtons en 3D qui présentent les relations entre deux ou trois variables, en mettant en évidence des motifs ou des concentrations de points de données.

Par exemple, un nuage de points en 2D permet d'afficher le poids et l'âge de tous les patients participant à une étude clinique.

Pour plus d'informations, voir ["Sélectionner un format de sortie pour les graphiques" page 6](#).

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GPLOT, G3D
Produits SAS requis	SAS Base, SAS/GRAPH

Nom du prérequis	Noms des procédures et des produits
Autres produits SAS recommandés	Néant

Nuage de points : Sélectionner un type de graphe

Dans la liste, cliquez sur **Nuage de points** pour accéder à ces options.

Nom du rôle	Description
Nuage de points 2D	Crée un graphe en deux dimensions affichant les valeurs de deux variables sous forme de points de données sur un axe horizontal (x) et un axe vertical (y). Chaque paire de valeurs x et y forme un point de données.
Nuage de points 3D	Crée un graphique en 3D qui affiche les valeurs de trois variables numériques sous forme de point de données. Chaque groupe de valeurs X, Y et Z forme un point de données.
Nuage de points 3D avec bâtons	Crée un graphique en 3D qui affiche les valeurs de trois variables numériques sous forme de point de données. Chaque groupe de valeurs X, Y et Z forme un point de données. Chaque point de données est relié au plan horizontal par une ligne appelée bâton.

Nuage de points : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la

source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Horizontal	La colonne que vous attribuez à cette fonction est la variable d'axe horizontal ou X du graphique.
Vertical	La colonne que vous attribuez à cette fonction est la variable d'axe vertical ou Y du graphique. Vous pouvez créer un récapitulatif statistique des valeurs uniques de la variable d'axe horizontal. Sélectionnez la variable sous la fonction Vertical. Sélectionnez Agréger pour chaque valeur horizontale distincte et sélectionnez la fonction d'agrégation. Par exemple, vos données d'entrée répertorient les ventes et les dépenses pour chaque pays et par région. Vous attribuez la variable Country à la fonction Horizontal et la variable Sales à la fonction Vertical. Si vous voulez représenter la somme des ventes pour chaque pays, activez la case à cocher, puis sélectionnez Somme dans la liste déroulante. La tâche ajoute les ventes régionales pour un pays et présente une seule valeur par pays dans le résultat final.
Vertical (Droite)	La colonne que vous attribuez à cette fonction est la variable d'axe vertical ou Y du côté droit du graphique. Cette variable est représentée par rapport à la variable d'axe X de sorte qu'un autre graphe soit généré dans le diagramme. Vous pouvez créer un récapitulatif statistique des valeurs uniques de la variable d'axe horizontal. Sélectionnez la variable sous la fonction Vertical. Sélectionnez Agréger pour chaque valeur horizontale distincte et sélectionnez la fonction d'agrégation. Par exemple, vos données d'entrée répertorient les ventes et les dépenses pour chaque pays et par région. Vous attribuez la variable Country à la fonction Horizontal et la variable Expenses à la fonction Vertical (Droite). Si vous voulez représenter la somme des dépenses pour chaque pays, activez la case à cocher, puis sélectionnez Somme dans la liste déroulante. La tâche ajoute les dépenses régionales pour un pays et présente une seule valeur par pays dans le résultat final. Remarque : cette fonction est disponible uniquement pour un nuage de points en 2D.
Profondeur	La colonne que vous attribuez à cette fonction est la variable de troisième dimension ou d'axe z du graphique.

Nom du rôle	Description
	Remarque : cette option est disponible uniquement pour les graphiques en trois dimensions.
Grouper les graphiques par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de graphes créés. Un graphique distinct est créé pour chaque valeur unique. Activez la case à cocher Produire la même mise à l'échelle pour tous les graphiques si vous voulez que l'échelle de chaque axe soit la même pour tous les graphiques créés. Par défaut, la plage de valeurs de chaque axe est basée sur les valeurs minimales et maximales des données représentées, et peuvent donc varier d'un graphique à l'autre. Si vous activez cette option, la plage de chaque axe est la même pour tous les graphiques.

Nuage de points : Définir les options d'apparence

Définir les options du nuage de points en 3D

Dans la liste, sous l'en-tête Apparence, cliquez sur **Nuage de points 3D** pour accéder à ces options.

Sélectionnez un type, une couleur et une taille de symbole et indiquez si vous souhaitez afficher des bâtons pour chaque point de données.

Utilisez les curseurs de défilement ou les contrôles de **Rotation** et d'**Inclinaison** pour définir l'angle de visualisation. Par défaut, l'angle de rotation est 70, et l'angle d'inclinaison est 70. Cliquez sur **Réinitialiser** pour rétablir l'angle d'affichage à ses paramètres par défaut.

Remarque : si vous désactivez la case à cocher **Indiquer les valeurs de rotation et d'inclinaison**, la rotation et l'inclinaison du graphique sont déterminés par le format de sortie graphique. Par défaut, cette option est sélectionnée et vous pouvez indiquer l'angle d'affichage.

Définir d'autres options de tracé

Dans la liste, sous l'en-tête Apparence, cliquez sur **Graphes** pour accéder à ces options.

Remarque : ces options ne sont disponibles que pour un nuage de points en 2D.

Pour contrôler l'apparence d'un graphe pour une variable spécifique :

- 1 Sélectionnez une variable dans la liste. Les variables répertoriées dans la liste sont celles que vous avez attribuées aux fonctions **Vertical** et **Vertical (Droite)**.
- 2 Dans la zone **Trait**, indiquez le style, l'épaisseur et la couleur du trait.
- 3 Sélectionnez une couleur de contour. Cette option spécifie la couleur pour ces éléments :
 - Contour des zones délimitées
 - Intervalle de confiance d'une analyse de régression
 - Traits ou contours des courbes avec écart-type
- 4 Dans la zone **Symbole pour un point de données**, spécifiez le type du symbole, le symbole à utiliser, la hauteur de chaque symbole, ainsi que la couleur des symboles des points de données.

Remarque : certains symboles ne sont pas pris en charge par tous les formats de sortie graphique. Si vous sélectionnez un symbole non pris en charge par le format de sortie, un autre symbole (pris en charge) s'affiche.

- 5 Indiquez si vous souhaitez interrompre les lignes de la courbe aux valeurs manquantes.

Remarque : la case à cocher **Interrompre la seconde ligne de la courbe aux valeurs manquantes** n'est disponible que si vous attribuez une variable à la fonction **Vertical (Droite)**.

Options d'apparence supplémentaires

Certaines options sont communes aux tâches graphiques. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Sélectionner la méthode d'interpolation" page 738](#)

- “Définir les options d’un axe” page 729
- “Ajouter des lignes de référence” page 732
- “Personnaliser l’aire” page 734

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S’il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.
.....

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d’une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l’heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),

- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Matrice nuage de points

<i>À propos de la tâche Matrice du nuage de points</i>	617
<i>Matrice nuage de points : Créer une matrice de nuages de points</i>	618

À propos de la tâche Matrice du nuage de points

La tâche Matrice du nuage de points crée un graphique de nuage de points pour plusieurs combinaisons de variables. Vous pouvez utiliser les options pour superposer les courbes d'ajustement et les ellipses dans vos nuages de points.

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	SGSCATTER
Produits SAS requis	SAS Base, SAS/GRAPH
Autres produits SAS recommandés	Néant

Matrice nuage de points : Créer une matrice de nuages de points

Nom d'option	Description
Libellé	Identifie la tâche dans l'arborescence du projet et dans le flux de processus. Dans la zone Libellé , saisissez le nom de la tâche.
Source de données	Permet de modifier la source des données en entrée pour la tâche. Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée. Le nom de la source de données en entrée apparaît dans le champ Source de données . Tout filtre spécifié dans la source de données apparaît dans le champ Filtre de tâche .
Variables de la matrice	spécifie les variables à représenter les unes par rapport aux autres pour la création des nuages de points. Vous devez attribuer au moins deux variables à cette fonction.
Variable de groupe	spécifie une variable de classification pour diviser les valeurs en groupes. L'attribution d'une variable à cette fonction n'est pas obligatoire.
Titre du graphique	Spécifie le titre qui apparaît dans les résultats. Par défaut, le titre des résultats est Matrice du nuage de points. Vous pouvez néanmoins personnaliser ce titre. Cliquez sur la flèche vers le bas pour modifier le texte du titre. Vous pouvez utiliser des macro-variables dans le titre. Les valeurs des macro-variables sont déterminées lorsque vous exécutez la tâche.
Diagonale	Ajoute des graphiques aux cellules de diagonales de la matrice. Vous pouvez placer les histogrammes dans les cellules de diagonales, et superposer ces histogrammes avec des courbes de densité normales ou des estimations de la densité du noyau. Vous pouvez également spécifier si la diagonale doit démarrer dans le coin supérieur gauche ou dans le coin inférieur gauche. Si vous ne sélectionnez pas de graphique à inclure dans la diagonale, les cellules de la diagonale contiennent les noms des variables.

Nom d'option	Description
Ellipse	Ajoute une ellipse de confiance ou de prédiction dans chaque cellule qui contient un nuage de points. Les deux types d'ellipse supposent une distribution normale à deux variables. Vous pouvez choisir parmi ces types d'ellipses : ■ Néant - aucune ellipse n'est ajoutée. ■ Moyenne - ajoute une ellipse de confiance pour la moyenne de la population. ■ Prédiction - ajoute une ellipse de prédiction pour une nouvelle observation. Vous pouvez aussi indiquer le niveau de confiance pour l'ellipse.
Légende	Indique si vous voulez afficher une légende pour la matrice du nuage de points. Les légendes ne sont disponibles que lorsque vous attribuez une variable de groupe. Lorsque vous attribuez une variable de groupe, une légende est affichée par défaut. Pour supprimer la légende, désactivez la case à cocher Afficher la légende. Spécifiez un titre pour la légende, puis sélectionnez une position pour la légende dans la liste Position. Vous pouvez également indiquer si vous souhaitez afficher une bordure autour de la légende.
Aire	Spécifie la hauteur et la largeur du graphe. Vous pouvez également indiquer si vous souhaitez afficher les info-bulles qui apparaissent lorsque vous déplacez le pointeur sur un point de données dans le nuage de points. Par défaut, l'aire a une dimension de 800 x 800, et aucune info-bulle n'est affichée.
Note de bas de page pour le graphique	Spécifie le pied de page qui apparaît dans les résultats. Par défaut, la date et l'heure de l'exécution de la tâche apparaissent dans le pied de page du graphe. (Dans le champ de saisie Note de bas de page pour le graphique, les macro-variables SAS pour la date et l'heure courantes s'affichent.) Cependant, vous pouvez personnaliser la note de bas de page pour vos résultats. Cliquez sur la flèche vers le bas pour modifier le texte de la note de bas de page. Vous pouvez utiliser des macro-variables dans le pied de page. Pour supprimer le pied de page, supprimez tout le texte dans le champ de saisie et cliquez sur OK. Aucun pied de page ne sera alors inclus dans les résultats.

Trier les données

<i>À propos de la tâche Trier les données</i>	621
<i>Trier les données : Attribuer des variables aux fonctions d'analyse</i>	622
<i>Trier les données : Définir les options de tri des données</i>	622
<i>Trier les données : Définition des options des résultats</i>	624
<i>Afficher les propriétés</i>	625

À propos de la tâche Trier les données

La tâche Trier les données permet de trier une table de votre projet en fonction de l'une de ses colonnes. Vous pouvez trier les données par une ou plusieurs colonnes et pouvez sélectionner l'ordre croissant ou décroissant pour chacune d'elle. La tâche Trier les données place la table triée en tant que nouvel élément de données dans votre projet.

Pour exécuter cette tâche, les prérequis sont :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	SORT
Produits SAS requis	Base SAS
Autres produits SAS recommandés	Néant

Trier les données : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Trier par	Lorsque vous attribuez une ou plusieurs variables à cette fonction, la table est groupée en fonction de la variable ou des variables sélectionnées. L'ordre dans lequel les variables apparaissent dans cette fonction détermine la variable utilisée comme clé de tri primaire, celle utilisée comme clé de tri secondaire, et ainsi de suite. La clé de tri primaire est toujours la première variable répertoriée dans la fonction Trier par .
Colonnes à supprimer	Lorsque vous attribuez une ou plusieurs variables à cette fonction, la sortie générée ne contient pas les variables spécifiées. Vous pouvez attribuer jusqu'à $(n - 1)$ variables à cette fonction, n étant le nombre total de variables triées de la table.

Trier les données : Définir les options de tri des données

Dans la liste, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Ordre de sortie	Vous pouvez déterminer l'ordre dans lequel les données apparaîtront dans la sortie en modifiant les options suivantes : ■ Indique l'ordre de classement à utiliser lors du tri des variables alphanumériques. Vous pouvez utiliser l'ordre de classement défini sur le serveur, ou bien sélectionner les ordres de classement ASCII ou EBCDIC, ou encore utiliser les standards nationaux danois, finlandais, norvégien ou suédois. Vous pouvez également trier les variables alphanumériques à l'aide d'un ordre de classement personnalisé, défini par votre site d'installation. Remarque : si vous sélectionnez Personnaliser, un ordre de tri national personnalisé doit être défini sur votre site. Votre administrateur peut vérifier si c'est le cas. Si aucun ordre de tri n'est disponible, l'ordre de classement par défaut défini pour le serveur, est utilisé. ■ Maintenir l'ordre des données d'origine au sein des groupements 'Trier par groupe les données selon l'ordre que vous avez défini pour les fonctions d'analyse. Si vous ne sélectionnez pas cette option, les données de sortie sont groupées dans un ordre indéfini au sein des groupes de clés triées.
Enregistrements en double	Vous pouvez indiquer comment les enregistrements en double doivent être traités dans la sortie à l'aide des options suivantes : ■ Conserver tous les enregistrements conserve tous les enregistrements de la table de sortie, y compris les doublons. ■ Éliminer les observations en double ayant les mêmes valeurs pour le groupe Trier par élimine les observations en double ayant les mêmes valeurs pour le groupe Trier par. Si l'option Maintenir l'ordre des données d'origine au sein des groupements 'Trier par' est sélectionnée, l'observation retenue pour chaque groupe Trier par sera la première lue dans la table d'origine. En revanche, si cette option n'est pas sélectionnée, l'observation conservée pour chaque groupe Trier par ne peut pas être prédéterminée. ■ Ne pas garder les enregistrements en double adjacents compare chaque enregistrement au précédent dans la table de sortie. En cas de correspondance exacte, l'enregistrement en double n'est pas copié dans la table de sortie.

Nom d'option	Description
	Remarque : si vous n'attribuez pas toutes les variables à la fonction Trier par, certains enregistrements en double ne seront peut-être pas supprimés.
Options de tri avancées	Vous pouvez définir les options avancées pour le tri des données dans la sortie. Vous pouvez modifier les options suivantes : ■ Mémoire de tri indique la quantité de mémoire maximale utilisable pour la tâche Trier les données. Vous pouvez indiquer la quantité de mémoire en octets (o), kilo-octets (Ko), mégaoctets (Mo) ou gigaoctets (Go). Vous pouvez également choisir d'utiliser la totalité de la mémoire disponible ou la quantité de mémoire par défaut allouée sur le serveur. ■ Réduire l'espace disque temporaire nécessaire indique que lors du processus de tri des données, seules les variables Trier par et les numéros d'observation sont stockés dans des fichiers temporaires, ce qui réduit la capacité de stockage nécessaire pour procéder au tri. Au cours de la phase finale du tri, le fichier temporaire est utilisé comme index pour accéder à la table d'origine et envoyer, dans le bon ordre, les données dans la table des résultats. ■ Forcer le tri des données indexées indique que vous voulez trier toutes les tables, y compris celles susceptibles d'avoir déjà été triées dans l'ordre souhaité ou contenant un index créé par l'utilisateur avec des clés qui reflètent celles spécifiées dans la fonction Trier par. Si vous spécifiez cette option, la table est triée quel que soit son ordre actuel, et qu'elle contienne ou non un index.

Trier les données : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Vous pouvez désigner un autre nom ou un autre emplacement de stockage pour la table de sortie. Vous pouvez également sauvegarder les enregistrements en double dans une table si vous choisissez d'éliminer les observations en double comportant les mêmes valeurs pour le groupe **Trier par** ou de ne conserver qu'un seul des

enregistrements dont les valeurs correspondent exactement. Pour plus d'informations, consultez les options concernant les **enregistrements en double**.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Fractionner des colonnes

<i>À propos de la tâche Fractionner des colonnes</i>	627
<i>Fractionner des colonnes : Attribuer des variables aux fonctions d'analyse</i>	629
<i>Fractionner des colonnes : Définition des options des résultats</i>	630
<i>Afficher les propriétés</i>	630

À propos de la tâche Fractionner des colonnes

La tâche Fractionner des colonnes crée une table de sortie en fractionnant la combinaison unique de valeurs des colonnes sélectionnées dans la table d'entrée en de multiples colonnes. Vous pouvez utiliser la table de sortie pour analyser individuellement les colonnes qui contiennent plusieurs lignes de la table d'entrée.

Cette tâche est utile lorsqu'une table comporte une colonne qui contient plusieurs observations pour différents sous-groupes et que vous voulez fractionner les mesures des sous-groupes dans des colonnes distinctes.

Par exemple, la table suivante contient une seule colonne avec les températures relevées mensuellement dans divers endroits de l'état de la Caroline du Nord :

	 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

La tâche Fractionner des colonnes permet de créer la table suivante en transposant les températures mensuelles de chaque endroit dans une colonne pour chaque mois :

	 Location	 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

Pour exécuter cette tâche, les prérequis sont :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	TRANSPOSE
Produits SAS requis	Base SAS
Autres produits SAS recommandés	SAS/STAT

Fractionner des colonnes : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Colonnes à fractionner	La colonne que vous attribuez à cette fonction contient les valeurs qui sont transposées dans les nouvelles colonnes. Vous ne devez attribuer qu'une seule variable à cette fonction.
Colonne de l'identifiant de la valeur	La colonne que vous attribuez à cette fonction fournit le nom des nouvelles colonnes contenant les valeurs transposées de Colonnes à fractionner. L'ordre des valeurs uniques de cette colonne identifie quelles lignes seront transposées dans les nouvelles colonnes. Si vous n'attribuez pas de variables de groupe, cette colonne ne doit contenir qu'un ensemble de valeurs uniques, et les valeurs de Colonnes à fractionner sont transposées dans une ligne de la table de sortie. Vous devez attribuer une colonne à la fonction Colonne de l'identifiant de la valeur. Pour sélectionner une variable de groupe, attribuez une colonne à la fonction Grouper l'analyse par.
Colonne du libellé de la valeur	La colonne que vous attribuez à cette fonction fournit les libellés des nouvelles colonnes contenant les valeurs transposées de Colonnes à fractionner. Remarque : pour voir l'effet de cette colonne, affichez les propriétés des nouvelles colonnes dans la table de sortie.
Grouper l'analyse par	Chaque variable que vous attribuez à cette fonction est utilisée pour segmenter les lignes de Colonne à fractionner en sous-groupes qui sont transposés dans les nouvelles colonnes. Chaque sous-groupe, défini par une combinaison unique de valeurs des variables de groupe, devient une ligne dans la table de sortie. Le

Nom du rôle	Description
	nombre de lignes formant un sous-groupe est déterminé par le nombre de valeurs uniques dans la Colonne de l'identifiant de la valeur .

Fractionner des colonnes : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Vous pouvez attribuer un autre nom à la table de sortie.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur **Parcourir**.

Vous pouvez ajouter un préfixe aux noms des colonnes de la table de sortie. Par défaut, les noms des colonnes sont créés à partir des valeurs uniques de la **Colonne de l'identifiant de la valeur**. Pour indiquer un préfixe, activez la case à cocher **Utiliser le préfixe** et saisissez-le dans le champ de saisie. Le préfixe par défaut est **Colonne**.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche

- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Empiler colonnes

<i>À propos de la tâche Empiler des colonnes</i>	633
<i>Empiler des colonnes : Attribuer des variables aux fonctions d'analyse</i>	635
<i>Empiler des colonnes : Définir les options des noms de colonne</i>	636
<i>Empiler des colonnes : Définition des options des résultats</i>	637
<i>Afficher les propriétés</i>	637

À propos de la tâche Empiler des colonnes

La tâche Empiler colonnes crée une table de sortie en restructurant les colonnes sélectionnées dans la table d'entrée pour que ces colonnes soient transposées dans les observations. Vous pouvez utiliser la table de sortie pour analyser les valeurs de plusieurs colonnes de la table d'entrée. Si vous groupez les observations, les colonnes sélectionnées sont divisées en sous-groupes qui sont basés sur les combinaisons uniques des valeurs de groupe. Chaque sous-groupe forme une ligne de la table de sortie.

Cette tâche est utile si vous disposez d'une table dans laquelle chaque observation contient le même type de données dans plusieurs colonnes et que vous voulez analyser ces données sur plusieurs colonnes. Par exemple, la table suivante contient les températures relevées mensuellement à divers endroits de l'état de la Caroline du Nord :

	 Location	 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

La structure des données permet de calculer des statistiques descriptives pour chaque mois, basées sur les différents endroits de l'état. Vous ne pouvez cependant pas calculer les statistiques descriptives pour un endroit en particulier, car les statistiques doivent être calculées sur une observation.

Pour calculer des statistiques sur toutes les températures ou sur chaque endroit, vous pouvez empiler les colonnes avec les températures mensuelles par endroit afin de créer la table de sortie suivante :

	 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

Pour exécuter cette tâche, les prérequis sont :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	TRANSPOSE

Nom du prérequis	Noms des procédures et des produits
Produits SAS requis	Base SAS
Autres produits SAS recommandés	SAS/STAT

Empiler des colonnes : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Colonnes à empiler	Chaque colonne que vous attribuez à cette fonction devient une ou plusieurs lignes de la table de sortie. Lorsque vous sélectionnez une ou plusieurs variables de groupe, ces variables sont utilisées pour segmenter chaque colonne en sous-colonnes, dont chacune est transformée en une ligne. Dans ce cas, une colonne est transposée dans un nombre de lignes égal au nombre de groupes définis par les variables de groupe. Vous devez attribuer au moins une colonne à la fonction Colonnes à empiler. Pour sélectionner une variable de groupe, attribuez une colonne à la fonction Grouper l'analyse par. Remarque : si vous n'attribuez pas de variables de groupe, la table d'entrée doit contenir une seule observation et chaque colonne devient une ligne de la table de sortie.
Grouper l'analyse par	Chaque variable que vous attribuez à cette fonction est utilisée pour segmenter les colonnes à empiler dans des sous-groupes qui sont transposés séparément. Chaque sous-groupe, défini par une combinaison unique de valeurs des variables de groupe, devient une ligne de la table de sortie. Chaque combinaison unique des valeurs de la variable de groupe n'existe qu'une fois.

Empiler des colonnes : Définir les options des noms de colonne

Cliquez sur **Noms de colonne** dans la liste pour indiquer les noms et libellés des nouvelles colonnes de la table de sortie.

Nom du rôle	Description
Indiquer un nom et un libellé pour la nouvelle colonne	Vous créez une nouvelle colonne en transposant les valeurs d'une observation de la table d'entrée dans les valeurs d'une variable de la table de sortie. Pour indiquer un en-tête pour la colonne de sortie contenant ces valeurs transposées, saisissez un en-tête dans la zone Nom de la zone Nouvelle colonne de valeurs. Au départ, la zone Nom contient l'en-tête par défaut, Valeurs empilées. Pour indiquer une description pour la colonne qui est stockée avec la table de sortie, saisissez un libellé dans la zone Libellé de la zone Nouvelle colonne de valeurs. La description peut comporter 256 caractères maximum. Laissez ce champ à blanc si vous ne voulez pas de libellé.
Indiquer un nom et un libellé pour la colonne dérivée	Chaque ligne de la table de sortie contient le nom de la variable de la table d'entrée qui est transposée pour créer l'observation en cours. Afin d'indiquer un en-tête pour la colonne de sortie contenant ces noms de variables, saisissez l'en-tête dans la zone Nom de la zone Colonne dérivée. Au départ, la zone Nom contient l'en-tête par défaut, Source de la valeur. Afin d'indiquer une description pour la colonne qui est stockée avec la table de sortie, saisissez un libellé dans la zone Libellé de la zone Colonne dérivée. La description peut comporter 256 caractères maximum. Laissez ce champ à blanc si vous ne voulez pas de libellé.
Indiquer un nom et un libellé pour la colonne Description	Chaque ligne de la table de sortie contient le libellé de la variable de la table d'entrée à laquelle les valeurs de cette sortie appartiennent. Afin d'indiquer un en-tête pour la colonne de sortie contenant ces noms de variables, saisissez l'en-tête dans la zone Nom de la zone Colonne Description. Au départ, la zone Nom contient l'en-tête par défaut, Description de la valeur. Afin d'indiquer une description pour la colonne qui est stockée avec la table de sortie, saisissez un libellé dans la zone Libellé de la zone Colonne Description. La description peut comporter 256 caractères maximum. Laissez ce champ à blanc si vous ne voulez pas de libellé.

Nom du rôle	Description
	Remarque : pour voir le résultat de l'indication d'un libellé pour une colonne, affichez les propriétés de la nouvelle colonne dans la table de sortie.

Empiler des colonnes : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Vous pouvez attribuer un autre nom à la table de sortie.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur **Parcourir**.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),

- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Normaliser les données

<i>À propos de la tâche Normaliser les données</i>	639
<i>Normaliser les données : Attribuer des variables aux fonctions d'analyse</i>	640
<i>Normaliser les données : Définir les options de normalisation</i>	641
<i>Normaliser les données : Définition des options des résultats</i>	641
<i>Afficher les propriétés</i>	642

À propos de la tâche Normaliser les données

La tâche Normaliser des données normalise les données d'une table SAS par rapport à une moyenne et un écart-type donnés et crée une nouvelle table SAS qui contient les données standardisées.

Par exemple, vous pouvez utiliser cette tâche pour normaliser les scores de test en spécifiant une moyenne et un écart-type.

Pour exécuter cette tâche, les prérequis sont :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	STANDARD
Produits SAS requis	Base SAS
Autres produits SAS recommandés	Néant

Normaliser les données : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables d'analyse	effectue des normalisations sur chaque variable que vous attribuez à cette fonction. Vous devez attribuer au moins une variable à cette fonction.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Poids relatif	spécifie les valeurs à utiliser pour calculer une moyenne pondérée et une variance pondérée pour les variables d'analyse. Vous ne pouvez attribuer qu'une seule variable à cette fonction.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : Vous ne pouvez pas grouper la table en fonction d'une variable déjà attribuée à une autre fonction.

Normaliser les données : Définir les options de normalisation

Cliquez sur **Normaliser** dans la liste pour accéder à ces options.

Nom d'option	Description
Normaliser avec	Par défaut, cette tâche normalise les données à une moyenne de 0 et un écart-type de 1. Vous pouvez spécifier une nouvelle moyenne et un nouvel écart-type pour la ou les variables que vous analysez, ou bien choisir de conserver la moyenne ou l'écart-type en cours. Remarque : les mêmes moyenne et écart-type sont appliqués à toutes les variables d'analyse. Vous pouvez choisir de garder la moyenne et l'écart-type de chaque variable en désactivant les cases à cocher Nouvelle moyenne ou Nouvel écart-type.
Remplacer valeurs manquantes par la valeur de moy. de la col.	Par défaut, les valeurs manquantes d'une variable d'analyse d'entrée sont exclues du processus de normalisation. Vous pouvez remplacer les valeurs manquantes d'entrée par la moyenne des valeurs non manquantes de cette variable en activant la case à cocher Remplacer valeurs manquantes par la valeur de moyenne de la col.

Normaliser les données : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Par défaut, l'option **Ajouter les nouvelles variables normalisées** est activée afin que les variables normalisées soient incluses dans la table de sortie avec les variables d'origine. Pour remplacer les variables d'origine par les variables normalisées, sélectionnez **Remplacer les variables existantes**.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre

les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Statistiques descriptives

<i>À propos de la tâche Statistiques descriptives</i>	643
<i>Statistiques descriptives : Attribuer des variables aux fonctions d'analyse</i>	644
<i>Statistiques descriptives : Sélectionner des statistiques descriptives</i>	645
Sélectionner les statistiques descriptives de base	645
Sélectionner les statistiques descriptives du centile	646
Sélectionner des statistiques descriptives supplémentaires	646
<i>Statistiques descriptives : Définir les options de tracé</i>	647
<i>Statistiques descriptives : Définition des options des résultats</i>	648
<i>Indiquer des titres et des notes de bas de page</i>	650
<i>Afficher les propriétés</i>	651

À propos de la tâche Statistiques descriptives

La tâche Statistiques descriptives fournit des outils permettant de calculer des statistiques descriptives pour les variables sur toutes les observations et dans les groupes d'observations. Vous pouvez également agréger vos données sous la forme d'un graphique tel qu'un histogramme.

Par exemple, vous pouvez utiliser cette tâche pour créer un rapport sur les nouvelles ventes par type de produit et par pays.

Pour exécuter cette tâche, les prérequis sont :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	MEANS, UNIVARIATE
Produits SAS requis	Base SAS

Nom du prérequis	Noms des procédures et des produits
Autres produits SAS recommandés	SAS/GRAPH

Statistiques descriptives : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables d'analyse	spécifie la variable sur laquelle vous voulez que les statistiques soient générées. Vous devez attribuer au moins une variable à cette fonction.
Variables de classification	spécifie les variables à utiliser pour diviser les données d'entrée en catégories ou sous-groupes. Ces variables peuvent être alphanumériques ou numériques. Les statistiques sont calculées sur toutes les variables d'analyse sélectionnées pour chaque combinaison unique de variables de classification. Vous pouvez activer l'option Autoriser formats multi-libellés pour que les libellés de formats principaux et secondaires d'un intervalle donné, et d'intervalles se chevauchant, soient utilisés pour créer des combinaisons de sous-groupes lorsque le format d'un multi-libellé est attribué à cette variable.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.

Nom du rôle	Description
Poids relatif	spécifie la variable à utiliser pour calculer les moyennes pondérées, les variances et les sommes.
Copier des variables	Spécifie les variables à inclure dans la table de sortie. Vous pouvez utiliser l'option Valeur à copier pour la fonction Copier des variables pour copier dans la table de sortie la valeur maximale ou minimale de la variable dans les observations correspondantes des données d'entrée.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe.

Statistiques descriptives : Sélectionner des statistiques descriptives

Sélectionner les statistiques descriptives de base

Sous l'en-tête Statistiques de la liste, cliquez sur **Base** pour accéder à ces options.

Par défaut, le rapport Statistiques descriptives comprend les statistiques suivantes : Moyenne, Ecart-type, Minimum, Maximum et Nombre d'observations. Vous pouvez contrôler les statistiques à calculer, le nombre de décimales à afficher dans le rapport, ainsi que la méthode à utiliser pour calculer les statistiques du quantile.

Nom d'option	Description
Statistiques de base	spécifie les statistiques de base à inclure dans l'analyse.
Nombre max. de décimales	spécifie le nombre maximal de décimales à utiliser dans les statistiques calculées. Par défaut, une statistique est affichée selon le meilleur ajustement, qui est généralement de sept décimales.
Diviseur pour écart-type et variance	précise le diviseur utilisé pour le calcul des écarts-types et des variances. Par défaut, le diviseur est le degré de liberté.

Sélectionner les statistiques descriptives du centile

Sous l'en-tête Statistiques de la liste, cliquez sur **Centiles** pour accéder à ces options.

Nom d'option	Description
Statistiques du centile	spécifie les centiles à calculer. Vous pouvez sélectionner le calcul du quartile inférieur (25ème centile), la médiane (50ème centile) ainsi que le quartile supérieur (75ème centile). Par défaut, aucun centile n'est sélectionné.
Méthode du quantile	indique la méthode utilisée pour calculer les quantiles, la médiane et les centiles. Vous avez le choix entre ces deux méthodes : ■ Classer les statistiques pour utiliser une méthode qui enregistre toutes les données et les trie en fonction des valeurs uniques. ■ Algorithme par morceaux de courbe ou parabolique pour utiliser une méthode consommant moins de mémoire et qui s'approche du quantile.

Sélectionner des statistiques descriptives supplémentaires

Sous l'en-tête Statistiques de la liste, cliquez sur **Autres** pour accéder à ces options.

Vous pouvez également choisir de calculer ces statistiques :

- **Intervalle de confiance de la moyenne**
- **Statistique t et probabilité > |t|**
- **Coefficient de variation**
- **Somme des carrés corrigée**
- **Somme des carrés non corrigée**

Vous pouvez aussi calculer le niveau de confiance de l'intervalle de confiance de la moyenne.

Statistiques descriptives : Définir les options de tracé

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Remarque : si une variable a été attribuée à la fonction **Variable de classification**, les graphiques sont générés uniquement pour la combinaison à n critères des variables de classification.

Nom du graphe	Description
Histogramme	L'histogramme est un graphique utilisé pour déterminer la distribution d'un ensemble de données. Vous pouvez indiquer les milieux dans l'histogramme. Les options disponibles dans la liste déroulante varient selon que vous créez un histogramme ou un histogramme comparatif. Si vous attribuez une variable à la fonction Variables de classification, vous créez un histogramme comparatif. Les options disponibles sont alors les suivantes : ■ Spécifier les milieux permet d'indiquer les milieux de début et de fin. Vous pouvez également spécifier l'intervalle d'escalier à utiliser pour calculer les autres milieux. Le milieu correspond à la valeur située au milieu de la plage de valeurs que représente chaque barre. ■ Clé détermine les milieux des données de la cellule clé. Par défaut, les niveaux de la variable de classification s'affichent de haut en bas (de gauche à droite) dans l'ordre croissant des valeurs internes (non formatées) de la première variable de classification. Si vous n'avez spécifié qu'une seule de ces variables, la cellule clé correspond au niveau qui apparaît le premier dans cet ordre. Si vous en avez spécifié deux, la cellule clé correspond à une combinaison des niveaux des variables 1 et 2 qui apparaissent en premier dans cet ordre. Ainsi, le choix de la cellule clé détermine l'axe horizontal uniforme utilisé pour les autres cellules. La liste des milieux de la cellule clé est donc étendue dans l'une ou l'autre direction si besoin jusqu'à ce qu'elle couvre les données dans les cellules restantes. ■ Uniforme détermine le nombre de milieux en fonction de la taille de l'échantillon complet.

Nom du graphe	Description
	Si vous n'attribuez pas de variable à la fonction Variables de classification, vous créez un histogramme. Pour les histogrammes, ces options sont disponibles : ■ Spécifier les milieux permet d'indiquer les milieux de début et de fin. Vous pouvez également spécifier l'intervalle d'escalier à utiliser pour calculer les autres milieux. Le milieu correspond à la valeur située au milieu de la plage de valeurs que représente chaque barre. ■ Indiquer le nombre de classes permet d'indiquer le nombre de classes (également appelées intervalles d'histogrammes) des données. Par défaut, le nombre de classes affichées dépend de l'étendue des données. Cependant, vous pouvez indiquer un maximum.
Boîte à moustaches	Les boîtes à moustaches représentent la mesure de la position centrale (la médiane), deux mesures de dispersion (l'intervalle et l'écart interquartile), le coefficient d'asymétrie (Skewness) (de l'orientation de la médiane relative aux quartiles), ainsi que les valeurs hors norme potentielles. Les boîtes à moustaches sont particulièrement utiles pour la comparaison de deux ensembles de données.

Statistiques descriptives : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Nom d'option	Description
Enregistrer les statistiques dans la table	spécifie si les statistiques doivent être enregistrées dans une table de données. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de

Nom d'option	Description
	stockage différent pour vos données de sortie., cliquez sur Parcourir .
Sélectionnez les statistiques à inclure dans les résultats	spécifie les statistiques à inclure dans les résultats. Vous avez le choix entre ces options : ■ Valeur à copier pour la fonction Copier des variables - si vous attribuez une variable à la fonction Copier des variables, vous pouvez choisir de copier dans la table de sortie la valeur maximale ou minimale de la variable dans les observations correspondantes de la table d'entrée. ■ Combinaisons des variables de classification - la fonction Variable de classification permet de grouper les statistiques descriptives des variables de classification dans des tables. Vous pouvez créer des tables à n critères de classification, n étant le nombre de variables de classification. Les groupes sont déterminés par les combinaisons uniques des variables de classification. L'ordre des variables dans la table dépend de l'ordre des variables dans la fonction Variables de classification. Vous pouvez utiliser les options suivantes pour diviser les données de sortie en sous-groupes : □ N critères de classification uniquement crée une table unique affichant les statistiques descriptives de la combinaison à n critères des variables de classification. Supposons, par exemple, que vous voulez comparer les ventes résultant de deux campagnes marketing différentes. Vous attribuez la variable Profit à la fonction Variables d'analyse et les variables Month, Marketing Campaign et Region à la fonction Variables de classification. Si vous sélectionnez ensuite l'option N critères de classification uniquement, vous obtenez la table de Month*Campaign*Region. □ Tous les critères crée une table de statistiques descriptives distincte pour chaque combinaison unique de variables de classification. Par exemple, l'utilisation des variables Month, Campaign et Region, puis la sélection de l'option Tous les cas possibles donnent les tables suivantes : ■ une table des ventes totales. Cela s'appelle une table à zéro critère de classification, c'est-à-dire ne comprenant pas les variables de classification. ■ une table à un critère de classification affichant les statistiques descriptives pour chaque mois (par ex., May, June et July) ■ une table à un critère de classification affichant les statistiques descriptives pour chaque région (par ex., north, south, east et west)

Nom d'option	Description
	■ une table à un critère de classification affichant les statistiques descriptives pour chaque type de campagne Marketing (par ex., TV commercial, magazine ad) ■ une table à deux critères de classification affichant les statistiques descriptives de Month*Region ■ une table à deux critères de classification affichant les statistiques descriptives de Month*Campaign ■ une table à deux critères de classification affichant les statistiques descriptives de Region*Campaign ■ une table à trois critères de classification affichant les statistiques descriptives de Month*Campaign*Region □ Indiquer les critères crée une table de statistiques descriptives distincte pour chaque combinaison unique des variables de classification que vous désignez. Par exemple, si vous saisissez 1, 2 ou 3 dans la zone Indiquer les critères, les résultats contiennent toutes les tables à un, deux ou trois critères de classification possibles. Si vous spécifiez 0 dans ce champ, les résultats contiennent une table à zéro critère de classification, qui n'inclut pas les variables de classification.
Afficher les libellés d'analyse	afficher les libellés des variables d'analyse.
Supprimer l'affichage des résultats	n'affiche pas les résultats. Si vous choisissez de supprimer la sortie affichée, vous devez enregistrer vos résultats dans une table de sortie.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.

- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.
.....

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

.....
Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.
.....

Assistant Statistiques descriptives

<i>À propos de l'assistant Statistiques descriptives</i>	653
<i>Assistant Statistiques descriptives : Sélectionner les données</i>	653
<i>Assistant Statistiques descriptives : Attribuer des variables aux fonctions</i>	654
<i>Assistant Statistiques descriptives : Indiquer les statistiques du rapport et les résultats de la sortie</i>	655
<i>Assistant Statistiques descriptives : Indiquez le titre et les notes de bas de page</i>	656

À propos de l'assistant Statistiques descriptives

L'Assistant **Statistiques descriptives** permet de créer des tables de statistiques descriptives ainsi que des histogrammes et des boîtes à moustaches, avec des variables de groupe facultatives.

Assistant Statistiques descriptives : Sélectionner les données

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la

source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Cliquez sur **Suivant** pour attribuer des variables aux fonctions.

Assistant Statistiques descriptives : Attribuer des variables aux fonctions

- 1 Avant d'attribuer des variables aux fonctions, vous devez sélectionner les données.
- 2 Attribuez au moins une variable à la fonction **Statistiques descriptives de (Variable d'analyse)**. Les variables que vous attribuez à cette fonction sont les variables numériques sur lesquelles vous voulez générer des statistiques.

.....

Remarque : vous devez attribuer au moins une variable à cette fonction. Toutes les autres fonctions sont facultatives.

.....

- 3 Attribuez au moins une variable à la fonction **Pour chaque valeur de (Variable de classification)**. Les variables que vous attribuez à cette fonction sont des variables alphanumériques ou numériques discrètes, utilisées pour diviser les données d'entrée par catégories ou par sous-groupes. Les statistiques sont calculées sur toutes les variables d'analyse sélectionnées pour chaque combinaison unique de variables de classification.
- 4 Attribuez au moins une variable à la fonction **Tables distinctes pour les valeurs de (Variable de groupe)**. Les variables que vous attribuez à cette fonction sont utilisées pour calculer des statistiques différentes pour chaque valeur distincte ou combinaison de valeurs des variables **Grouper l'analyse par**. Les données sont automatiquement triées sur les variables de cette fonction avant le calcul des statistiques. Pour sélectionner un ordre de tri pour chaque variable, utilisez les flèches vers le haut et vers le bas.
- 5 (Facultatif) Cliquez sur **Avancées** pour affecter des variables à ces fonctions :
 - **Indiquer la fréquence par : (Nombre d'occurrences)** - lorsque vous attribuez une variable à cette fonction, chaque observation de la table est supposée représenter n observations, où n est la valeur du nombre d'occurrence pour cette ligne. Les statistiques sont calculées en conséquence. Vous ne pouvez attribuer qu'une seule variable à cette fonction.
 - **Analyse de poids par : (Poids relatif)** - si vous attribuez une variable à cette fonction, la valeur de la variable pour chaque observation est utilisée pour calculer les moyennes pondérées, les variances et les sommes. Vous ne pouvez attribuer qu'une seule variable à cette fonction.

- **Inclure les libellés de : (Copier variable)** - les variables que vous attribuez à cette fonction sont incluses en tant que variables supplémentaires dans la table de sortie.
- 6 Cliquez sur **Suivant** pour indiquer des titres et des notes de bas de page.

Assistant Statistiques descriptives : Indiquer les statistiques du rapport et les résultats de la sortie

- 1 Avant d'indiquer statistiques du rapport et les résultats de la sortie, vous devez attribuer les variables aux fonctions.
- 2 Sélectionnez les statistiques à inclure dans les résultats. Par défaut, la moyenne, l'écart-type, le minimum, le maximum et le nombre d'observations sont inclus. Vous pouvez inclure des statistiques supplémentaires en utilisant les options de ces onglets :

Classique

inclut les options descriptives. Vous pouvez également sélectionner le nombre maximal de décimales pour les statistiques calculées. Par défaut, une statistique est affichée selon le meilleur ajustement, qui est généralement de sept décimales.

Centile

spécifie les centiles à calculer. Vous pouvez sélectionner le calcul du quartile inférieur (25ème centile), la médiane (50ème centile) ainsi que le quartile supérieur (75ème centile). Par défaut, aucun centile n'est sélectionné.

Utilisez l'option **Méthode du quantile** pour sélectionner la méthode à utiliser pour calculer les quantiles, la médiane et les centiles. Vous avez le choix entre ces deux méthodes :

- **Classer les statistiques** pour utiliser une méthode qui enregistre toutes les données et les trie en fonction des valeurs uniques.
- **Algorithme par morceaux de courbe ou parabolique** pour utiliser une méthode consommant moins de mémoire et qui s'approche du quantile.

Supplémentaires

comprend ces statistiques :

- **Intervalle de confiance de la moyenne**
- **Statistique t et probabilité > |t|**
- **Coefficient de variation**
- **Somme des carrés corrigée**

■ Somme des carrés non corrigée

Vous pouvez aussi calculer le niveau de confiance de l'intervalle de confiance de la moyenne.

- 3 Sélectionnez l'option **Afficher les statistiques** pour afficher les résultats. Vous devez choisir entre afficher les statistiques dans les résultats ou bien les enregistrer de façon permanente dans une table de données, ou faire les deux.
- 4 Activez la case à cocher **Histogramme** ou **Boîte à moustaches** pour sélectionner le type de graphique que vous voulez générer. Vous pouvez en sélectionner un seul ou bien les deux. Par défaut, aucun graphe n'est sélectionné.

.....

Remarque : si une variable a été attribuée à la fonction **Variable de classification**, les graphiques sont générés uniquement pour la combinaison à n critères des variables de classification.

.....

- 5 Activez la case à cocher **Enregistrer les statistiques dans une table** pour enregistrer la sortie. Vous devez choisir entre afficher les statistiques dans les résultats ou bien les enregistrer de façon permanente dans une table de données, ou faire les deux.

.....

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

.....

- 6 Cliquez sur **Suivant** pour indiquer des titres et des notes de bas de page.

Assistant Statistiques descriptives : Indiquez le titre et les notes de bas de page

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

- 1 Avant d'indiquer le titre et les notes de bas de page, vous devez indiquer les statistiques du rapport et les résultats de la sortie.

- 2 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie. Vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

.....
Remarque : certains titres peuvent ne pas être disponibles, selon les graphiques que vous avez sélectionnés pour les résultats.
.....

- 3 Cliquez sur **Terminer** pour exécuter la tâche.

Tables de synthèse

<i>À propos de la tâche Tables de synthèse</i>	659
<i>Tables de synthèse : Attribuer des variables aux fonctions d'analyse</i>	660
<i>Tables de synthèse : Définir la table de synthèse</i>	661
Créer la table de synthèse	661
Personnaliser l'apparence de la table	661
<i>Tables de synthèse : Définition des options des résultats</i>	664
<i>Indiquer des titres et des notes de bas de page</i>	664
<i>Afficher les propriétés</i>	665

À propos de la tâche Tables de synthèse

La tâche Tables de synthèse présente des statistiques descriptives sous forme de tableau, en utilisant tout ou partie des variables d'une table. Vous pouvez créer différentes tables, allant de la plus simple à la plus sophistiquée.

Par exemple, supposons que vous disposiez d'une table SAS qui contient des données sur les ventes d'un produit pour deux types de client, les particuliers et les professionnels, dans des départements distincts de deux régions. Vous pouvez utiliser la tâche Tables de synthèse pour additionner les ventes pour chaque type de client dans chacun des départements, par région.

Pour exécuter cette tâche, les prérequis sont :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	TABULATE
Produits SAS requis	Base SAS
Autres produits SAS recommandés	Néant

Tables de synthèse : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables d'analyse	spécifie les variables pour l'analyse. Vous devez attribuer au moins une variable à cette fonction ou à la fonction Variables de classification .
Variables de classification	spécifie les variables alphanumériques ou numériques discrètes à utiliser pour diviser les données d'entrée en catégories ou sous-groupes. Les statistiques sont calculées sur toutes les variables d'analyse sélectionnées pour chaque combinaison unique de variables de classification. Vous pouvez activer l'option Formats multi-libellés pour que les libellés de formats principaux et secondaires d'un intervalle donné, et d'intervalles se chevauchant, soient utilisés pour créer des combinaisons de sous-groupes lorsque le format d'un multi-libellé est attribué à cette variable.
Pages	génère un tableau récapitulatif distinct pour chaque valeur distincte de cette variable.
Poids relatif	spécifie la variable à utiliser pour calculer les moyennes pondérées, les variances et les sommes.
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.

Tables de synthèse : Définir la table de synthèse

Créer la table de synthèse

Dans la liste, cliquez sur **Tables de synthèse** pour accéder à ces options.

Pour définir une table de synthèse, vous devez décider quelles informations vous souhaitez faire apparaître et comment. La tâche Tables de synthèse est très flexible ; vous pouvez organiser les variables de classification et d'analyse de la façon dont vous le souhaitez.

Pour inclure une variable :

- 1 Sélectionnez la variable dans la zone **Variables disponibles**.
- 2 Faites glisser la variable dans une cellule dans la partie supérieure ou inférieure de la case **Aperçu**. Si vous lâchez la nouvelle variable sur une variable existante, la nouvelle variable remplace l'ancienne. Pour ajouter la nouvelle variable aux variables existantes, lâchez-la sur la bordure d'une cellule. Lorsque le pointeur de la souris est correctement placé, la bordure devient bleue.

Remarque : par défaut, la statistique appliquée aux variables d'analyse que vous utilisez est le nombre d'observations (N). Vous pouvez ajouter ou remplacer des statistiques en utilisant les mêmes méthodes décrites ci-dessus.

Personnaliser l'apparence de la table

Faites un clic droit sur la table d'aperçu pour définir des options permettant de personnaliser l'apparence de la table de synthèse. À l'exception de **Supprimer des cellules**, chaque option du menu contextuel ouvre une fenêtre permettant de définir différentes options.

Nom du graphe	Description
Supprimer des cellules	Cette option est activée lorsque vous faites un clic droit sur une variable d'analyse, une variable de classification ou une statistique de la table d'aperçu. Elle permet de retirer la variable ou la statistique de la table de synthèse.

Nom du graphe	Description
Propriétés de la table	Cette option est toujours active. Les options définies dans la boîte de dialogue Propriétés de la table sont appliquées à l'ensemble de la table de synthèse. ■ Onglet Général : Saisissez un libellé pour les valeurs manquantes. La valeur par défaut est un point ("."). Utilisez les cases à cocher pour inclure les valeurs manquantes en tant que valeurs légitimes des variables de classification et pour ne pas afficher les en-têtes (libellés) des variables de classification de ligne. ■ Onglet Format : Attribuez un format SAS standard ou un format créé par l'utilisateur à toutes les valeurs de données de la table de synthèse. Sélectionnez le format, la largeur des champs et le nombre de positions décimales. ■ Onglet Police : Indiquez la police, la taille, le style, la couleur de premier plan, la couleur d'arrière-plan, la justification et l'alignement vertical de l'ensemble du texte de la table.
Propriétés de la case	Cette option est toujours active. Elle permet de définir des options pour l'angle supérieur gauche de la table. Ces options remplacent celles éventuellement sélectionnées dans la boîte de dialogue Propriétés de la table. ■ Onglet Général : Saisissez le texte à utiliser dans la case. Si vous avez attribué une ou plusieurs variables à la fonction Pages, vous pouvez choisir d'afficher les en-têtes de page dans la case. Vous pouvez en outre choisir de renvoyer à la ligne le texte de la case ou de le tronquer si sa longueur est supérieure à celle de la case. ■ Onglet Format : Cet onglet ne s'applique pas à la case. ■ Onglet Police : Indiquez la police, la taille, le style, la couleur de premier plan, la couleur d'arrière-plan, la justification et l'alignement vertical du texte de la case.
Propriétés de l'en-tête	Pour activer cette option, vous devez faire un clic droit sur un en-tête de colonne ou de ligne. Elle permet de définir des options pour des en-têtes de colonne ou de ligne particuliers. Ces options remplacent celles éventuellement sélectionnées dans la boîte de dialogue Propriétés de la table. ■ Onglet Général : Saisissez un libellé pour l'en-tête de ligne ou de colonne. Vous pouvez en outre choisir de renvoyer à la ligne le texte de l'en-tête ou de le tronquer si sa longueur est supérieure à l'espace prévu pour l'en-tête. ■ Onglet Format : Cet onglet ne s'applique pas aux valeurs de variable de classification. Pour attribuer un format d'en-tête à une variable de classification, cliquez sur Données dans la liste et choisissez la variable de classification. Pour attribuer le format, sélectionnez l'option Format de l'en-tête.

Nom du graphe	Description
	■ Onglet Police : Indiquez la police, la taille, le style, la couleur de premier plan, la couleur d'arrière-plan, la justification et l'alignement vertical du texte d'en-tête.
Propriétés du niveau de classe	Cette option n'est active que si vous faites un clic droit sur une variable de classification dans la table d'aperçu. Elle permet de définir des options pour les valeurs de variable de classification. ■ Onglet Général : Indiquez si le texte de la valeur de la variable de classification doit être renvoyé à la ligne ou non. ■ Onglet Format : Cet onglet ne s'applique pas aux valeurs de variable de classification. Pour attribuer un format d'en-tête à une variable de classification, cliquez sur Données dans la liste et choisissez la variable de classification. Pour attribuer le format, sélectionnez l'option Format de l'en-tête. ■ Onglet Police : Indiquez la police, la taille, le style, la couleur de premier plan, la couleur d'arrière-plan, la justification et l'alignement vertical du texte de la valeur de la variable de classification.
Propriétés de la valeur de données	Cette option n'est active que si vous faites un clic droit sur une variable ou une statistique dans la table d'aperçu. Elle permet de définir des options pour les valeurs de données de la ligne ou de la colonne sélectionnée dans la table de synthèse. Ces options remplacent celles éventuellement sélectionnées dans la boîte de dialogue Propriétés de la table. ■ Onglet Général : Indiquez si le texte de la valeur de données doit être renvoyé à la ligne ou non. ■ Onglet Format : Attribuez un format SAS standard ou un format créé par l'utilisateur à toutes les valeurs de données de la ligne ou de la colonne sélectionnée dans la table de synthèse. Sélectionnez le format, la largeur des champs et le nombre de positions décimales. ■ Onglet Police : Indiquez la police, la taille, le style, la couleur de premier plan, la couleur d'arrière-plan, la justification et l'alignement vertical du texte des valeurs de données.

Tables de synthèse : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Pour enregistrer vos résultats dans une table de sortie, activez la case à cocher **Enregistrer les résultats dans une table**.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Assistant Tables de synthèse

<i>À propos de l'assistant Tables de synthèse</i>	667
<i>Assistant Tables de synthèse : Sélectionner les données</i>	667
<i>Assistant Tables de synthèse : Attribuer des variables aux fonctions</i>	668
<i>Assistant Tables de synthèse : Sélectionner les variables de classification</i>	669
<i>Assistant Tables de synthèse : Spécifier les totaux</i>	669
<i>Assistant Tables de synthèse : Sélectionner une autre sortie</i>	670
<i>Assistant Tables de synthèse : Indiquez le titre et les notes de bas de page</i>	670

À propos de l'assistant Tables de synthèse

L'Assistant **Tables de synthèse** permet de créer différentes configurations de table, notamment des tables à double entrée, avec des options pour les statistiques de synthèse.

Assistant Tables de synthèse : Sélectionner les données

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Cliquez sur **Suivant** pour attribuer les variables d'analyse et les statistiques.

Assistant Tables de synthèse : Attribuer des variables aux fonctions

- 1 Avant d'attribuer des variables aux fonctions, vous devez sélectionner les données.
- 2 Pour ajouter une variable dans la zone **Variables d'analyse**, cliquez sur **Ajouter**. Dans le menu contextuel, sélectionnez la variable à ajouter dans cette fonction. Les variables que vous affectez à ce rôle sont les variables numériques pour les calculs.
- 3 Dans la zone **Variables d'analyse**, sélectionnez la statistique à calculer pour chaque variable d'analyse. Par défaut, la somme est calculée. Pour calculer une autre statistique, sélectionnez dans la colonne **Statistique** la cellule correspondant à la statistique que vous voulez modifier. Dans la liste déroulante, sélectionnez la statistique à calculer.

Remarque : vous ne pouvez pas attribuer de variable d'analyse si les données ne contiennent aucune variable numérique. Dans ce cas, la statistique de fréquence est utilisée à la place d'une variable d'analyse.

- 4 Dans la liste déroulante **Libellés des variables d'analyse**, sélectionnez un emplacement pour les libellés des variables d'analyse. Les options disponibles sont les suivantes : **dans les colonnes**, **dans les lignes**, **sur les pages** et **masqué**.
- 5 Dans la liste déroulante **Libellés des statistiques**, sélectionnez un emplacement pour les libellés des statistiques. Les options disponibles sont les suivantes : **dans les colonnes**, **dans les lignes**, **sur les pages** et **masqué**.
- 6 Cliquez sur **Suivant** pour indiquer des titres et des notes de bas de page.

Assistant Tables de synthèse : Sélectionner les variables de classification

- 1 Avant de sélectionner les variables de classification, vous pouvez attribuer les variables d'analyse et les statistiques.
- 2 Pour ajouter une variable de classification dans les colonnes, lignes ou pages, cliquez sur **Ajouter** dans la zone appropriée. Dans le menu contextuel, sélectionnez la variable à ajouter. Les variables de classification que vous attribuez sont des variables alphanumériques ou numériques discrètes qui permettent de diviser les données d'entrée par catégories. Les statistiques seront calculées sur toutes les variables d'analyse sélectionnées pour chaque combinaison unique de variables de classification. Une table de synthèse distincte est générée pour chaque variable que vous attribuez à la zone **Pages**.
- 3 Cliquez sur **Autres options** pour personnaliser la structure de la table. Dans la boîte de dialogue **Structure de la table**, vous pouvez masquer les en-têtes des colonnes, des lignes et des pages. Vous pouvez également imbriquer ou concaténer les colonnes, les lignes et les tables.
- 4 Cliquez sur **Suivant** pour indiquer des titres et des notes de bas de page.

Assistant Tables de synthèse : Spécifier les totaux

- 1 Avant de spécifier les totaux, vous devez spécifier les variables de classification.
- 2 Utilisez les listes déroulantes **Colonnes**, **Lignes** et **Pages** pour indiquer si vous voulez afficher le total général pour chaque colonne, ligne ou page. Pour afficher le total, sélectionnez **Total général uniquement**. Si vous ne souhaitez pas l'afficher, sélectionnez **Néant**. Pour les pages, vous pouvez également indiquer si vous voulez afficher les totaux au début ou à la fin de la page.

Remarque : si vous n'avez pas attribué de variable de classification aux colonnes, lignes ou pages, l'option correspondante n'est pas disponible.

- 3 Saisissez le libellé du total dans la zone **Libellé des totaux**.
- 4 Cliquez sur **Suivant** pour indiquer des titres et des notes de bas de page.

Assistant Tables de synthèse : Sélectionner une autre sortie

- 1 Avant de sélectionner une autre sortie, vous pouvez spécifier les totaux.
- 2 Activez la case à cocher **Enregistrer les résultats dans une table**.

.....
Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

.....

- 3 Cliquez sur **Suivant** pour indiquer des titres et des notes de bas de page.

Assistant Tables de synthèse : Indiquez le titre et les notes de bas de page

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

- 1 Avant de pouvoir spécifier le titre et les notes de bas de page, vous pouvez spécifier si les résultats doivent être enregistrés dans une table de sortie.
- 2 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie. Vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

.....
Remarque : certains titres peuvent ne pas être disponibles, selon les graphiques que vous avez sélectionnés pour les résultats.

.....

- 3 Cliquez sur **Terminer** pour exécuter la tâche.

Surface

<i>À propos de la tâche Surface</i>	671
<i>Surface : Sélectionner un type de graphe</i>	672
<i>Surface : Attribuer des variables aux fonctions d'analyse</i>	673
<i>Surface : Définir les options d'apparence</i>	674
Définir les options de la surface	674
Options supplémentaires	674
Lisser les données	674
<i>Indiquer des titres et des notes de bas de page</i>	675
<i>Afficher les propriétés</i>	676

À propos de la tâche Surface

La tâche Surface crée des courbes de structure filaire tridimensionnelles, des courbes lissées tridimensionnelles et des courbes de gradient tridimensionnelles qui présentent les relations mathématiques entre trois variables numériques. Dans le graphique en 3D, une variable verticale (z) est tracée pour une position du plan spécifiée par deux variables horizontales (x et y). Les coordonnées de chaque point correspondent à la valeur des trois variables numériques dans une observation de la table d'entrée. L'observation peut contenir des valeurs sous la forme $z = f(x,y)$ ou des valeurs indépendantes telles que l'altitude pour une longitude et une latitude données.

Pour plus d'informations, voir "[Sélectionner un format de sortie pour les graphiques](#)" page 6.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	G3D

Nom du prérequis	Noms des procédures et des produits
Produits SAS requis	SAS Base, SAS/GRAPH
Autres produits SAS recommandés	Aucun

Surface : Sélectionner un type de graphe

Dans la liste, cliquez sur **Courbe de surface** pour accéder à ces options.

Nom du rôle	Description
Surface 3D en fil de fer	Crée un graphique en 3D dans lequel la structure filaire affiche les relations mathématiques entre trois variables.
Surface 3D avec lissage	Crée un graphique en 3D dans lequel la surface est entièrement de la même couleur.
Surface du gradient 3D avec bandes	Crée un graphique en 3D dans lequel la couleur indique un changement de valeur. Les régions sont clairement définies.
Surface du gradient lissé 3D	Crée un graphique en 3D dans lequel la couleur indique un changement de valeur. Les régions ne sont pas aussi clairement définies que dans une courbe de surface de gradient avec bandes.

Le format de sortie graphique sélectionné détermine le mode d'affichage de la surface.

ActiveX et Image ActiveX
représente tous les types de courbe de surface.

Java et Image Java
représente les courbes de surface de structure filaire tridimensionnelles. Les courbes de type surface 3D avec lissage, surface du gradient 3D avec bandes et surface du gradient lissé 3D sont représentées comme des courbes de surface du gradient lissé 3D.

GIF, JPEG, PNG et SAS EMF
représente tous les types de courbe de surface comme des courbes de surface de structure filaire 3D.

Surface : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Horizontal	La colonne que vous attribuez à cette fonction est la variable d'axe horizontal ou X du graphique.
Vertical	La colonne que vous attribuez à cette fonction est la variable d'axe vertical ou Y du graphique.
Profondeur	La colonne que vous attribuez à cette fonction est la variable de troisième dimension ou d'axe z du graphique. Remarque : cette option est disponible uniquement pour les graphiques en trois dimensions.
Grouper les graphiques par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de graphes créés. Un graphique distinct est créé pour chaque valeur unique.

Surface : Définir les options d'apparence

Définir les options de la surface

Dans la liste, sous l'en-tête Apparence, cliquez sur **Surface** pour accéder à ces options.

Sélectionnez une couleur pour le haut (orienté vers le Z positif) et le bas (orienté vers le Z négatif) du graphique.

Indiquez si vous voulez tracer les côtés du graphique.

Utilisez les curseurs de défilement ou les contrôles de **Rotation** et d'**Inclinaison** pour définir l'angle de visualisation. Par défaut, l'angle de rotation est 70, et l'angle d'inclinaison est 70. Cliquez sur **Réinitialiser** pour rétablir l'angle d'affichage à ses paramètres par défaut. Note

Remarque : si vous désactivez la case à cocher **Indiquer les valeurs de rotation et d'inclinaison**, la rotation et l'inclinaison du graphique sont déterminés par le format de sortie graphique. Par défaut, cette option est sélectionnée et vous pouvez indiquer l'angle d'affichage.

Pour les courbes de surface de gradient avec bandes 3D et de gradient lissé 3D, vous pouvez indiquer si un contour des données doit être tracé avec la surface. Vous pouvez afficher le contour au-dessus ou au-dessous de la surface.

Options supplémentaires

Certaines options sont communes aux tâches graphiques. Pour plus d'informations sur ces options, voir ces rubriques :

- [“Définir les options d'un axe” page 729](#)
- [“Personnaliser l'aire” page 734](#)

Lisser les données

Dans la liste, sous l'en-tête Apparence, cliquez sur **Données de la grille** pour accéder à ces options.

Vous pourrez avoir à créer une grille rectangulaire de valeurs interpolées ou lissées dans les cas suivants :

- Les points des données de votre table d'entrée sont espacés de façon irrégulière
- La table d'entrée comporte un nombre de lignes insuffisant pour pouvoir créer une courbe de surface à trois dimensions
- Les données de votre table d'entrée sont bruyantes.

Par défaut, aucune méthode d'interpolation n'est utilisée. Cependant, vous pouvez sélectionner ces méthodes d'interpolation :

Jointure

Utilise une interpolation linéaire au sein d'un ensemble de régions triangulaires constitué à partir de la table d'entrée. Cette méthode d'interpolation crée des valeurs dans la plage des valeurs initiales de la variable verticale, mais la surface résultant de l'interpolation peut ne pas être lisse.

Sans mise à l'échelle

Indique que les variables X et Y ne seront pas mises à la même échelle avant l'interpolation. Par défaut, l'interpolation est réalisée une fois que les deux variables sont mises à l'échelle de façon similaire, car les méthodes d'interpolation supposent que les échelles de X et Y sont comparables.

Partiel

Indique qu'un spline sera utilisé pour estimer les dérivées de l'interpolation polynomiale biquintique. Une courbe spline à deux variables est ajustée par rapport aux voisins les plus proches et est utilisée pour estimer les dérivées nécessaires.

Spline

Indique qu'un spline bivarié sera utilisé pour l'interpolation. Cette option aboutit sur l'utilisation d'un algorithme d'ordre n^3 , où n est le nombre de données d'entrée. Cette méthode peut être longue.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Test t

<i>À propos de la Tâche Test t</i>	677
<i>Test t : Choisir le type du test t</i>	678
<i>Test t : Attribuer des variables aux fonctions d'analyse</i>	678
<i>Test t : Définir les options de l'analyse</i>	680
<i>Test t : Générer des graphes</i>	680
<i>Indiquer des titres et des notes de bas de page</i>	681
<i>Afficher les propriétés</i>	682

À propos de la Tâche Test t

La tâche Test t réalise des tests t pour les observations à un échantillon, à deux échantillons ou appariées. Le test t à un échantillon compare la moyenne de l'échantillon à une valeur donnée. Le test t à deux échantillons compare la moyenne du premier échantillon moins celle du second échantillon à une valeur donnée. Le test t à observations appariées compare la moyenne des différences entre les observations à une valeur donnée.

Il est parfois impossible de supposer l'existence de deux groupes de données indépendants et l'appariement naturel des données. Dans ce cas, le test t apparié est avantageux. Par exemple, supposons l'examen d'un stimulus en vue d'en déterminer l'effet sur la pression systolique. L'étude porte sur douze hommes dont la pression systolique est mesurée avant et après l'application du stimulus. La tâche Test t permet de déterminer si la modification moyenne de la pression systolique est sensiblement différente de zéro.

Pour exécuter cette tâche, les prérequis sont :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	TTEST

Nom du prérequis	Noms des procédures et des produits
Produits SAS requis	SAS Base, SAS/STAT
Autres produits SAS recommandés	SAS/GRAPH

Test t : Choisir le type du test t

Dans la liste, cliquez sur **Type du test t** pour accéder à ces options.

Nom du prérequis	Noms des procédures et des produits
À deux échantillons	compare la moyenne du premier échantillon moins la moyenne du second avec un nombre donné.
Pour échantillons appariés	compare la moyenne des différences dans les observations avec un nombre donné.
À un échantillon	compare la moyenne de l'échantillon avec un nombre donné.

Test t : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Variables de classification	Disponible pour les tests t à deux échantillons. Attribuez une variable de classification à cette fonction pour l'analyse de deux échantillons indépendants. La variable ne doit avoir que deux niveaux, un par échantillon. Vous ne pouvez attribuer qu'une seule variable à cette fonction.
Variables d'analyse	Disponible pour les tests t à un et à deux échantillons. Les tests t sont effectués sur chaque variable de cette liste. Vous devez attribuer au moins une variable à cette fonction.
Variables appariées	Disponible avec les tests t pour échantillons appariés. Pour réaliser un test t pour échantillons appariés, attribuez les deux variables à analyser à cette fonction. Du fait que les tests t pour échantillons appariés sont réalisés en soustrayant chaque valeur de la seconde variable à partir de la valeur correspondante de la première variable, l'ordre des variables est important. Remarque : vous devez attribuer deux variables précisément à cette fonction.
Grouper l'analyse par	Lorsque vous attribuez une ou plusieurs variables à cette fonction, la table est triée en fonction de la(des) variable(s) sélectionnée(s) et un test t est effectué pour chacun des groupes. Remarque : vous ne pouvez pas grouper les tests t en fonction d'une variable que vous avez déjà sélectionnée pour les Variables d'analyse ou pour les Variables appariées.
Nombre d'occurrences	Lorsque vous attribuez une variable à cette fonction, chaque ligne de la table est supposée représenter n observations, n étant le nombre d'occurrences de cette ligne. Vous ne pouvez attribuer qu'une seule variable à cette fonction.
Poids relatif	Lorsque vous attribuez une variable à cette fonction, la valeur de la variable de chaque observation est utilisée pour calculer les statistiques pondérées. Vous ne pouvez attribuer qu'une seule variable à cette fonction. Remarque : si la table d'entrée de cette tâche contient uniquement des statistiques descriptives, la fonction Poids relatif n'est pas disponible. Lorsque la table d'entrée contient des statistiques descriptives, l'option de poids est sans effet.

Test t : Définir les options de l'analyse

Dans la liste, cliquez sur **Analyse** pour accéder à ces options.

Nom d'option	Description
Indiquer une hypothèse nulle	Par défaut, la valeur de l'hypothèse nulle est 0. Pour définir une nouvelle valeur d'hypothèse nulle, entrez la nouvelle valeur dans la zone $H_{0=}$.
Définir les options de l'intervalle de confiance de l'écart-type	Définissez les options de l'intervalle de confiance en procédant comme suit : ■ La case à cocher Répartition symétrique est activée par défaut. L'intervalle de confiance à répartition symétrique place un montant identique de chaque queue de la distribution du Chi-deux. Cette option requiert l'affichage d'un intervalle de confiance à répartition symétrique. ■ Activez la case à cocher UMPU pour aboutir à un intervalle de confiance déduit du test de type UMPU. Remarque : si vous ne voulez pas afficher l'intervalle de confiance, désactivez les cases à cocher Répartition symétrique et UMPU.

Test t : Générer des graphes

Dans la liste, cliquez sur **Graphes** pour accéder à ces options.

Nom du graphe	Description
Graphique de synthèse	génère un graphique de synthèse. Le contenu du graphique de synthèse varie selon le type de test t que vous avez choisi d'effectuer. ■ Pour les tests t à un échantillon, le graphique de synthèse contient un histogramme avec densités normale et du noyau superposées, une boîte à moustaches et une bande représentant l'intervalle de confiance.

Nom du graphe	Description
	■ Pour les tests t à deux échantillons, il contient des histogrammes comparatifs avec densités superposées et des boîtes à moustaches. ■ Pour les tests t appariés, il contient un histogramme, des densités, une boîte à moustaches et un intervalle de confiance de la différence ou du rapport.
Histogramme	génère un histogramme ou des histogrammes comparatifs avec densités normale et du noyau superposées. Pour les tests t à un échantillon et les tests t appariés, l'histogramme et les densités sont basés sur le critère de test (qui correspond à la différence moyenne ou au rapport d'une expérience appariée). Pour les tests t à deux échantillons, le graphique contient des histogrammes comparatifs (un par classe).
Boîte à moustaches	génère une boîte à moustaches ou des boîtes à moustaches comparatives. La boîte s'étend du 25ème centile (quartile inférieur) au 75ème centile (quartile supérieur). La ligne verticale dans la boîte représente la médiane. Pour les tests t à un échantillon et appariés, un intervalle de confiance de la moyenne apparaît sous forme de bande à l'arrière plan. Pour les tests t à deux échantillons, le graphique contient des boîtes à moustaches comparatives (une par classe).
Courbe d'intervalle de confiance	génère des graphes de l'intervalle de confiance des moyennes.
Graphique Q-Q (quantile-quantile) normal	génère un graphique Q-Q (quantile-quantile) normal. Pour les tests t à deux échantillons, des graphiques distincts apparaissent pour chaque classe dans un seul et même panel.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.
.....

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

.....
Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.
.....

Analyse d'une table

<i>À propos de la tâche Analyse d'une table</i>	683
<i>Analyse d'une table : Attribuer des variables aux fonctions d'analyse</i>	684
<i>Analyse d'une table : Définir des tables</i>	685
<i>Analyse d'une table : Sélectionner les statistiques des cellules</i>	685
<i>Analyse d'une table : Spécifier les statistiques des tables</i>	687
Effectuer des tests d'association	687
Effectuer des tests de concordances	688
Effectuer des tests de différences ordonnées	689
Effectuer des tests de tendance	690
Définir les options de calcul	690
<i>Analyse d'une table : Définition des options des résultats</i>	692
Enregistrer les statistiques d'une cellule	692
Enregistrer les statistiques d'une table	693
<i>Indiquer des titres et des notes de bas de page</i>	693
<i>Afficher les propriétés</i>	694

À propos de la tâche Analyse d'une table

Cette tâche permet de générer des tables à double entrée, également connues sous le nom de tables de contingence, à partir de vos données.

Par exemple, supposons que vous vouliez comparer la probabilité d'affections coronariennes pour deux types de régime. Une table SAS contient des données pour une étude cas-témoins d'un régime riche en matières grasses sur le risque d'affections coronariennes. Les données sont enregistrées dans les cellules, la variable *Compte* contenant les fréquences de chaque combinaison de régime et de risque. En utilisant la tâche *Analyse d'une table*, vous pouvez créer une table de contingence 2x2 qui affiche les valeurs des fréquences pour chaque combinaison de régime (riche et pauvre en matières grasses) et de risques (risque fort et faible).

Pour exécuter cette tâche, les prérequis sont :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	FREQ
Produits SAS requis	Base SAS
Autres produits SAS recommandés	Néant

Analyse d'une table : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Nombre d'occurrences	spécifie la variable qui contient le nombre de fréquences pour chaque ligne de la table. Chaque fonction est supposée représenter n observations, où n est la valeur de la fréquence pour cette ligne.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : Vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez attribuée à d'autres fonctions.
Variables de table	utilise ces variables pour créer des tables de fréquences et à double entrée de deux à n critères. Des statistiques sont calculées pour les variables de table. Vous devez attribuer au moins deux variables à cette fonction.

Analyse d'une table : Définir des tables

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Pour générer une nouvelle table :

- 1 Dans la zone **Variables autorisées dans la table**, sélectionnez la variable que vous souhaitez ajouter à la table. Les variables de la liste ont été attribuées à la fonction **Variables de table**.
- 2 Faites glisser la variable sélectionnée sur la case **Aperçu**. Les valeurs de la première variable que vous faites glisser comprennent les variables de la table à double entrée. Les valeurs de la seconde variable que vous faites glisser comprennent les lignes de la table à double entrée.

Si vous faites glisser plus de deux variables vers la zone **Aperçu**, plusieurs tables à double entrée seront générées, une pour chaque combinaison de valeurs pour la troisième variable (et variables suivantes) que vous faites glisser.

Remarque : une fois la variable ajoutée dans la table, elle est supprimée dans la zone **Variables autorisées dans la table**.

Pour supprimer une variable dans la table, sélectionnez la variable dans la zone **Aperçu**, puis faites-la glisser sur **Variables autorisées dans la table**.

Pour démarrer une nouvelle table, sélectionnez **sélectionner pour commencer la définition d'une nouvelle table** dans la liste des tables définies.

Analyse d'une table : Sélectionner les statistiques des cellules

Cliquez sur **Statistiques d'une cellule** pour accéder à ces options.

Les statistiques sélectionnées sont affichées dans chaque cellule de la table à double entrée.

Nom d'option	Description
Pourcentages cumulés de la colonne	affiche le pourcentage cumulé de la colonne de chaque cellule.
Pourcentages de la ligne	affiche le pourcentage de la ligne de chaque cellule.
Pourcentages de la colonne	affiche le pourcentage de la colonne de chaque cellule. Par défaut, ces statistiques sont incluses dans la sortie.
Fréquences de la colonne	affiche le nombre d'occurrences pour chaque cellule. Par défaut, ces statistiques sont incluses dans la sortie.
Pourcentages de la cellule	affiche le pourcentage, le pourcentage de ligne et le pourcentage de colonne des tables à double entrée, ou bien les pourcentages et les pourcentages cumulés des tables de fréquences à n critère de classification.
Fréquences de la valeur manquante	affiche les fréquences des valeurs manquantes.
Contribution de la cellule au Khi-2 de Pearson	affiche la contribution de chaque cellule aux statistiques totales du Chi-deux de Pearson.
Écart absolu de la cellule	affiche l'écart entre la fréquence de chaque cellule et la valeur prédite.
Fréquence de la cellule requise	affiche la fréquence prédite pour chaque cellule.
Pourcentage de la fréquence totale	affiche le pourcentage de la fréquence totale dans les tables à n critères de classification lorsque $n > 2$.
Inclure les pourcentages dans la table	affiche les pourcentages pour les colonnes et les lignes de la table de sortie.

Analyse d'une table : Spécifier les statistiques des tables

Effectuer des tests d'association

Sous l'en-tête Statistiques de la table de la liste, cliquez sur **Association** pour accéder à ces options.

Remarque : notez que si vous sélectionnez **P-value exacte**, le calcul de la p -value exacte peut nécessiter une grande quantité de mémoire et prendre beaucoup de temps. Utilisez l'option **Délai du calcul** pour spécifier la durée maximale de chaque calcul de p -value exacte.

Nom d'option	Description
Tests d'association	Vous pouvez spécifier : ■ Tests du Chi-deux calcule les tests d'homogénéité ou d'indépendance du Chi-deux et les mesures d'association basées sur le Chi-deux. Ces tests comprennent les Chi-deux de Pearson, du rapport de vraisemblance et de Mantel-Haenszel. Pour les tables 2×2, ils incluent le test exact de Fisher et le Chi-deux ajusté de la continuité. ■ p-values exactes calcule les p-values exactes des statistiques suivantes : test de critère qualitatif de lissage du Khi-2 pour les tables à un critère de classification et tests du Khi-2 de Pearson, du rapport de cotes et de Mantel-Haenszel pour les tables à deux critères de classification. ■ Test exact de Fisher calcule le test exact de Fisher pour les tables plus étendues que 2×2. Ce test est également appelé test de Freeman-Halton.
Mesures d'association	Vous pouvez spécifier : ■ Mesures calcule plusieurs mesures d'association et leurs erreurs types asymptotiques (ASE). Ces mesures incluent les coefficients gamma, de corrélation tau-b de Kendall, tau-c de Stuart, D (C/R) de Somers, D (R/C) de Somers, de Pearson et de Spearman, lambda (symétriques et asymétriques) et d'incertitude (symétriques et asymétriques).

Nom d'option	Description
	Pour les tables 2×2, ces mesures indiquent également le rapport de cotes, le risque relatif des colonnes 1 et 2, ainsi que l'intervalle de confiance correspondant. ■ P-values exactes et intervalle de confiance pour le rapport de cotes calcule les <i>p</i>-values exactes pour les corrélations de Pearson et de Spearman, ainsi que l'intervalle de confiance du rapport de cotes pour les tables 2×2. ■ Tester que la mesure est égale à zéro calcule des tests asymptotiques pour les coefficients gamma, tau-<i>b</i> de Kendall, tau-<i>c</i> de Stuart, <i>D</i> (<i>C/R</i>) de Somers, <i>D</i> (<i>R/C</i>) de Somers et de corrélation de Pearson et de Spearman. ■ Différences de risque pour les tables 2 x 2 fournit les estimations des risques (ou proportions binomiales) et des différences de risque pour les tables 2×2. Cette analyse peut être utile en cas de comparaison de la proportion de certaines caractéristiques pour deux groupes, les lignes 1 et 2 correspondant aux deux groupes, et les colonnes correspondant à deux caractéristiques ou résultats possibles. ■ Risque relatif pour les tables 2 x 2 calcule les mesures du risque relatif et leur intervalle de confiance pour les tables 2×2. Ces mesures incluent le rapport de cotes et les risques relatifs des colonnes 1 et 2.
Statistiques de Cochran-Mantel-Haenszel	Pour calculer toutes les statistiques de Cochran-Mantel-Haenszel, activez la case à cocher Statistiques CMH. Ces statistiques incluent la statistique de corrélation CMH, les scores de moyenne de ligne (ANOVA), les risques relatifs ajustés et les rapports de cotes. Le type de score par défaut est Tab1e. Pour changer le type de score, utilisez l'option Type de score.

Effectuer des tests de concordances

Sous l'en-tête Statistiques de la table de la liste, cliquez sur **Concordances** pour accéder à ces options.

Vous pouvez sélectionner ces tests et mesures de concordances pour les tables *n*×*n*.

Nom d'option	Description
Mesures	calcule des tests et des mesures de concordance de classification pour les tables carrées. Cette option propose le test de McNemar pour les tables 2x2 et le test de symétrie de Bowker pour les tables comportant plus de deux catégories de réponses. Elle produit

Nom d'option	Description
	également les coefficients kappa simples et pondérés, les erreurs types asymptotiques des kappas simples et pondérés, ainsi que l'intervalle de confiance correspondant. En présence de plusieurs niveaux et de deux catégories de réponses, cette option calcule également le test Q de Cochran.
P-values exactes	calcule les p-values exactes du test de McNemar pour les tables 2×2, ainsi que les coefficients kappa simples et pondérés. Remarque : notez que si vous sélectionnez P-value exacte, le calcul de la p-value exacte peut nécessiter une grande quantité de mémoire et prendre beaucoup de temps. Utilisez l'option Délai du calcul pour spécifier la durée maximale de chaque calcul de p-value exacte.
Tester que la mesure est égale à zéro	calcule les tests asymptotiques pour les coefficients kappa simples et pondérés.

Effectuer des tests de différences ordonnées

Dans la liste, sous l'en-tête Statistiques de la table, cliquez sur **Différence ordonnée** pour accéder à ces options.

Nom d'option	Description
Test de Jonckheere-Terpstra	calcule le test de Jonckheere-Terpstra, qui est un test non paramétrique des différences triées entre classes. Il teste l'hypothèse nulle selon laquelle la distribution de la variable de réponse ne diffère pas parmi les classes.
P-values exactes	calcule la p-value exacte pour le test de Jonckheere-Terpstra. Remarque : notez que si vous sélectionnez P-value exacte, le calcul de la p-value exacte peut nécessiter une grande quantité de mémoire et prendre beaucoup de temps. Utilisez l'option Délai du calcul pour spécifier la durée maximale de chaque calcul de p-value exacte.

Effectuer des tests de tendance

Sous l'en-tête Statistiques de la table de la liste, cliquez sur **Test de tendance** pour accéder à ces options.

Nom d'option	Description
Test de Cochran-Armitage	calcule le test de tendance de Cochran-Armitage, qui teste la tendance en proportions binomiales à tous les niveaux d'un facteur ou d'une covariable unique. Ce test convient pour une table de contingence dont une variable possède deux niveaux et l'autre variable est ordinale. La variable à deux niveaux représente la réponse, tandis que l'autre représente une variable explicative avec des niveaux classés.
P-values exactes	calcule les p-values exactes pour le test de tendance de Cochran-Armitage. Remarque : notez que si vous sélectionnez P-value exacte, le calcul de la p-value exacte peut nécessiter une grande quantité de mémoire et prendre beaucoup de temps. Utilisez l'option Délai du calcul pour spécifier la durée maximale de chaque calcul de p-value exacte.

Définir les options de calcul

Sous l'en-tête Statistiques de la table de la liste, cliquez sur **Options de calcul** pour accéder à ces options.

Nom d'option	Description
Type de score	Si vous effectuez des tests d'association, de concordances ou de tendance, sélectionnez un type de score. Le type de score correspond au type des scores de ligne et de colonne à utiliser avec la statistique du Chi-deux de Mantel-Haenszel, les statistiques de Cochran-Mantel-Haenszel, la corrélation de Pearson, le test de tendance de Cochran-Armitage et le coefficient kappa pondéré. Vous avez le choix entre ces options : ■ Table - pour les variables numériques, les scores de table sont les valeurs des niveaux ligne et colonne. Si les variables de ligne ou de colonne sont formatées, le score de la table est la valeur numérique interne correspondant à ce niveau. Si plusieurs valeurs numériques sont classées dans le même niveau formaté,

Nom d'option	Description
	la valeur numérique interne de ce niveau est la plus petite de ces valeurs. Pour les variables alphanumériques, les scores de table sont définis en tant que numéros de ligne et de colonne, c'est-à-dire 1 pour la première ligne, 2 pour la deuxième ligne et ainsi de suite. ■ Rang - Les scores de rang sont définis par les scores de ligne et les scores de colonne. ■ RIDIT - les scores Ridit sont calculés à partir des scores de rang. ■ RIDIT modifié - les scores RIDIT modifiés (MODRIDIT) représentent les valeurs attendues des statistiques d'ordre de la distribution uniforme sur (0,1). Ils sont calculés à partir des scores de rang.
Calculs exacts	Si vous calculez les p-values exactes dans l'un de vos tests, cette opération peut consommer une quantité importante de mémoire et de temps de traitement. Dans la zone Calculs exacts, activez la case Délai limite de calcul et spécifiez un délai limite de calcul (en secondes) pour chaque p-value de chaque table. La valeur par défaut est 900 secondes. Pour calculer l'estimation Monte Carlo des p-values exactes au lieu de calculer directement ces valeurs, activez la case à cocher Utiliser l'estimation Monte Carl. L'estimation Monte Carlo peut être utile pour résoudre des problèmes importants dont les calculs exacts exigent beaucoup de temps et de mémoire, mais pour lesquels des approximations asymptotiques risquent de ne pas être suffisantes.
Inclure les valeurs manquantes dans les calculs	Vous pouvez choisir d'inclure les valeurs manquantes dans les calculs des pourcentages et des autres statistiques.
Niveau de confiance	Vous pouvez indiquer un niveau de confiance pour les calculs. Par défaut, le niveau de confiance est de 95 pour cent.
Classer les valeurs par	Sélectionnez l'ordre dans lequel les valeurs des variables des tables à double entrée sont affichées. Classer les valeurs par peut être défini par les options suivantes : ■ Ordre de la table pour classer les valeurs en fonction de leur ordre d'apparence dans la table ■ Valeurs formatées pour classer les valeurs dans l'ordre croissant de leur valeur formatée ■ Fréquences décroissantes pour classer les valeurs de sorte que les niveaux ayant le plus grand nombre d'occurrences apparaissent en premier

Nom d'option	Description
	■ Valeurs non formatées pour classer les valeurs en fonction de leurs valeurs non formatées. Il s'agit de la présentation par défaut.

Analyse d'une table : Définition des options des résultats

Enregistrer les statistiques d'une cellule

Sous l'en-tête Résultats de la liste, cliquez sur **Résultats des statistiques de cellule** pour accéder à ces options.

Nom d'option	Description
Enregistrer les données de sortie	Activez la case à cocher de la ou des tables à double entrée voulues. La table répertoriée correspond à la variable attribuée à la fonction Variables de table. Dans la zone Table de statistiques des cellules, indiquez la table dans laquelle les statistiques de cellule de la table sélectionnée seront enregistrées. Le nom de la table est différent pour chaque table que vous sélectionnez. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Supprimer l'affichage des résultats	N'affiche pas les résultats.

Enregistrer les statistiques d'une table

Sous l'en-tête Résultats de la liste, cliquez sur **Résultats des statistiques de table** pour accéder à ces options.

Nom d'option	Description
Enregistrer les statistiques pour la table suivante	enregistre les statistiques d'une table à double entrée. Dans la zone Table de statistiques pour, indiquez la table dans laquelle les statistiques de table de la table sélectionnée seront enregistrées. Le nom de la table est différent pour chaque table que vous sélectionnez. Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone Tâches > Bibliothèque de sortie de la boîte de dialogue Options et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve. SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie., cliquez sur Parcourir.
Affiche les tables sous forme de liste	affiche toute la table à entrée multiple dans une seule table, plutôt que d'afficher une table à double entrée pour chacun des niveaux. Remarque : l'option Affiche les tables sous forme de liste est ignorée si vous sélectionnez une option qui génère une statistique sur la table.
Supprimer l'affichage des résultats	n'affiche pas les résultats.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.
.....

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

.....
Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.
.....

Diagramme en mosaïque

<i>À propos de la tâche Diagramme en mosaïque</i>	695
<i>Diagramme en mosaïque : Attribuer des variables aux fonctions d'analyse</i>	696
<i>Diagramme en mosaïque : Définir les options d'apparence</i>	697
Définir l'apparence d'un diagramme en mosaïque	697
Définir les options de légende	699
<i>Diagramme en mosaïque : Définir les options avancées</i>	699
<i>Indiquer des titres et des notes de bas de page</i>	701

À propos de la tâche Diagramme en mosaïque

La tâche Diagramme en mosaïque vous permet de créer ce type de graphique de manière à pouvoir visualiser une grande quantité de données hiérarchiques dans un espace limité. Chaque combinaison de catégories unique est représentée par une mosaïque rectangulaire dont la taille et la couleur sont déterminées par les variables de réponse. Ces mosaïques sont disposées selon une organisation hiérarchique.

Un diagramme en mosaïque est une arborescence rectangulaire. Le diagramme en mosaïque étant affiché sous forme d'arborescence, il est particulièrement utile dans un environnement interactif tel que le Web. Voici les avantages des diagrammes en mosaïque dans un environnement interactif :

- Pour consulter en détail de petites parties du graphique d'origine, vous pouvez utiliser la fonctionnalité de zoom avant.
- Lorsque vous placez le pointeur de la souris sur une mosaïque, des info-bulles de données affichent les mesures de cette mosaïque.
- Vous pouvez créer un lien vers un rapport ou une page web à partir des mosaïques d'un niveau particulier. Lorsque vous cliquez sur une mosaïque de ce niveau, le rapport ou la page web associé à cette mosaïque apparaît.

Seuls les formats de graphique ActiveX, Java, Image ActiveX et Image Java sont valides avec la tâche Diagramme en mosaïque. Si le format de graphique actuel n'est pas l'un de ces formats, la tâche Diagramme en mosaïque utilise le format ActiveX.

Pour les résultats HTML, certaines fonctionnalités ne sont disponibles que pour des formats de graphique spécifiques.

- La fonctionnalité de zoom avant est disponible si vous avez sélectionné le format de sortie graphique ActiveX ou Java.
- La fonctionnalité de création d'un lien avec une page web à partir d'un diagramme en mosaïque est disponible si vous avez sélectionné le format de sortie graphique Java.

Vous pouvez spécifier le format graphique et le format des résultats dans la boîte de dialogue **Options**.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	GTILE
Produits SAS requis	SAS Base, SAS/GRAPH
Autres produits SAS recommandés	Aucun

Diagramme en mosaïque : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Fonctions de la tâche** pour accéder à ces options.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Remarque : les fonctions disponibles dépendent du type de graphique sélectionné.

Nom du rôle	Description
Variables de classification	Désigne les variables de catégorie hiérarchique (ou variables ID). Une variable de classification est obligatoire et vous pouvez attribuer jusqu'à six variables à cette fonction. Si vous attribuez plusieurs variables à cette fonction, l'ordre des variables définit la hiérarchie. Par exemple, si l'ordre des

Nom du rôle	Description
	variables de la fonction Variables de classification est Pays, Région et Division, la hiérarchie obtenue est Pays > Région > Division.
Variable d'analyse de couleur	Désigne les valeurs de variable qui déterminent la couleur des mosaïques. La valeur de liste de couleurs minimale est attribuée à la plus petite réponse de couleur, tandis que la valeur de liste de couleurs maximale est attribuée à la réponse de couleur la plus élevée. Les autres réponses de couleur se voient attribuer une valeur de façon linéaire dans un gradient entre les couleurs de la liste. Vous ne pouvez attribuer qu'une seule variable à cette fonction. Remarque : si vous n'attribuez aucune variable d'analyse de couleur, la variable que vous attribuez à la fonction Variable d'analyse de taille est également utilisée comme variable d'analyse de couleur.
Variable d'analyse de taille	Désigne les valeurs de variable qui déterminent la taille relative des mosaïques. La taille globale de toutes les mosaïques recouvre entièrement la zone de mosaïques. Si une valeur manquante, nulle ou négative est attribuée à une mosaïque, celle-ci est ignorée. Une variable d'analyse de taille est obligatoire et vous ne pouvez attribuer qu'une seule variable à cette fonction.
Grouper les graphiques par	Les valeurs de la colonne que vous attribuez à cette fonction déterminent le nombre de graphes créés. Un graphique distinct est créé pour chaque valeur unique. Vous pouvez attribuer plusieurs variables à cette fonction.

Diagramme en mosaïque : Définir les options d'apparence

Définir l'apparence d'un diagramme en mosaïque

Dans la liste, sous l'en-tête Apparence, cliquez sur **Diagramme en mosaïque** pour accéder à ces options.

Nom du rôle	Description
Disposition des mosaïques	Choisissez la disposition des mosaïques du diagramme. Vous avez le choix entre ces options : ■ Flux répartit la disposition des mosaïques sur plusieurs lignes alignées. Dans le diagramme, les mosaïques sont lisibles de gauche à droite, puis de haut en bas. ■ Carré est l'arborescence rectangulaire classique. En général, les mosaïques sont organisées de la plus petite à la plus grande, de l'angle inférieur gauche du diagramme à l'angle supérieur droit. Le tri des variables de données est ignoré et la forme des mosaïques est aussi carrée que possible. ■ Bascule fractionne la direction pour chaque niveau de la hiérarchie issue des variables attribuées à la fonction Variables de classification. Le niveau hiérarchique le plus élevé est fractionné en secteurs verticaux de l'espace disponible et affiché dans l'ordre de tri le long d'un axe unique. Aux niveaux inférieurs, la direction du fractionnement est modifiée (ou inversée), de sorte que le deuxième niveau est fractionné horizontalement.
Indiquer le niveau de classification	spécifier le nombre de niveaux à afficher dans les résultats. Par exemple, si l'ordre des variables de la fonction Variables de classification est Pays, Région et Division, la hiérarchie obtenue est Pays > Région > Division. Le nombre de niveaux affichés dans le diagramme en mosaïque dépend de la valeur de l'option Indiquer le niveau de classification. ■ Si la valeur de cette option est 1, un seul niveau de la hiérarchie est affiché dans les résultats. Le diagramme en mosaïque affiche les mosaïques de chaque pays (Canada, Allemagne et États-Unis, par exemple). Si vous effectuez un zoom avant sur la mosaïque Canada, le diagramme en mosaïque affiche les mosaïques des deux régions (Ouest et Est) du Canada. ■ Si la valeur de cette option est 2, deux niveaux de la hiérarchie sont visibles dans les résultats. Le diagramme en mosaïque affiche les mosaïques de chaque pays et région (Canada Ouest, Canada Est, Allemagne Ouest, Allemagne Est, États-Unis Ouest et États-Unis Est). Si vous effectuez un zoom avant sur la mosaïque Canada, les mosaïques de toutes les divisions des régions Canada Ouest et Canada Est sont visibles. Par défaut, tous les niveaux de la hiérarchie sont affichés.
Indiquer les couleurs	personnalise la palette de couleurs du diagramme en mosaïque. La palette de couleurs est la répartition des couleurs sur un intervalle de valeurs de données. Par défaut, vous devez indiquer au moins deux couleurs. Ces couleurs sont les couleurs de début et de fin de la palette.

Nom du rôle	Description
	Par défaut, les couleurs de la palette de couleurs sont réparties équitablement sur la légende du dégradé. Vous pouvez utiliser les points de couleurs pour spécifier comment les couleurs doivent être réparties dans les données. Pour spécifier des points de couleur, cliquez sur Modifier. Dans la boîte de dialogue Modifier les points de couleur, les valeurs peuvent s'étendre de 0 à 1 et doivent être dans l'ordre croissant. Par exemple, si vous avez indiqué trois couleurs pour la palette de couleurs, vous pouvez spécifier 0, .75, et 1 comme points de couleur. Il est possible d'indiquer la même valeur pour différents points de couleur. Par exemple, si vous avez indiqué cinq couleurs pour la palette de couleurs, vous pouvez spécifier 0, .2, .4, .4, et .6 comme points de couleur. Vous devez indiquer le même nombre de points de couleur que le nombre de couleurs dans la palette de couleurs.

Définir les options de légende

Pour indiquer les points de début et de fin de la légende, activez la case **Spécifier une plage de légende**. Supposons que les valeurs de la plage de données aillent de 0 à 100 000. Néanmoins, sur le diagramme en mosaïque, vous souhaitez que le code couleur mette en valeur les valeurs de 22 000 à 26 000. Dans cet exemple, vous pouvez spécifier 22 000 comme valeur minimale (ou de début) et 26 000 comme valeur maximale (ou de fin). Les mosaïques avec des valeurs inférieures à 22 000 ou supérieures à 26 000 apparaissent en gris. Les mosaïques avec des valeurs de 22 000 à 26 000 utilisent la palette de couleurs que vous avez indiquée pour le diagramme en mosaïque.

Par défaut, le code couleur de la légende reprend les couleurs des données de réponse. Toutefois, vous pouvez indiquer si le code couleur doit être symétrique de chaque côté d'une valeur fixe à l'aide de la case à cocher **Définir le code couleur de la légende comme étant symétrique autour d'une valeur fixe**. Cette option permet de spécifier le milieu de l'échelle de couleurs. Par défaut, le point de symétrie ou le milieu est 0. Vous pouvez spécifier un nouveau point de symétrie ou milieu.

Diagramme en mosaïque : Définir les options avancées

Dans la liste, cliquez sur **Options avancées** pour accéder à ces options.

Remarque : Vous pouvez créer un lien vers des pages web à partir de votre diagramme en mosaïque à condition d'avoir sélectionné le format de sortie graphique Java et le format de résultats HTML dans SAS Enterprise Guide. Pour spécifier le format du graphique et des résultats, consultez les panneaux **Résultats** ⇒ **Graphique** de la boîte de dialogue **Options**. En raison des spécifications Java, les liens ne sont pas pris en charge dans SAS Add-In for Microsoft Office.

Vous pouvez ajouter des liens vers des pages web à partir de votre diagramme en mosaïque. Un lien est disponible dans le menu contextuel qui s'affiche lorsque vous faites un clic droit sur la mosaïque d'un niveau correspondant à la variable de lien sélectionnée. (Les variables de lien sont les variables auxquelles vous avez attribué la fonction **Variables de classification**). Le niveau où ces liens apparaissent dépend de si vous avez spécifié un niveau de classification.

- Si vous n'avez pas spécifié de niveau de classification, les liens sont disponibles à tous les niveaux du diagramme en mosaïque. Par exemple, si la hiérarchie du diagramme en mosaïque est Pays > Région > Division, les liens sont disponibles lorsque vous faites un clic droit sur les mosaïques des niveaux Pays, Région et Division.
- Si vous avez indiqué un niveau de classification, les liens ne sont disponibles qu'aux niveaux les plus bas du diagramme en mosaïque. Supposons, par exemple, que la hiérarchie du diagramme en mosaïque comporte trois niveaux : Pays, région et division. Si le niveau de classification est 2, les liens sont disponibles dans les deux niveaux inférieurs, à savoir Région et Division. Si le niveau de classification est 1, les liens ne sont disponibles que à partir du niveau le plus bas, Division.

Dans la tâche Diagramme en mosaïque, les variables de lien sont répertoriées dans le champ **Variables disponibles**. La tâche attribue automatiquement un numéro de substitution à chaque variable de lien. Ce numéro apparaît en regard de la variable associée, dans la zone **Variables disponibles**. Vous pouvez insérer le numéro de substitution dans l'URL du lien web ou dans le libellé afin de créer un lien unique pour chaque valeur d'une variable de lien.

Supposons, par exemple, que dans vos données, les valeurs de Pays soient fr (France), jp (Japon) et uk (Royaume-Uni). Si 1 est le numéro de substitution automatiquement attribué pour la variable Pays et si vous indiquez `http://www.sales.&1;.com` comme URL du lien Web, les liens suivants seront disponibles à partir du diagramme en mosaïque :

- `http://www.sales.fr.com`
- `http://www.sales.jp.com`
- `http://www.sales.uk.com`

Pour inclure des liens dans votre graphique :

- 1 Indiquez le libellé du lien. Si vous souhaitez créer un libellé unique pour chaque valeur de la variable de lien, vous pouvez insérer dans le libellé le numéro de substitution de la variable de lien. Pour ajouter le numéro de substitution d'une variable de lien dans le libellé, cliquez sur **Insérer une variable**.

Voici quelques exemples de libellé de lien :

- Page web d'état pour {& 1}

■ Page Web Wikipedia pour {&2}

- 2 Indiquez l'URL de la page Web. Vous pouvez créer un lien web unique pour chaque valeur de la variable de lien en insérant dans l'URL le numéro de substitution de la variable de lien. La syntaxe de la variable de liaison est : {&numéro_substitution}. Cette valeur de substitution peut figurer n'importe où dans l'URL du lien web. Pour ajouter le numéro de substitution d'une variable de lien dans l'URL, cliquez sur **Insérer une variable**.

Voici quelques exemples d'URL de lien Web :

- `http://www.state.{&1}.us`
- `http://en.wikipedia.org/wiki/{&2}`

- 3 Cliquez sur **Ajouter** pour ajouter ce lien Web à la liste des liens Web du diagramme en mosaïque en cours.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.

Transposer

<i>À propos de la tâche Transposer</i>	703
<i>Transposer : Attribuer des variables aux fonctions d'analyse</i>	704
<i>Transposer : Définir les options de l'analyse</i>	705
<i>Transposer : Définition des options des résultats</i>	706
<i>Afficher les propriétés</i>	706

À propos de la tâche Transposer

La tâche Transposer transforme les colonnes sélectionnées d'une table d'entrée en lignes d'une table de sortie. Si vous n'utilisez pas de variables de groupe, chaque colonne sélectionnée devient une seule ligne. Si vous utilisez des variables de groupe, les colonnes sélectionnées sont divisées en sous-colonnes basées sur les valeurs des variables de groupe, et chaque sous-colonne devient une ligne de la table de sortie.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	TRANSPOSE
Produits SAS requis	Base SAS
Autres produits SAS recommandés	SAS/STAT

Transposer : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Transposer les variables	Chaque colonne que vous attribuez à cette fonction devient une ou plusieurs lignes de la table de sortie. Si vous ne sélectionnez pas de variables de groupe, une colonne entière est transformée en une seule ligne. Si vous sélectionnez une ou plusieurs variables de groupe, ces variables sont utilisées pour fractionner chaque colonne en sous-colonnes, dont chacune est transformée en une ligne. Dans ce cas, une colonne est transposée dans un nombre de lignes égal au nombre de groupes définis par les variables de groupe. Vous devez attribuer au moins une colonne à la fonction Transposer les variables. Pour sélectionner une variable de groupe, attribuez une colonne à la fonction Grouper l'analyse par.
Copier des variables	Chaque colonne que vous attribuez à cette fonction est copiée directement de la table d'entrée vers la table de sortie sans être transposée. Étant donné que ces colonnes sont copiées directement dans la table de sortie, le nombre de lignes de la table de sortie est égal au nombre de lignes de la table d'entrée. Si le nombre de lignes de la table d'entrée est différent du nombre de variables qu'elle transpose, la table de sortie est complétée par des valeurs manquantes.
Nouveaux noms de colonnes	Si vous attribuez une colonne d'entrée à cette fonction, les valeurs de la colonne sont utilisées pour former les suffixes des en-têtes de colonnes de la table de sortie. Si vous n'attribuez pas de colonne d'entrée à cette fonction, des nombres entiers sont utilisés pour former les suffixes des en-têtes de colonne de la table de sortie. Pour spécifier le préfixe des en-têtes de

Nom du rôle	Description
	colonne, cliquez sur Options dans la liste, puis activez la case à cocher Utiliser le préfixe. Saisissez ensuite le préfixe dans la zone Préfixe du nom de la colonne. En général, une colonne, utilisée pour fournir les noms de colonnes de la table de sortie, ne doit pas contenir des valeurs en double car cela aurait pour effet de créer des doublons de noms de colonnes. Si vous voulez utiliser une colonne contenant des valeurs en double, sélectionnez la colonne après l'avoir attribuée, puis sélectionnez la case à cocher Autoriser doublons s'affichant sur la droite. Dans ce cas, seule la dernière ligne contenant une valeur donnée sera utilisée dans la transposition. Si vous utilisez une colonne contenant des doublons et ne sélectionnez pas la case à cocher Autoriser doublons, la transposition ne peut être effectuée.
Grouper l'analyse par	Chaque variable que vous attribuez à cette fonction est utilisée pour fractionner les colonnes sur le point d'être transposées en sous-colonnes qui seront transposées séparément. Chaque sous-colonne, définie par un ensemble de valeurs des variables de groupe, devient une ligne de la table de sortie.
Nouveaux libellés de colonnes	Cette fonction devient disponible si vous attribuez une colonne d'entrée à la fonction Nouveaux noms de colonnes. Si vous attribuez une colonne d'entrée à cette fonction, les valeurs de cette colonne sont utilisées en tant que libellés des variables de données dans la table de sortie. Si vous n'attribuez pas de colonne d'entrée à cette fonction, les libellés des variables de données de la table de sortie seront identiques aux noms de ces variables.

Transposer : Définir les options de l'analyse

Dans la liste, cliquez sur **Options** pour accéder à ces options.

Chaque ligne de la table de sortie contient le nom et le libellé de la variable de la table d'entrée. Vous pouvez spécifier un en-tête pour la colonne de sortie qui contient ces noms ou libellés de variables. Le nom ou le libellé peut comporter des caractères spéciaux, des chiffres à gauche et un espace, mais ne peut pas comprendre plus de 32 caractères. L'en-tête par défaut de la colonne du nom de la

variable est **Source**. L'en-tête par défaut de la colonne de libellé variable est **Libellé**.

Les en-têtes de toutes les colonnes de données de la table de sortie commencent par le même préfixe. Le préfixe peut comporter des caractères spéciaux, des chiffres à gauche et un espace, mais il ne peut pas comprendre plus de 32 caractères. Le préfixe par défaut est **Colonne**.

Remarque : l'indication d'un préfixe est facultative si vous attribuez une variable à la fonction **Nouveaux noms de colonnes**. Si vous n'attribuez pas de variable à cette fonction, un préfixe est requis et la case à cocher **Utiliser le préfixe** n'est pas disponible.

Transposer : Définition des options des résultats

Dans la liste, cliquez sur **Résultats** pour accéder à ces options.

Vous pouvez attribuer un autre nom à la table de sortie.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,

- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.

Carte u

<i>À propos de la Carte u</i>	709
<i>Carte u : Attribuer des variables aux fonctions d'analyse</i>	710
<i>Cartes de contrôle : Indiquer des limites de contrôle</i>	711
<i>Carte de contrôle : Sélectionner les tests à effectuer</i>	713
<i>Cartes de contrôle : Définir les options d'apparence du graphe</i>	714
<i>Cartes de contrôle : Enregistrer les résultats d'analyse</i>	714
<i>Indiquer des titres et des notes de bas de page</i>	715
<i>Afficher les propriétés</i>	716

À propos de la Carte u

La carte de contrôle de Shewhart est un outil graphique et analytique qui permet de déterminer si un processus est sous contrôle statistique. La tâche Carte u crée des cartes u pour le nombre de non-conformités (défauts) par unité d'inspection dans les échantillons du sous-groupe contenant un nombre arbitraire d'unités.

Vous pouvez utiliser cette tâche pour déterminer le nombre de défauts par unité d'inspection résultant d'un processus de fabrication. Par exemple, supposons qu'un fabricant de textiles utilise une carte u pour suivre le nombre de défauts par mètre carré de tissu. Lors de la recherche des défauts, le tissu est bobiné en rouleaux. Chaque pièce de tissu mesure 1 m de largeur sur 30 m de longueur. Le décompte des défauts sur 20 rouleaux est enregistré dans une table SAS utilisée pour créer la carte u. Chaque point de la carte u représente le nombre de non-conformités par unité d'inspection pour un sous-groupe particulier. Par exemple, la valeur représentée pour le premier sous-groupe est $12/30=0,4$ (sachant que le premier rouleau présente 12 défauts et compte 30 mètres carrés du tissu). Si aucun des points ne dépasse la limite de contrôle (égale à 3 sigma par défaut), la carte u indique que le processus de fabrication du tissu est sous contrôle statistique.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	SHEWHART
Produits SAS requis	SAS Base, SAS/QC, SAS/GRAPH
Autres produits SAS recommandés	SAS/STAT

Carte u : Attribuer des variables aux fonctions d'analyse

Dans la liste, cliquez sur **Données** pour accéder à ces options.

Par défaut, la source de données que vous avez sélectionnée avant d'ouvrir la tâche est la source de données en entrée de la tâche. Le nom de la source de données en entrée apparaît dans le champ **Source de données**. Tout filtre spécifié dans la source de données apparaît dans le champ **Filtre**. Pour modifier ou filtrer la source de données d'entrée, cliquez sur **Modifier**.

Après avoir sélectionné une source de données, vous pouvez attribuer des variables aux fonctions.

Nom du rôle	Description
Nombre de non-conformités	indique les variables contenant le nombre de non-conformités. Vous devez attribuer au moins une variable à cette fonction
Identificateur du sous-groupe	indique la variable qui identifie les sous-groupes dans les données. Les valeurs de la variable indiquent de quelle manière les observations de la table d'entrée sont organisées en sous-groupes rationnels. En général, les valeurs sont les suivantes : ■ les indices qui donnent l'ordre dans lequel les sous-échantillons sont collectés ■ les dates ou heures auxquelles les sous-échantillons ont été collectés ■ les libellés qui identifient de façon unique les échantillons de sous-groupe Un identificateur de sous-groupe est requis et vous ne pouvez attribuer qu'une seule variable à cette fonction. Pour trier les données selon les sous-groupes identifiés, activez la case à cocher Trier par sous-groupe.

Nom du rôle	Description
Taille de l'échantillon du sous-groupe	spécifie la taille de l'échantillon du sous-groupe comme valeur de la variable. La taille de l'échantillon du sous-groupe est nécessaire et vous n'attribuez qu'une seule variable numérique à cette fonction. Activez la case à cocher Utiliser une valeur numérique pour le sous-groupe (au lieu d'une var.) lorsque la taille de l'échantillon de sous-groupe est déterminée et que vous devez définir une valeur constante pour la taille. Saisissez la valeur à utiliser dans la zone Taille de l'échantillon du sous-groupe. La taille de l'échantillon du sous-groupe doit être une valeur comprise entre la valeur maximale de la variable de traitement et 99999. Remarque : cette fonction n'est pas nécessaire si vous avez activé la case à cocher Utiliser une valeur numérique pour le sous-groupe (au lieu d'une var.) dans le volet Options.
Grouper l'analyse par	produit des analyses distinctes sur les observations des groupes définis par les variables de groupe. Par défaut, les données sont triées par ordre croissant des valeurs de ces variables de groupe. Remarque : pour empêcher le tri de la table d'entrée, désactivez la case à cocher Trier par variables. Remarque : vous ne pouvez pas grouper les analyses en fonction d'une variable que vous avez déjà sélectionnée pour une autre fonction.
Variables de bloc	indique deux variables maximum, qui regroupent les données par blocs de sous-groupes consécutifs. Ces blocs sont libellés dans une légende, et chaque variable de bloc fournit un niveau pour les libellés de la légende.

Cartes de contrôle : Indiquer des limites de contrôle

Dans la liste, cliquez sur **Limites de contrôle** pour accéder à ces options.

Nom d'option	Description
Limites Sigma	spécifie la largeur des limites de contrôle commun multiple de l'erreur type (sigma) des statistiques récapitulatives du sous-groupe tracé dans le graphique. La largeur doit être positive. Le multiple par défaut est 3.
Sélectionner une méthode de calcul	spécifie la méthode de calcul pour les limites de contrôle. Vous avez le choix entre ces options : ■ Calculer les limites de contrôle à partir des données actives ■ Calculer les limites de contrôle à partir de la table sélectionnée Pour désigner la table contenant les limites de contrôle ou les paramètres à partir desquels les limites de contrôle peuvent être calculées, cliquez sur Parcourir. Sélectionnez la table à utiliser et cliquez sur Ouvrir. Remarque : la table que vous sélectionnez doit être une table LIMITS=. Pour plus d'informations sur les variables requises dans une table LIMITS=, consultez la documentation des procédures du graphique que vous êtes en train de créer. Pour représenter les statistiques descriptives de tous les sous-groupes, activez la case à cocher Trace les statistiques descriptives de tous les sous-groupes, que la taille de l'échantillon du sous-groupe soit égale à la taille nominale des limites de contrôle ou non. Vous pouvez représenter ces statistiques lorsque la taille de l'échantillon est la même pour pratiquement tous les sous-groupes et lorsque vous voulez afficher les limites de contrôle fixes (non soumises à variation) correspondant à la taille d'échantillon nominale n. Si vous sélectionnez cette option alors que les tailles des échantillons de vos sous-groupes sont très différentes, l'interprétation des limites de contrôle n'a de sens que pour les sous-groupes dont la taille de l'échantillon est égale à n. Si vous souhaitez ajouter des symboles spéciaux pour les points de données où la taille de l'échantillon est différente de n, activez la case à cocher Ajouter des symboles spéciaux pour les points correspondant aux tailles d'échantillon différentes de la taille nominale pour les limites de contrôle fixes. ■ Spécifier les limites de contrôle Dans les champs Supérieure, Centrale, et Inférieure, indiquez respectivement la limite de contrôle supérieure, la ligne centrale et la limite de contrôle

Nom d'option	Description
	inférieure. Si l'effectif du sous-groupe varie, vous devez indiquer un effectif nominal. Remarque : si vous attribuez des variables à la fonction Grouper l'analyse par , vous ne pouvez pas saisir des valeurs pour la limite de contrôle.

Carte de contrôle : Sélectionner les tests à effectuer

Dans la liste, cliquez sur **Tests** pour accéder à ces options.

Remarque : ces tests ne s'appliquent pas aux limites de contrôle qui ne sont pas trois limites sigma ou qui varient en fonction de la taille de l'échantillon du sous-groupe.

Nom d'option	Description
Sélectionner un test	Utilisez les cases à cocher pour demander un ou plusieurs tests pour des sujets spéciaux, également connus sous les noms de test de dérive : Run test, Pattern et Western Electric tests. Ces tests détectent les dérives particulières non aléatoires dans les points représentés sur le graphique de contrôle principal. L'occurrence du modèle non aléatoire est appelé signal. Les tests peuvent indiquer la présence d'une cause spéciale de variation. Lorsque vous sélectionnez l'un des tests, une description de ce test apparaît dans la zone Description .
Libellé	Utilisez les champs de saisie afin de saisir des libellés pour les points où les tests pour causes spéciales sont positifs. Les libellés ne peuvent contenir plus de 16 caractères. Chaque libellé apparaît à tous les points du test où un signal est rencontré.
Texte identifiant les signaux de test	■ Utilisez la liste déroulante afin de sélectionner une couleur pour les libellés que vous avez indiqués. ■ Utilisez la case à cocher Afficher lignes de zone pour inclure des lignes qui représentent les zones A, B et C

Nom d'option	Description
	sur un graphique principal. Ces zones se construisent en divisant l'intervalle entre les limites de contrôle en six sous-intervalles espacés de façon égale.
	■ Utilisez la case Surpasser la limite 3 sigma pour tester des sujets spéciaux si vous spécifiez une limite de contrôle sigma différente du multiple par défaut de 3. Pour changer la limite sigma, cliquez sur Limites de contrôle dans le volet de sélection.
	■ Cochez la case Appliquer le test aux dérives des points se chevauchant afin d'appliquer des tests pour les causes spéciales aux dérives de points se chevauchant.

Cartes de contrôle : Définir les options d'apparence du graphe

Certaines options d'apparence sont communes à plusieurs tâches de la carte de contrôle. Pour plus d'informations sur ces options, voir ces rubriques :

- ["Définir les options d'un axe" page 743](#)
- ["Ajouter des lignes de référence" page 744](#)
- ["Définir d'autres options de tracé" page 744](#)
- ["Définir les options de variable de bloc" page 746](#)

Cartes de contrôle : Enregistrer les résultats d'analyse

Dans la liste, cliquez sur **Tables** pour accéder à ces options.

Pour enregistrer les résultats de l'analyse, vous pouvez créer des tables de sortie contenant plusieurs statistiques. Activez les cases à cocher pour les données de sortie que vous voulez enregistrer de façon permanente.

Nom du rôle	Description
Statistiques de sous-groupe et table des limites de contrôle	crée une table de sortie contenant les informations représentées sur le graphe. Cela inclut les valeurs de la variable de sous-groupe ainsi que les statistiques descriptives correspondantes et les limites de contrôle. Remarque : cette option crée une table OUTTABLE=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option TABLE= décrite par la procédure SHEWHART. Pour plus d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des statistiques de sous-groupe	crée une table de sortie contenant les statistiques descriptives du sous-groupe. Remarque : cette option crée une table OUTHISTORY=. Vous pouvez réenregistrer cette table dans SAS à l'aide de l'option HISTORY= décrite par la procédure SHEWHART. Pour plus d'informations, veuillez consulter l'aide de la procédure SHEWHART.
Table de sortie des limites de contrôle	crée une table de sortie qui enregistre les limites de contrôle. Vous pouvez utiliser cette table lorsque vous sélectionnez la méthode de calcul pour les limites de contrôle lors d'une utilisation ultérieure de cette tâche.

Remarque : SAS Enterprise Guide parcourt la liste de bibliothèques définie dans la zone **Tâches > Bibliothèque de sortie** de la boîte de dialogue **Options** et enregistre les données de sortie dans la première bibliothèque accessible en écriture qu'il y trouve.

SAS Add-In for Microsoft Office stocke les données de sortie dans la bibliothèque WORK. Pour désigner un emplacement de stockage différent pour vos données de sortie, cliquez sur **Parcourir**.

Indiquer des titres et des notes de bas de page

Dans la liste, cliquez sur **Titres** pour accéder à ces options.

S'il vous convient, utilisez tel quel le texte par défaut qui décrit les résultats dans le titre et dans la note de bas de page.

Pour indiquer vos propres titres et notes de bas de page :

- 1 Sélectionnez le titre ou la note de bas de page dans la zone **Section**.
- 2 Désactivez la case à cocher **Utiliser le texte par défaut**.
- 3 Modifiez le texte du titre ou de la note de bas de page dans le champ de saisie.

.....
Remarque : vous pouvez utiliser des macro-variables dans les titres et les notes de bas de page.
.....

Afficher les propriétés

Dans la liste, cliquez sur **Propriétés** pour accéder à ces options.

Vous pouvez afficher la synthèse des propriétés d'une tâche à partir du volet **Propriétés**. Pour certaines tâches, vous pouvez modifier ces propriétés en cliquant sur **Modifier**. La boîte de dialogue **Propriétés** apparaît.

Dans le volet **Propriétés**, vous pouvez consulter ces informations :

- le libellé de la tâche,
- le serveur utilisé pour exécuter la tâche
- les date et heure de création de la tâche,
- les date et heure de dernière modification de la tâche,
- l'heure de la dernière exécution de la tâche,
- les limites définies pour les données (le cas échéant),
- les invites utilisées,
- le format des résultats,
- le format de sortie graphique des résultats.

.....
Remarque : si vous modifiez le libellé d'une tâche, le nouveau libellé est enregistré lorsque vous cliquez sur **OK** dans la boîte de dialogue **Propriétés**. Le nouveau libellé est utilisé même si vous annulez la tâche.
.....

Actualiser les métadonnées de la bibliothèque

<i>À propos de l'Assistant Actualiser les métadonnées de la bibliothèque</i>	717
<i>Actualiser les métadonnées de la bibliothèque : Sélectionnez le SAS Server et la bibliothèque</i>	718
<i>Actualiser les métadonnées de la bibliothèque : Exécuter un rapport sur les métadonnées ou actualisez ces dernières.</i>	718

À propos de l'Assistant Actualiser les métadonnées de la bibliothèque

L'Assistant **Actualiser les métadonnées de la bibliothèque** permet d'exécuter un rapport sur les définitions de table figurant dans le référentiel de métadonnées ou d'actualiser ces définitions. Vous pouvez créer un rapport qui répertorie le contenu de la bibliothèque à partir des définitions de table figurant dans le référentiel de métadonnées, ou un rapport répertoriant les différences entre les tables de la bibliothèque physique et les définitions de table du référentiel de métadonnées. Outre l'exécution de rapports, l'Assistant **Actualiser les métadonnées de la bibliothèque** permet d'actualiser, d'ajouter ou de supprimer des définitions de table dans le référentiel de métadonnées.

Par défaut, SAS Enterprise Guide affiche uniquement les tables enregistrées dans les métadonnées. L'Assistant **Actualiser les métadonnées de la bibliothèque** utilise la procédure METALIB pour enregistrer ces tables dans le référentiel de métadonnées par programmation et synchroniser les métadonnées de ces tables.

Remarque : cette tâche n'est disponible que dans SAS Enterprise Guide.

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	METALIB
Produits SAS requis	Base SAS
Autres produits SAS recommandés	Néant

Actualiser les métadonnées de la bibliothèque : Sélectionnez le SAS Server et la bibliothèque

- 1 Sélectionnez un serveur dans la liste des serveurs disponibles.
- 2 Sélectionner une bibliothèque.
- 3 Cliquez sur **Suivant** pour exécuter un rapport ou actualiser les métadonnées de la bibliothèque sélectionnée.

Actualiser les métadonnées de la bibliothèque : Exécuter un rapport sur les métadonnées ou actualisez ces dernières.

- 1 Avant d'exécuter un rapport ou d'actualiser les métadonnées, vous devez sélectionner le SAS Server et la bibliothèque.
- 2 Sélectionnez l'une de ces options :

Nom d'option	Description
Rapport sur son contenu, tel qu'il est enregistré dans le référentiel de métadonnées	génère un rapport du contenu de la bibliothèque à partir des définitions de table figurant dans le référentiel de métadonnées.
Rapport sur les différences entre les tables physiques et le référentiel de métadonnées	génère un rapport contenant ces informations : ■ définitions de table du référentiel de métadonnées qui seraient actualisées pour coïncider avec les tables physiques correspondantes ; ■ tables de la bibliothèque physique pour lesquelles il n'existe pas de définition dans le référentiel de métadonnées ; ■ tables qui ne figurent plus dans la bibliothèque physique mais pour lesquelles il existe une définition dans le référentiel de métadonnées.
Actualiser et ajouter les définitions de tables dans les métadon. avec les tables et col. réelles	réalise ces étapes : ■ Pour les tables de la bibliothèque physique qui contiennent des métadonnées, les définitions de table du référentiel de métadonnées sont actualisées afin d'être identiques aux tables physiques correspondantes. Ces mises à jour incluent des informations relatives aux colonnes, index, clés uniques, clés étrangères et associations de clés des tables. ■ Pour les tables de la bibliothèque physique qui ne contiennent pas de métadonnées, les définitions de ces tables sont ajoutées au référentiel de métadonnées.
N'actualiser que définitions de table existant dans les métadon. avec les info. de col. en cours	actualise dans le référentiel de métadonnées les définitions des tables de la bibliothèque physique qui contiennent des métadonnées.
Effacer les entrées obsolètes de la bibliothèque de métadonnées (tables qui n'existent plus)	supprime du référentiel de métadonnées les définitions des tables qui n'existent plus dans la bibliothèque physique.

- 3 Avant d'actualiser ou de supprimer des métadonnées, il est possible que vous deviez saisir l'identifiant et le mot de passe d'un compte doté de droits d'administration. Si aucun identifiant ni mot de passe n'est requis, ces options ne sont pas disponibles.
- 4 Cliquez sur **Terminer** pour exécuter le rapport ou actualiser les métadonnées de la bibliothèque.

Téléchargement vers CAS

À propos de la tâche Téléchargement vers CAS	721
Télécharger des données vers un SAS CAS Analytic Server	722

À propos de la tâche Téléchargement vers CAS

La tâche Téléchargement vers CAS permet de charger des données SAS 9.4 vers un Cloud Analytics Services (CAS) Server. Cette tâche utilise SAS/CONNECT pour déplacer les données vers un serveur distant. À partir de la version SAS 9.4M5, il existe des façons plus efficaces de déplacer les données vers un CAS Server. Par exemple, vous pouvez utiliser le moteur CAS LIBNAME ou PROC CASUTIL. Pour plus d'informations, voir [SAS Cloud Analytic Services: User's Guide](#).

Remarque : Pour exécuter cette tâche dans SAS Enterprise Guide ou SAS Add-In for Microsoft Office, vous devez d'abord configurer votre environnement. Pour plus d'informations, voir [SAS Intelligence Platform: Desktop Application Administration Guide](#).

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	Néant
Produits SAS requis	SAS Base, SAS Viya
Autres produits SAS recommandés	Néant

Télécharger des données vers un SAS CAS Analytic Server

- 1 Sélectionnez le CAS Server.
- 2 Sélectionner la bibliothèque où enregistrer les données.
- 3 Indique le nom du membre de la table. Par défaut, le nom du membre est le même que le nom de la table. Vous ne pouvez pas changer le nom du membre si vous utilisez un chargement distribué.
- 4 Indique où localiser les informations d'identification pour la table sélectionnée. Si vous êtes connecté à un SAS Metadata Server, vous pouvez utiliser les informations d'identification stockées dans les métadonnées. Si vous téléchargez des données locales vers le CAS Server, vous devez saisir l'identifiant et le mot de passe pour le CAS Server.

Télécharger des fichiers de données vers un serveur

<i>À propos de la tâche Télécharger des fichiers de données vers un serveur</i>	723
<i>Télécharger des fichiers de données vers un serveur : Sélectionner les fichiers de données SAS à télécharger</i>	724
<i>Télécharger des fichiers de données vers un serveur : Spécifier les options de téléchargement</i>	724

À propos de la tâche Télécharger des fichiers de données vers un serveur

La tâche Télécharger des fichiers de données vers un serveur vous permet de sélectionner un ou plusieurs fichiers de données SAS (*.sas7bdat) dans votre système de fichiers local Microsoft Windows et de les copier vers un SAS Server.

Cette tâche génère un journal qui récapitule les fichiers copiés, leur taille en octets, la temporisation de chaque fichier et de tous les fichiers regroupés, ainsi que les éventuelles erreurs apparues au cours du transfert. Par défaut, tous les fichiers de données sont ajoutés en tant que données de sortie à votre projet SAS Enterprise Guide afin de pouvoir être utilisés dans ce projet.

Télécharger des fichiers de données vers un serveur : Sélectionner les fichiers de données SAS à télécharger

- 1 Cliquez sur **Ajouter** pour sélectionner les fichiers de données SAS que vous voulez télécharger.
- 2 Sélectionnez le nom du fichier à copier.

.....
Remarque : les fichiers doivent être au format SAS7BDAT.
.....

- 3 Cliquez sur **Ouvrir** pour sélectionner les fichiers de données que vous voulez copier.

.....
Remarque : un nom en caractères minuscules est attribué à chaque fichier copié sur le serveur cible.
.....

- 4 Cliquez sur **Suivant** pour définir les options de téléchargement.

Télécharger des fichiers de données vers un serveur : Spécifier les options de téléchargement

- 1 Avant de définir les options de téléchargement, vous devez sélectionner les fichiers de données que vous souhaitez télécharger.
- 2 Dans la liste déroulante **Sélectionner un SAS server**, choisissez le serveur sur lequel vous voulez copier les tables sélectionnées.
- 3 Dans la liste déroulante **Sélectionnez une bibliothèque de destination**, choisissez la bibliothèque SAS que vous souhaitez utiliser.

- 4 Si vous souhaitez que les fichiers à copier remplacent des fichiers du même nom sur le SAS Server, activez la case à cocher **Remplacer les fichiers existants portant le même nom**. Par défaut, les fichiers existants ne sont pas remplacés
- 5 Si vous ne souhaitez pas ajouter les fichiers de données copiées dans le projet SAS Enterprise Guide en cours, désactivez la case à cocher **Ajouter les fichiers de données dans votre projet en cours**. Par défaut, les fichiers de données copiés sont ajoutés au projet en cours.
- 6 Cliquez sur **Terminer** pour copier les fichiers de données.

Remarque : si la bibliothèque cible est une bibliothèque de moteur SAS Base ne comportant qu'un seul chemin de fichier, les fichiers seront directement transférés vers le chemin de cette bibliothèque. S'il ne s'agit pas d'une bibliothèque de moteur SAS Base ou s'il s'agit d'une bibliothèque concaténée, les fichiers sont dans un premier temps transférés vers une zone de transit de la bibliothèque Work, puis copiés vers le dossier cible au moyen d'une instruction PROC SQL.

Téléchargement vers LASR

À propos de la tâche Téléchargement vers LASR	727
Télécharger des données vers un SAS LASR Analytic Server	728

À propos de la tâche Téléchargement vers LASR

Le SAS LASR Analytic Server est une plate-forme analytique in-memory spécifiquement conçue pour accélérer les temps de traitements analytiques exploratoires. Le serveur est disponible dans une configuration non distribuée pour des déploiements qui n'ont pas de gros volumes de données. Le serveur distribué utilise un environnement de calcul distribué et un traitement parallèle pour l'analyse des tables volumineuses. La tâche Téléchargement vers LASR permet de télécharger des données vers un LASR Analytic Server et donne accès aux données dans SAS Visual Analytics.

Remarque : Pour exécuter cette tâche, SAS Enterprise Guide ou SAS Add-In for Microsoft Office doit être connecté au même Metadata Server que SAS Visual Analytics.

Pour obtenir des informations sur la définition d'une connexion avec un SAS LASR Analytic Server, voir [SAS Intelligence Platform: Desktop Application Administration Guide](#)

Pour exécuter cette tâche, vous devez répondre aux exigences suivantes :

Nom du prérequis	Noms des procédures et des produits
Procédures SAS utilisées	Néant
Produits SAS requis	Base SAS, SAS LASR Analytic Server

Nom du prérequis	Noms des procédures et des produits
Autres produits SAS recommandés	SAS Visual Analytics

Télécharger des données vers un SAS LASR Analytic Server

- 1 Sélectionnez la bibliothèque où enregistrer les données
- 2 Indique le nom du membre de la table. Par défaut, le nom du membre est le même que le nom de la table. Vous ne pouvez pas changer le nom du membre si vous utilisez un chargement distribué.
- 3 Spécifie le chemin du répertoire pour les données téléchargées. Par défaut, chaque bibliothèque a un emplacement prédéfini.

Annexe 1

Options d'apparence courantes pour les tâches graphiques

Définir les options d'un axe	729
Voir aussi	732
Ajouter des lignes de référence	732
Voir aussi	733
Définir les options de légende	733
Voir aussi	734
Personnaliser l'aire	734
Voir aussi	735
Définir l'apparence des barres	735
Voir aussi	736
Définir les options des barres	736
Voir aussi	737
Définir les options des courbes	737
Voir aussi	738
Sélectionner la méthode d'interpolation	738
Voir aussi	739
Définir les options avancées	739
Voir aussi	741

Définir les options d'un axe

Dans la liste, ces options apparaissent sous l'en-tête Apparence > Axes.

Remarque : les options de l'axe ne sont pas appliquées si vous créez une surface en utilisant le format de sortie graphique GIF. Pour appliquer ces options à une surface, utilisez le format de sortie graphique ActiveX.

Nom du rôle	Description
Désactiver axes et graduations	supprime les axes et les graduations.
Axe	personnalise l'axe horizontal ou vertical. ■ Pour certains graphiques, vous pouvez également définir les options de l'axe vertical droit. ■ Pour les nuages de points 3D, les diagrammes en bâtons 3D et les surfaces 3D, vous pouvez définir ces options pour l'axe de profondeur. ■ Pour les graphiques groupés ou les histogrammes à mesures multiples, vous pouvez définir ces options pour l'axe de groupe. Pour personnaliser l'apparence d'un axe : ■ Pour les axes horizontaux, verticaux et de profondeur, indiquez la couleur, la largeur et le style de trait de l'axe. ■ Pour inverser l'ordre des valeurs de l'axe horizontal ou vertical, activez la case à cocher Inverser les axes. Si vous activez cette case à cocher pour l'axe horizontal, l'ordre des valeurs X est inversé. Si vous l'activez pour l'axe vertical, l'ordre des valeurs Y est inversé. Remarque : cette option n'est disponible que pour les aires, les boîtes à moustaches, les graphiques à bulles, les courbes de niveau, les courbes et les nuages de points en 2D. ■ Pour personnaliser le libellé d'un axe, cliquez sur l'onglet Libellé. Saisissez le texte du libellé personnalisé dans la zone Libellé. Utilisez la liste déroulante Rotation du libellé pour définir l'affichage du libellé par rapport à l'axe. Vous pouvez également définir le type, la taille et la couleur de la police. Vous pouvez souligner le texte ou le mettre en gras ou en italique. Utilisez les boutons pour modifier la justification. Remarque : si vous créez un histogramme à mesures multiples, le libellé par défaut de l'axe de réponse est la statistique utilisée pour le calcul de la longueur de chaque barre. ■ Pour personnaliser les valeurs d'un axe, cliquez sur l'onglet Valeurs. Vous pouvez définir le type, la taille et la couleur de la police des valeurs. Vous pouvez souligner le texte ou le mettre en gras ou en italique. Utilisez les boutons pour modifier la justification. Utilisez la liste déroulante

Nom du rôle	Description
Graduations principales	Rotation des valeurs pour définir l'affichage des valeurs par rapport à l'axe. spécifie les graduations principales pour l'axe horizontal ou vertical. Pour certains graphiques, vous pouvez également définir les options de l'axe vertical droit. Pour les nuages de points 3D, les diagrammes en bâtons 3D et les courbes de surface 3D, vous pouvez créer des graduations sur l'Axe de profondeur. Pour les histogrammes, vous pouvez spécifier des graduations uniquement pour l'axe de réponse. Par exemple, si vous créez un histogramme vertical, vous pouvez spécifier des graduations pour l'axe vertical mais pas pour l'axe horizontal. Sélectionnez Afficher les graduations principales nom-axe et indiquez la manière de déterminer le nombre de graduations à utiliser. ■ L'option Automatique laisse le logiciel déterminer le nombre et la fréquence des graduations principales. Si vous créez une aire, un graphique à bulles, une courbe ou un nuage de points en 2D, activez la case à cocher Commencer à zéro pour forcer le démarrage des graduations à 0. ■ Pour indiquer le nombre de graduations à utiliser, sélectionnez Utiliser. Saisissez le nombre souhaité dans le champ de saisie. ■ Pour spécifier un schéma de graduation logarithmique, sélectionnez Log. Sélectionnez une base, une puissance et indiquez si les graduations doivent être organisées en fonction des valeurs de la base ou de la puissance. ■ Pour indiquer des emplacements particuliers pour les graduations, sélectionnez Spécifier. Saisissez une valeur dans le champ de saisie, puis cliquez sur Ajouter. Pour modifier une valeur, sélectionnez-la dans la liste et cliquez sur Modifier. Une fois que vous avez terminé, cliquez sur Enregistrer.
Graduations secondaires	spécifie les graduations secondaires de l'axe horizontal ou vertical. ■ Pour certains graphiques, vous pouvez également définir les options de l'axe vertical droit. ■ Pour les nuages de points 3D, les diagrammes en bâtons 3D et les courbes de surface 3D, vous pouvez créer des graduations sur l'Axe de profondeur. ■ Pour les histogrammes, vous pouvez spécifier des graduations uniquement pour l'axe de réponse. Par exemple, si vous créez un histogramme vertical, vous

Nom du rôle	Description
	pouvez spécifier des graduations pour l'axe vertical mais pas pour l'axe horizontal.

Voir aussi

- Chapitre 3, "Aire" page 15
- Chapitre 6, "Histogramme" page 37
- Chapitre 8, "Courbe-histogramme" page 47
- Chapitre 11, "Boîte à moustaches" page 71
- Chapitre 12, "Graphique à bulles" page 79
- Chapitre 19, "Courbe de niveau" page 133
- Chapitre 42, "Courbe" page 305
- Chapitre 77, "Nuage de points" page 609
- Chapitre 87, "Surface" page 671

Ajouter des lignes de référence

Dans la liste, ces options apparaissent sous l'en-tête Apparence > Référence.

Remarque : ces options ne sont pas disponibles pour les nuages de points en 3D.

Nom du rôle	Description
Afficher la grille	affiche les lignes de la grille. Remarque : cette option n'est pas disponible pour les histogrammes.
Lignes de référence	ajoute des lignes de référence à l'axe horizontal ou vertical. Pour certains graphiques, vous pouvez également ajouter des lignes de référence pour l'axe vertical droit. Remarque : pour les histogrammes, des lignes de référence sont tracées pour l'axe de réponse. Elles se présentent sous forme de lignes horizontales pour les histogrammes verticaux et de lignes verticales pour les histogrammes

Nom du rôle	Description
	horizontaux. Pour les graphiques en 3D, aucune ligne de référence n'est disponible.

Voir aussi

- [Chapitre 3, "Aire" page 15](#)
- [Chapitre 6, "Histogramme" page 37](#)
- [Chapitre 8, "Courbe-histogramme" page 47](#)
- [Chapitre 11, "Boîte à moustaches" page 71](#)
- [Chapitre 12, "Graphique à bulles" page 79](#)
- [Chapitre 19, "Courbe de niveau" page 133](#)
- [Chapitre 42, "Courbe" page 305](#)
- [Chapitre 77, "Nuage de points" page 609](#)

Définir les options de légende

Dans la liste, sous l'en-tête Apparence, cliquez sur **Légende** pour accéder à ces options.

Vous pouvez créer des légendes pour plusieurs aires, plusieurs courbes, courbes-histogrammes, courbes de niveau, cartes, histogrammes à barres empilées, diagrammes circulaires, et graphiques en anneau. Les légendes sont disponibles pour plusieurs histogrammes de mesures si vous avez attribué une variable à la fonction **Empiler**.

Pour les courbes-histogrammes, vous pouvez créer deux légendes : une pour l'histogramme et une pour la courbe.

Elles sont également proposées si vous créez un nuage de points et si vous attribuez des variables aux fonctions **Vertical** et **Vertical droit**. Dans la mesure où plusieurs nuages de points sont représentés sur la même graphique, la légende permet de distinguer les points de données de la variable **verticale** de ceux de la variable **verticale droite**.

Voir aussi

- [Chapitre 3, "Aire" page 15](#)
- [Chapitre 6, "Histogramme" page 37](#)
- [Chapitre 8, "Courbe-histogramme" page 47](#)
- [Chapitre 19, "Courbe de niveau" page 133](#)
- [Chapitre 30, "Anneau" page 203](#)
- [Chapitre 42, "Courbe" page 305](#)
- [Chapitre 61, "Diagramme circulaire" page 471](#)

Personnaliser l'aire

Dans la liste, sous l'en-tête Apparence, cliquez sur **Aire** pour accéder à ces options.

Nom du rôle	Description
Indiquer taille personnalisée du graphe	spécifie la taille de l'image. La taille de votre écran est indiquée à titre de référence.
Couleur d'arrière-plan du graphe	spécifie la couleur d'arrière-plan. Remarque : si vous créez une carte, vous sélectionnez sa couleur d'arrière-plan dans la liste déroulante Couleur d'arrière-plan .
Tracer un cadre autour de la courbe	trace un cadre autour du graphique. Sélectionnez une couleur pour la courbe dans la liste Couleur d'arrière-plan de la courbe . Remarque : cette option n'est pas disponible pour les anneaux, les cartes, les diagrammes circulaires, les barres 3D groupées à trois dimensions, les nuages de points à trois dimensions, ni pour les surfaces.
Afficher des info-bulles de graphiques	spécifie si une info-bulle de graphique doit être affichée. Cette option s'applique à l'ensemble du graphique, et non à ses différents composants, tels que les barres ou les courbes. Remarque : l'info-bulle de graphique n'est pas visible pour les cartes ou si vous avez sélectionné le format de sortie graphique ActiveX ou Java.

Voir aussi

- Chapitre 3, "Aire" page 15
- Chapitre 6, "Histogramme" page 37
- Chapitre 8, "Courbe-histogramme" page 47
- Chapitre 11, "Boîte à moustaches" page 71
- Chapitre 12, "Graphique à bulles" page 79
- Chapitre 19, "Courbe de niveau" page 133
- Chapitre 30, "Anneau" page 203
- Chapitre 42, "Courbe" page 305
- Chapitre 61, "Diagramme circulaire" page 471
- Chapitre 69, "Carte radar" page 551
- Chapitre 77, "Nuage de points" page 609
- Chapitre 87, "Surface" page 671

Définir l'apparence des barres

Dans la liste, sous l'en-tête Apparence, cliquez sur **Barres** pour accéder à ces options.

Nom du rôle	Description
Définir la couleur des barres	indique la couleur des barres. Vous avez le choix entre ces options : ■ Jeu de couleurs par défaut utilise le jeu de couleurs par défaut. Ce dernier est déterminé par le style actuel. Vous pouvez indiquer le style à l'aide du Gestionnaire de styles ou de la boîte de dialogue Propriétés du résultat. Dans SAS Enterprise Guide, vous pouvez également spécifier le style de chaque type de résultat à partir de la boîte de dialogue Options. Dans SAS Add-In for Microsoft Office, vous pouvez appliquer des styles aux formats Rapport SAS et HTML à l'aide de la boîte de dialogue Options de SAS Add-In. ■ Couleurs existantes (Spécifiées dans une version précédente) permet d'utiliser les couleurs qui ont été spécifiées dans une version précédente de SAS Enterprise Guide ou SAS Add-In for Microsoft Office. Cette option

Nom du rôle	Description
	n'est disponible que pour le contenu SAS ayant été migré depuis une version antérieure. ■ Couleurs personnalisées permet de spécifier des couleurs personnalisées. Vous pouvez spécifier des couleurs personnalisées pour les histogrammes avec barres groupées, empilées et imbriquées, qui nécessitent plusieurs couleurs. En vous servant des listes déroulantes, vous pouvez indiquer jusqu'à 12 couleurs pour un histogramme. Si votre graphique nécessite plus de 12 couleurs, sélectionnez les couleurs supplémentaires dans le jeu de couleurs par défaut actuel. Cliquez sur Réinitialiser pour remplacer les valeurs personnalisées par les valeurs par défaut. Remarque : selon le type d'histogramme, il est possible que toutes les couleurs personnalisées ne puissent pas être utilisées.
Déterminer le nombre de barres	spécifie le nombre de barres dans le graphique. Par défaut, le nombre de barres est déterminé automatiquement. Pour les colonnes numériques, les valeurs sont divisées par intervalles, avec une barre par intervalle. Si la colonne numérique représente des valeurs discrètes, vous pouvez utiliser l'option Une barre pour chaque valeur de données unique. Remarque : vous ne pouvez pas spécifier le nombre de barres à utiliser dans les histogrammes à mesures multiples. Activez la case à cocher Spécifiez le nombre de barres pour indiquer le nombre de barres à faire apparaître dans le graphique.

Voir aussi

- [Chapitre 6, "Histogramme" page 37](#)
- [Chapitre 8, "Courbe-histogramme" page 47](#)

Définir les options des barres

Sous l'en-tête Apparence de la liste, cliquez sur **Options de la barre** pour accéder à ces options.

Utilisez la liste déroulante **Ordre** pour indiquer l'ordre de tri des barres. Par défaut, il est déterminé automatiquement. Vous pouvez toutefois choisir d'organiser les barres par ordre croissant ou décroissant de longueur ou de hauteur.

Activez la case à cocher **Couleur du contour** si vous voulez appliquer un contour à chaque barre. Vous pouvez sélectionner une couleur de contour dans la liste déroulante.

Utilisez la liste déroulante **Taille de la barre** pour définir la taille des barres. Par défaut, elle est déterminée automatiquement. Vous pouvez toutefois saisir la largeur de barre ou l'espacement entre les barres sous forme de pourcentage de la largeur totale du graphique.

Voir aussi

Chapitre 8, "Courbe-histogramme" page 47

Définir les options des courbes

Dans la liste, sous l'en-tête Apparence, cliquez sur **Graphes** pour accéder à ces options.

Remarque : ces options ne sont disponibles que si vous attribuez au moins une variable à la fonction **Somme de la courbe de**. Si vous n'avez pas affecté de variable à la fonction **Somme de la ligne de**, vous pouvez utiliser la liste déroulante **Statistique utilisée pour tracer la courbe** du volet **Avancé** pour sélectionner la statistique à utiliser pour tracer la courbe d'une courbe-histogramme.

Pour contrôler l'apparence de la courbe pour une variable spécifique :

- 1 Sélectionnez une variable dans la liste. Les variables répertoriées dans la liste sont celles que vous avez attribuées à la fonction **Somme de la courbe de**.
- 2 Dans la zone **Trait**, indiquez le style, l'épaisseur et la couleur du trait.
- 3 Sélectionnez une couleur de contour.
- 4 Dans la zone **Symbole pour un point de données**, spécifiez le type du symbole, le symbole à utiliser, la hauteur de chaque symbole, ainsi que la couleur des symboles des points de données.

Remarque : certains symboles ne sont pas pris en charge par tous les formats de sortie graphique. Si vous sélectionnez un symbole non pris en charge par le format de sortie, un autre symbole (pris en charge) s'affiche.

- 5 Dans la liste déroulante **Statistique utilisée pour tracer la courbe**, sélectionnez la statistique à utiliser pour tracer la courbe dans une courbe-histogramme. Vous avez le choix entre ces statistiques :
- **Moyenne** est la valeur moyenne de la variable attribuée au rôle **Somme de la courbe de**. Cette moyenne est calculée pour chaque valeur unique, représentée par un point de la courbe. Par exemple, si vous souhaitez afficher les ventes moyennes pour chaque ville de la table, attribuez la variable Ventes à la fonction **Somme de la courbe de** et sélectionnez la statistique **Moyenne** pour tracer la courbe.
 - **Somme** est la somme des valeurs de la variable attribuée à la fonction **Somme de la courbe de**. Ce total est calculé pour chaque valeur unique, représentée par un point de la courbe. Supposons, par exemple, que vous attribuez la variable Ventes à la fonction **Somme de la courbe de**. Si les valeurs de cette variable sont 8734, 982, 1504, 3207, 4502, 624 et 918 pour la ville de Nantes, la statistique Somme de la courbe représentant Nantes est 20.471.

Voir aussi

Chapitre 8, "Courbe-histogramme" page 47

Sélectionner la méthode d'interpolation

Dans la liste, sous l'en-tête Apparence, cliquez sur **Interpolations** pour accéder à ces options.

Pour sélectionner une méthode d'interpolation :

- 1 Sélectionner une variable. Si vous créez une courbe ou un nuage de points en 2D, les variables de la liste sont celles que vous avez attribuées aux fonctions **Vertical** et **Vertical (droite)**. Si vous avez sélectionné le type de courbe **Plusieurs courbes par colonne de groupe**, la liste contient les valeurs uniques de la variable que vous avez attribuée à la fonction **Groupe**.
- 2 Sélectionnez une méthode d'interpolation dans la liste déroulante. En fonction de la méthode d'interpolation que vous choisissez, vous devrez peut-être spécifier des paramètres supplémentaires. Ces méthodes d'interpolation sont disponibles :
 - **Lagrange** trace une courbe lissée reliant les points de données. Cette méthode est principalement utilisée pour les données correspondant à des valeurs précises mises en tableau. Un polynôme du degré indiqué (1, 3 ou 5) est ajusté sur les 2, 4 ou 6 points les plus proches. En général, la première valeur dérivée n'est pas continue. Si les valeurs de la variable X ne sont pas

strictement croissantes, la méthode paramétrique correspondante est utilisée.

- **Courbe** relie les points de données par des lignes droites. Les points sont reliés dans l'ordre dans lequel ils apparaissent dans les données d'entrée.
 - La méthode **Bâton** trace un trait vertical à partir de chaque point de données jusqu'à un trait horizontal à la valeur 0 sur l'axe vertical ou à la valeur minimale sur l'axe vertical si celle-ci est supérieure à 0. La ligne horizontale est tracée automatiquement.
 - **Régression** crée un graphique d'analyse de régression. Cette méthode d'interpolation n'est disponible que si vous attribuez des variables numériques aux fonctions Horizontal et Vertical.
 - **Nuage de points** supprime toute interpolation. Une courbe des points de données est créée.
 - **Courbe lissée** ajuste une courbe lissée par rapport aux données à l'aide d'une routine de lissage. Cette méthode permet de lisser les données bruyantes. Les points du tracé ne se trouvent pas systématiquement sur la ligne.
 - **Spline** trace le trait en utilisant une routine spline.
 - **Écart-type** utilise une ligne pleine pour relier la valeur Y moyenne à des écarts-types de ± 1 , 2 ou 3 pour chaque X. Par défaut, deux écarts-types sont utilisés.
 - **Escalier** représente les données sous forme d'un escalier en utilisant la fonction par paliers.
- 3 (Facultatif) Activez la case à cocher **Appliquer à tout** pour utiliser les mêmes méthode d'interpolation et propriétés pour toutes les variables de la liste. Cette option n'est disponible que si la liste comporte plusieurs variables.

Voir aussi

- [Chapitre 42, "Courbe" page 305](#)
- [Chapitre 77, "Nuage de points" page 609](#)

Définir les options avancées

Dans la liste, sous l'en-tête Apparence, cliquez sur **Options avancées** pour accéder à ces options.

Nom du rôle	Description
Statistique utilisée pour calculer la barre	spécifie comment calculer la longueur de la barre. Les statistiques disponibles varient selon le graphique créé. ■ Si aucune variable n'a été attribuée à la fonction Somme de dans la tâche Histogramme ou à la fonction Somme de la barre de dans la tâche Courbe-histogramme, les statistiques Fréquence, Fréquence cumulée, Pourcentage et Pourcentage cumulé sont disponibles. ■ Si vous avez attribué une variable à la fonction Somme de ou Somme de la barre de, ou si vous avez sélectionné le type d'histogramme Groupe (vertical ou horizontal) de plusieurs mesures, les statistiques Somme et Moyenne sont disponibles.
Statistique utilisée pour tracer la courbe	spécifie la statistique à utiliser lors du traçage de la ligne sur une courbe-histogramme. Les statistiques Fréquence, Effectifs cumulés, Pourcentage et Pourcentage cumulé sont alors disponibles. Si vous avez attribué une variable à la fonction Somme de la courbe de dans la tâche Courbe-histogramme, cette option n'est pas disponible dans le volet Avancé. En revanche, elle est proposée dans le volet Graphes, afin que vous puissiez déterminer la statistique de chaque courbe lorsque plusieurs variables sont attribuées à la fonction Somme de la courbe de.
Afficher les barres d'erreur	affiche les intervalles de confiance sur les barres. Dans la liste déroulante Type, sélectionnez l'une des options suivantes : ■ Barres représente les barres d'erreur sous forme de barres dont la largeur est égale à la moitié de celle des barres principales. ■ Les deux représente les barres d'erreur sous forme de deux graduations reliées par une ligne. ■ Haut représente la barre d'erreur sous la forme d'une graduation pour l'intervalle confiance supérieur relié au haut de la barre par une ligne. Sélectionnez une couleur pour les barres d'erreur et utilisez la zone CLM pour indiquer un niveau de confiance. Remarque : vous pouvez afficher les barres d'erreur uniquement lorsque : ■ Vous avez sélectionné la statistique Moyenne ou Pourcentage pour le calcul de la longueur des barres ■ Vous ne créez pas de graphe en 3D à barres groupées ■ Vous n'avez pas attribué de variable à la fonction Empiler

Nom du rôle	Description
Accepter les valeurs manquantes	accepte une valeur manquante comme valeur possible pour la colonne du graphique.
Ignorer les valeurs zéro	supprime toutes les valeurs de la colonne dont la longueur de barre est égale à zéro.
Calculer % + % cumulés pour chaque gpe (G100)	calcule les pourcentages et les pourcentages cumulés séparément pour chaque groupe dans des diagrammes à barres groupés. Lorsque vous sélectionnez cette option, les différents pourcentages reflètent la contribution du milieu au groupe ; ils totalisent 100 pour cent pour chaque groupe. Remarque : cette option n'est disponible que lorsque vous attribuez une variable à la fonction Grouper les barres par dans le cadre de graphes à barres groupées.
Afficher les statistiques près de la barre	affiche la valeur de la statistique utilisée pour calculer la barre. Remarque : cette option n'est disponible que pour les histogrammes horizontaux.
Indiquez une valeur statistique à afficher pour les barres	affiche la valeur d'une statistique que vous avez sélectionnée dans la liste déroulante. La statistique peut être la même ou non que celle utilisée pour calculer la longueur de la barre. Si vous créez un histogramme en 2D, vous pouvez choisir d'afficher la valeur de la statistique à l'intérieur de la barre.

Voir aussi

- [Chapitre 6, "Histogramme" page 37](#)
- [Chapitre 8, "Courbe-histogramme" page 47](#)
- [Chapitre 19, "Courbe de niveau" page 133](#)

Annexe 2

Options d'apparence courantes pour les tâches de la carte de contrôle

<i>Définir les options d'un axe</i>	743
<i>Ajouter des lignes de référence</i>	744
<i>Ajouter un encart</i>	744
<i>Définir d'autres options de tracé</i>	744
<i>Définir les options de variable de bloc</i>	746

Définir les options d'un axe

Dans la liste, sous l'en-tête **Apparence** ⇒ **Axes**, cliquez sur **Axes** pour accéder à ces options.

Vous pouvez définir ces options :

- la couleur des axes
- la largeur des axes
- la couleur des graduations
- un libellé pour les axes horizontaux et verticaux

Ajouter des lignes de référence

Dans la liste, sous l'en-tête **Apparence** ⇒ **Axes**, cliquez sur **Horizontal** ou **Vertical** pour accéder à ces options.

Pour utiliser les lignes de référence, sélectionnez **Utiliser les lignes de référence**. Vous pouvez formater la ligne de référence et spécifier l'emplacement de chaque ligne de référence.

Ajouter un encart

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Encart** pour accéder à ces options.

Vous pouvez améliorer un graphe en ajoutant une case ou une table (appelée encart) contenant les statistiques descriptives directement sur le graphe. Si vous créez un histogramme comparatif, un encart apparaît dans chaque histogramme.

Sélectionnez **Inclure l'encart** pour ajouter des statistiques descriptives au graphe. Sélectionnez ensuite les statistiques que vous voulez inclure dans l'encart.

Remarque : les statistiques permettant de calculer les indices de capacité ne sont disponibles que si vous avez créé des limites de spécification pour la variable attribuée à la fonction **Variables d'analyse**. Vous pouvez créer des limites de spécification pour les histogrammes, les courbes CDF, les courbes de probabilité et les diagrammes Q-Q. Pour plus d'informations sur le calcul d'indices de capacité, consultez l'aide relative à la procédure CAPABILITY.

Définir d'autres options de tracé

Dans la liste, sous l'en-tête **Apparence**, cliquez sur **Options** pour accéder à ces options.

Nom d'option	Description
Dans les limites de contrôle, Limites de contrôle, Connexion des segments, Segments en dehors des limites, Zones en dehors des limites et Cadre	spécifie la couleur de ces éléments graphiques. ■ Dans les limites de contrôle spécifie la couleur à l'intérieur des limites de contrôle inférieure et supérieure. ■ Limites de contrôle spécifie la couleur des limites de contrôle et de la ligne centrale. Spécifie également la couleur des libellés de ces lignes. ■ Relier les segments de ligne - spécifie la couleur des segments de ligne qui relient le point à la courbe. ■ Segments hors limites - spécifie la couleur des symboles du tracé et les portions des segments de lignes de connexion situées en dehors des limites de contrôle. Cette option est utile pour sélectionner les points hors de contrôle. ■ Zones en dehors des limites - indique la couleur de remplissage des zones extérieures aux limites de contrôle, situées entre les points connectés et les limites de contrôle et reliées par des lignes de connexion. Cette option est utile pour sélectionner les points hors de contrôle. ■ Cadre spécifie les couleurs de remplissage du rectangle délimité par les axes et le cadre.
Symbole	spécifie le symbole pour les points de données.
Nombre de pages affichées	spécifie le nombre de pages à utiliser pour afficher le graphique.
Étendue mobile	Spécifie le nombre de mesures consécutives à partir desquelles les étendues mobiles sont calculées. La valeur spécifiée doit être comprise entre 2 et 10 (inclus). La valeur par défaut est 2. L'intervalle de 2 à 10 est propre à cette tâche. Une valeur supérieure peut être utilisée avec la procédure SHEWHART de SAS. Pour plus d'informations, consultez la procédure SHEWHART dans le Base SAS Procedures Guide. Remarque : Cette option n'est disponible que pour la tâche Carte aux mesures individuelles.
Supprimer légende par défaut	supprime la légende par défaut de la taille de l'échantillon du sous-groupe. Cette option est utile si les tailles des échantillons des sous-groupes sont constantes et égales à la taille de l'échantillon de limite de contrôle, car cette dernière apparaît automatiquement dans l'angle supérieur droit du graphique.

Nom d'option	Description
Utiliser une valeur num. pour le sous-groupe (au lieu d'une var.)	spécifie une valeur constante pour la taille. Utilisez cette option lorsque la taille de l'échantillon de sous-groupe est fixe. Saisissez la valeur à utiliser dans la zone Taille de l'échantillon du sous-groupe. La taille de l'échantillon du sous-groupe doit être une valeur comprise entre la valeur maximale de la variable de traitement et 99999. Remarque : si vous attribuez une variable à la fonction Taille d'échantillon du sous-groupe, cette option est désélectionnée et la variable attribuée à cette fonction est employée comme taille de l'échantillon du sous-groupe. Remarque : cette option n'est disponible que pour les cartes np, p et u.
Omettre le graphique secondaire	empêche l'affichage d'un graphique secondaire. En général, cette option s'utilise avec une Carte aux mesures individuelles pour créer des mesures individuelles et pour supprimer le graphe d'accompagnement pour les étendues mobiles. Remarque : cette option n'est disponible que pour les tâches Carte aux mesures individuelles, Carte de la moyenne et de l'étendue et Carte de la moyenne et de l'écart-type.
Type de graphe en marge	spécifie le type du graphe en marge qui apparaîtra sur la Carte aux mesures individuelles. Un graphe en marge est un graphe univarié des statistiques de la carte de contrôle. Vous pouvez sélectionner un nuage de points, un histogramme ou une boîte à moustaches. Le graphe apparaît soit sur la gauche soit sur la droite de la carte aux mesures individuelles. La sélection par défaut est Aucun graphe partiel. Remarque : cette option n'est disponible que pour la tâche Carte aux mesures individuelles.

Définir les options de variable de bloc

Dans la liste, sous l'en-tête Apparence, cliquez sur **Options du bloc** pour accéder à ces options.

Remarque : ces options ne sont disponibles que si vous avez attribué une variable à la fonction **Variables de bloc**.

Nom du rôle	Description
Libellé du bloc	spécifie le libellé pour les variables de bloc.
Position de la légende	spécifie la position verticale de la légende pour les valeurs des variables du bloc.
Valeurs de durée	spécifie la manière dont les valeurs des variables de bloc de durée sont traitées lorsqu'il n'y a pas suffisamment de place pour les afficher dans la légende du bloc. Par défaut, les valeurs de durée ne sont pas affichées. Si vous activez l'option Réduire en hauteur de façon uniforme , la hauteur du texte des valeurs est réduite pour s'ajuster à la légende. Si vous sélectionnez l'option Tronquer sur la droite , les valeurs de durée sont tronquées sur la droite pour s'ajuster à la légende.
Position du libellé	spécifie la position d'un libellé des variables de bloc dans la légende du bloc. Les options disponibles varient selon la position de la légende. Vous avez le choix entre ces options : ■ Au-dessus - place le libellé immédiatement au-dessus de la légende. Il s'agit de la présentation par défaut. ■ Gauche - place le libellé à gauche de la légende ■ Droite - place le libellé à droite de la légende. Cette option n'est pas disponible si vous avez choisi de placer la légende au-dessus de la carte de contrôle. Ne sélectionnez la position Gauche ou Droite que pour les libellés suffisamment courts pour ne pas dépasser les marges situées de chaque côté du graphique. Les libellés trop longs sont en effet tronqués.
Couleur du libellé du bloc	spécifie la couleur du texte du libellé du bloc. Remarque : cette option n'est disponible que si vous sélectionnez Au-dessus pour la position du libellé.
Afficher des valeurs répétées dans la légende	affiche les valeurs de la variable de bloc pour tous les sous-groupes. Par défaut, seule la première variable de bloc est affichée ; les valeurs répétées ne le sont pas.

Annexe 3

Créer un modèle

<i>À propos des modèles</i>	749
Voir aussi	749
<i>Indiquer les effets principaux</i>	750
<i>Spécifier les effets croisés (interactions)</i>	750
<i>Spécifier des effets imbriqués</i>	750
<i>Créer plusieurs croisements de deux variables sélectionnées ou plus</i>	751
<i>Créer des croisements polynomiaux de variables quantitatives</i>	751
<i>Supprimer les effets de l'intercept dans l'analyse</i>	751

À propos des modèles

Dans la liste, cliquez sur **Modèle** pour accéder à ces options.

Par défaut, aucun effet n'est spécifié, et la tâche ajuste alors un modèle avec uniquement un intercept. Pour spécifier un effet, vous devez attribuer au moins une variable à la fonction **Variables quantitatives** ou **Variables de classification**. Vous pouvez sélectionner des combinaisons de variables pour créer des effets croisés, imbriqués, factoriels, ou polynomiaux.

Voir aussi

- Chapitre 37, "Régression linéaire High-Performance" page 253
- Chapitre 38, "Régression logistique High-Performance" page 265
- Chapitre 48, "Régression logistique" page 353

Indiquer les effets principaux

- 1 Sélectionnez le nom de la variable dans **Variables de classe et var. quantitatives**.
- 2 Cliquez sur **Principal** pour ajouter la variable dans **Effets**.

Spécifier les effets croisés (interactions)

- 1 Sélectionnez au moins un nom de variable. Pour sélectionner plus d'une variable, appuyez sur Ctrl tout en sélectionnant les noms de variables supplémentaires.
- 2 Cliquez sur **Croiser**. Si vous sélectionnez une seule variable, la variable sera croisée avec elle-même.

Spécifier des effets imbriqués

- 1 Sélectionnez un ou plusieurs effets principaux ou effets croisés que vous voulez imbriquer.
- 2 Sélectionnez une ou plusieurs variables de classe dans lesquelles vous voulez imbriquer les effets principaux ou les effets croisés.
- 3 Cliquez sur **Imbriquer**. L'effet imbriqué apparaît dans la zone **Effets** sous la forme *variable 1(variable 2)*.

Créer plusieurs croisements de deux variables sélectionnées ou plus

- 1 Vous pouvez indiquer des croisements à des degrés plus élevés en ajustant la valeur du champ **Degrés**.
- 2 Cliquez sur **Factoriel** pour ajouter les effets factoriels dans la zone **Effets**. Par défaut, chaque combinaison de deux variables est utilisée.

Créer des croisements polynomiaux de variables quantitatives

- 1 Vous pouvez indiquer des croisements à des degrés plus élevés en ajustant la valeur du champ **Degrés**.
- 2 Cliquez sur **Polynomial** pour ajouter les effets factoriels dans la zone **Effets**.

Supprimer les effets de l'intercept dans l'analyse

Désactivez la case à cocher **Avec intercept**. Dans ce cas, vous devez avoir spécifié au moins un effet. Par défaut, l'intercept est inclus.

Annexe 4

Tâches intégrées

Tâches intégrées et Procédures et licences SAS associées 753

Tâches intégrées et Procédures et licences SAS associées

Le tableau suivant répertorie les tâches intégrées pour SAS Enterprise Guide et SAS Add-In for Microsoft Office. Chaque tâche liste les procédures SAS principales pour cette tâche ainsi que les licences de logiciel SAS requises pour exécuter une tâche.

Remarque : votre version de SAS Enterprise Guide ou SAS Add-In for Microsoft Office n'inclut peut-être pas toutes les tâches de la liste ci-dessous. La disponibilité d'une tâche dépend des licences et installations des logiciels SAS nécessaires sur votre site. Pour plus d'informations, contactez votre administrateur système.

Si vous disposez de licences et installations de SAS Studio sur votre site, d'autres tâches peuvent être disponibles. Pour plus d'informations, voir Tâches de SAS Studio et Procédures et licences associées.

Tâche intégrée	Procédures SAS utilisées	Licence SAS requise
Ajouter une table	SQL	Base SAS
Aire	GPLOT	SAS Base, SAS/GRAPH
Modélisation et prévision ARIMA	ARIMA	SAS Base, SAS/ETS

Tâche intégrée	Procédures SAS utilisées	Licence SAS requise
Affecter une bibliothèque au projet	LIBNAME	Base SAS
Histogramme	GCHART	SAS Base, SAS/GRAPH
Assistant Histogramme page 43	GCHART	SAS Base, SAS/GRAPH
Courbe-histogramme	GBARLINE	SAS Base, SAS/GRAPH
Prévision de base	FORECAST	SAS Base, SAS/ETS
Carte de la moyenne avec superposition des boîtes de Tukey	SHEWHART	SAS Base, SAS/QC
Boîte à moustaches	GPLOT	SAS Base, SAS/GRAPH
Graphique à bulles	GPLOT	SAS Base, SAS/GRAPH
Carte c page 93	SHEWHART	SAS Base, SAS/QC
Corrélation canonique	CANCORR	SAS Base, SAS/STAT
Courbes CDF	CAPABILITY	SAS Base, SAS/QC
Caractéristiques des données	CONTENTS, UNIVARIATE, FREQ	Base SAS
Analyse de classification	CLUSTER, FASTCLUS, TREE	SAS Base, SAS/STAT
Comparer les données	COMPARE	Base SAS
Courbe de niveau	GCONTOUR, G3GRID	SAS Base, SAS/GRAPH
Corrélations	CORR	Base SAS
Créer un format	FORMAT	Base SAS
Créer un format à partir d'une table page 159	FORMAT	Base SAS
Créer une table de caractéristiques cartographiques page 163	GPROJECT	SAS Base, SAS/GRAPH

Tâche intégrée	Procédures SAS utilisées	Licence SAS requise
Créer les données d'une série chronologique	TIMESERIES	SAS Base, SAS/ETS
Attributs de la table	DATASETS	Base SAS
Supprimer les tables et les formats	SQL, CATALOG	Base SAS
Analyse discriminante	DISCRIM	SAS Base, SAS/STAT
Analyse de distribution	UNIVARIATE	Base SAS
Anneau	GCHART	SAS Base, SAS/GRAPH
Télécharger des fichiers de données vers un PC	SQL	Base SAS
Analyse factorielle	FACTOR	SAS Base, SAS/STAT
Filtrer et trier*	SQL	Base SAS
Créer un projet Forecast Studio	Néant	SAS Forecast Server
Ouvrir un projet Forecast Studio	Néant	SAS Forecast Server
Projet de valeur corrigée Forecast Studio page 237	Néant	SAS Forecast Server
Modèle linéaire généralisé	GENMOD	SAS Base, SAS/STAT
Régression linéaire High-Performance	HPREG	SAS Base, SAS High Performance Analysis Server, SAS/STAT
Régression logistique High Performance	HPLOGISTIC	SAS Base, SAS High Performance Analysis Server, SAS/STAT
Histogrammes page 277	CAPABILITY	SAS Base, SAS/QC
Importer des données (Microsoft Excel, Fichiers texte, Microsoft Access)*	Étape DATA, IMPORT	Base SAS, SAS/ACCESS to PC File Formats
Importer un fichier JMP*	Étape DATA, IMPORT	Base SAS, SAS/ACCESS to PC File Formats

Tâche intégrée	Procédures SAS utilisées	Licence SAS requise
Importer un fichier SPSS*	Étape DATA, IMPORT	Base SAS, SAS/ACCESS to PC File Formats
Importer un fichier Stata*	Étape DATA, IMPORT	Base SAS, SAS/ACCESS to PC File Formats
Carte aux mesures individuelles	SHEWHART	SAS Base, SAS/QC
Générateur d'applications stockées JMP*	étape DATA	JMP
Tables de survie	LIFETEST	SAS Base, SAS/STAT
Courbe	GPLOT	SAS Base, SAS/GRAPH
Assistant Courbe	GPLOT	SAS Base, SAS/GRAPH
Modèles linéaires	GLM	SAS Base, SAS/STAT
Régression linéaire	REG	SAS Base, SAS/STAT
Liste des données	PRINT	Base SAS
Assistant Rapport sous forme de liste page 349	REPORT	Base SAS
Régression logistique	LOGISTIC	SAS Base, SAS/STAT
Carte	GMAP	SAS Base, SAS/GRAPH
Carte de la moyenne et de l'étendue	SHEWHART	SAS Base, SAS/QC
Carte de la moyenne et de l'écart-type	SHEWHART	SAS Base, SAS/QC
Modèles mixtes	MIXED	SAS Base, SAS/STAT
Scoring du modèle	Néant	SAS Enterprise Miner
Régression non linéaire	NLIN	SAS Base, SAS/STAT
Analyse de variance non-paramétrique à une dimension	NPAR1WAY	SAS Base, SAS/STAT
Carte np	SHEWHART	SAS Base, SAS/QC

Tâche intégrée	Procédures SAS utilisées	Licence SAS requise
Analyse de variance à une dimension	ANOVA	SAS Base, SAS/STAT
Fréquences à un critère de classification	FREQ	Base SAS
SAS ODS Statistical Graphs	Néant	ODS Graphics Designer
Carte p	SHEWHART	SAS Base, SAS/QC
Graphique de Pareto	PARETO	SAS Base, SAS/QC
Diagramme circulaire	GCHART	SAS Base, SAS/GRAPH
Assistant Diagramme circulaire	GCHART	SAS Base, SAS/GRAPH
Courbes P-P	CAPABILITY	SAS Base, SAS/QC
Préparer les données d'une série chronologique	EXPAND	SAS Base, SAS/ETS
Composantes principales	PRINCOMP	SAS Base, SAS/STAT
Courbes de probabilité	CAPABILITY	SAS Base, SAS/QC
Hasards proportionnels	PHREG	SAS Base, SAS/STAT
Courbes Q-Q	CAPABILITY	SAS Base, SAS/QC
Générateur de requêtes*	SQL	Base SAS
Carte radar	GRADAR	SAS Base, SAS/GRAPH
Échantillon aléatoire	SURVEYSELECT	SAS Base, SAS/STAT
Rang	RANK	Base SAS
Rapid Predictive Modeler	macros SAS Enterprise Miner	SAS Enterprise Miner
Analyse RFM (Récence, Fréquence et Montant)	macro SAS	Base SAS
Analyse de régression des données du panel	TSCSREG	SAS Base, SAS/ETS

Tâche intégrée	Procédures SAS utilisées	Licence SAS requise
Analyse de la régression avec erreurs autorégressives	AUTOREG	SAS Base, SAS/ETS
Nuage de points	GPLOT, G3D	SAS Base, SAS/GRAPH
Matrice nuage de points	SGSCATTER	SAS Base, SAS/GRAPH
Afficher le graphique statistique ODS	SGDESIGN	Base SAS
Trier les données	SORT	Base SAS
Fractionner des colonnes	TRANSPOSE	Base SAS
Empiler colonnes	TRANSPOSE	Base SAS
Normaliser les données	STANDARD	Base SAS
Statistiques descriptives	MEANS	Base SAS
Assistant Statistiques descriptives	MEANS	Base SAS
Tables de synthèse page 659	TRANSPOSE	Base SAS
Assistant Tables de synthèse	TRANSPOSE	Base SAS
Surface	G3D	SAS Base, SAS/GRAPH
Test t	TTEST	SAS Base, SAS/STAT
Analyse d'une table	FREQ	Base SAS
Diagramme en mosaïque	GTILE	SAS Base, SAS/GRAPH
Transposer	TRANSPOSE	Base SAS
Carte u	SHEWHART	SAS Base, SAS/QC
Actualiser les métadonnées de la bibliothèque*	METALIB	Base SAS
Télécharger des fichiers de données vers un Serveur*	SQL	Base SAS

Remarque : les tâches suivies d'un astérisque ne sont disponibles que dans SAS Enterprise Guide.
