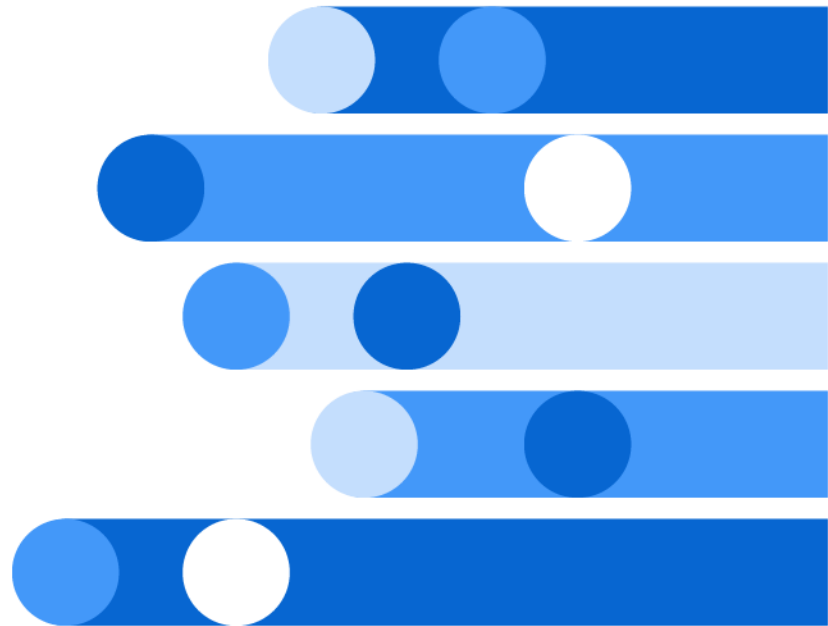




SAS[®] Enterprise Guide[®] 8.4 및 SAS[®] Add-In 8.4 for Microsoft Office의 SAS[®] 작업

8.4*



* 이 문서는 추가적인 소프트웨어 버전에 해당될 수 있습니다. 이 문서를 [SAS Help Center](#) 에서 열고 배너에서 버전을 클릭하여 사용 가능한 모든 버전을 확인하십시오.



이 설명서의 올바른 참고문헌 인용은 다음과 같습니다. SAS Institute Inc. 2024. *SAS® Enterprise Guide® 8.4 및 SAS® Add-In 8.4 for Microsoft Office의 SAS® 작업*. Cary, NC: SAS Institute Inc.

SAS® Enterprise Guide® 8.4 및 SAS® Add-In 8.4 for Microsoft Office의 SAS® 작업

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

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

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

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

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

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

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

April 2024

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

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

8.4-P1:egamotasks

목차

1장 / 작업 사용	1
빌트인 작업 및 SAS Studio 작업	1
작업 데이터 필터링	3
작업에 사용자 코드 추가	4
칼럼 속성 지정	4
그래픽 출력형식 선택	5
작업 템플릿 작업	5
2장 / 테이블 추가 작업	9
테이블 추가 작업 정보	9
테이블 추가: 입력 테이블 리스트 작성	10
테이블 추가: 결과 저장	10
3장 / 영역 도표	11
영역 도표 작업 정보	11
영역 도표: 도표 유형 선택	12
영역 도표: 분석 역할에 변수 할당	12
영역 도표: 모양 옵션 설정	13
제목 및 각주 지정	14
속성 보기	14
4장 / ARIMA 모델화 및 예측 작업	17
ARIMA 모델 개발 및 예측 작업 정보	17
ARIMA 모델화 및 예측: 분석 역할에 변수 할당	18
ARIMA 모델화 및 예측: 식별 옵션 설정	19
ARIMA 모델화 및 예측: 추정 옵션 설정	20
ARIMA 모델화 및 예측: 예측 옵션 설정	22
제목 및 각주 지정	24
속성 보기	24
5장 / 프로젝트 라이브러리 할당	25
프로젝트 라이브러리 할당 마법사 정보	25
프로젝트 라이브러리 할당: 이름 및 서버 지정	26
프로젝트 라이브러리 할당: 엔진 지정	26
프로젝트 라이브러리 할당: 구성 옵션 지정	28
프로젝트 라이브러리 할당: 요약 정보 검토	29
6장 / 막대 그래프	31
막대 그래프 작업 정보	31
막대 그래프: 그래프 유형 선택	32
막대 그래프: 분석 역할에 변수 할당	33
막대 그래프: 모양 옵션 설정	34
제목 및 각주 지정	35
속성 보기	35

7장 / 막대 그래프 마법사	37
막대 그래프 마법사 정보	37
막대 그래프 마법사: 데이터 선택	37
막대 그래프 마법사: 역할에 변수 할당	37
막대 그래프 마법사: 모양 옵션 설정	38
막대 그래프 마법사: 제목 및 각주 지정	39
8장 / 막대-선 그래프	41
막대-선 그래프 작업 정보	41
막대-선 그래프: 분석 역할에 변수 할당	42
막대-선 그래프: 모양 옵션 설정	43
제목 및 각주 지정	44
속성 보기	44
9장 / 기본 예측 모형	45
기본 예측 모형 작업 정보	45
기본 예측: 분석 역할에 변수 할당	46
기본 예측: 예측 옵션 설정	47
기본 예측: 도표 생성	49
기본 예측: 결과 옵션 설정	49
제목 및 각주 지정	50
속성 보기	50
10장 / 상자관리도	51
상자관리도 정보	51
관리도: 분석 역할에 변수 할당	52
관리도: 관리한계 지정	53
관리도: 수행할 검정 선택	54
관리도: 도표 모양 옵션 설정	55
관리도: 분석 결과 저장	56
제목 및 각주 지정	57
속성 보기	57
11장 / 상자 도표	59
상자 도표 작업 정보	59
상자 도표: 도표 유형 선택	60
상자 도표: 분석 역할에 변수 할당	60
상자 도표: 모양 옵션 설정	61
제목 및 각주 지정	63
속성 보기	63
12장 / 버블 도표	65
버블 도표 작업 정보	65
버블 도표: 분석 역할에 변수 할당	66
버블 도표: 모양 옵션 설정	67
제목 및 각주 지정	67
속성 보기	68
13장 / 정준상관분석	69
정준상관분석 작업 정보	69
정준상관분석: 분석 역할에 변수 할당	70
정준상관분석: 통계량 옵션 설정	71

정준상관분석: 도표 생성	72
정준상관분석: 결과 옵션 설정	72
제목 및 각주 지정	73
속성 보기	73
14장 / 결점수관리도	75
결점수관리도 작업 정보	75
결점수관리도: 분석 역할에 변수 할당	76
관리도: 관리한계 지정	77
관리도: 수행할 검정 선택	78
관리도: 도표 모양 옵션 설정	79
관리도: 분석 결과 저장	79
제목 및 각주 지정	80
속성 보기	80
15장 / CDF 도표	83
CDF 도표 작업 정보	83
공정능력분석: 분석 역할에 변수 할당	84
CDF 도표: 분포 선택	85
CDF 도표: 분포 사용자 정의	86
공정능력분석: 도표 모양 옵션 설정	86
공정능력분석: 분석 유형 지정	89
제목 및 각주 지정	90
속성 보기	90
16장 / 데이터 특성화	93
데이터 특성화 마법사 정보	93
데이터 특성화: 데이터 선택	93
데이터 특성화: 리포트 옵션 선택	94
데이터 특성화: 출력 한계 지정	94
17장 / 군집분석	95
군집분석 작업 정보	95
군집분석: 분석 역할에 변수 할당	96
군집분석: 군집 옵션 설정	97
군집분석: 도표 및 다이어그램 생성	98
군집분석: 결과 옵션 설정	99
제목 및 각주 지정	100
속성 보기	101
18장 / 데이터 비교	103
데이터 비교 마법사 정보	103
데이터 비교: 데이터 선택	104
데이터 비교: 비교할 변수 선택	104
데이터 비교: 비교 기준 지정	104
데이터 비교: 출력 옵션 지정	105
19장 / 등고선 도표	107
등고선 도표 작업 정보	107
등고선 도표: 도표 유형 선택	108
등고선 도표: 분석 역할에 변수 할당	108
등고선 도표: 모양 옵션 설정	109

제목 및 각주 지정	110
속성 보기	110
20장 / 파일 복사	113
파일 복사 작업 정보	113
로컬 컴퓨터에서 SAS 서버로 파일 복사	113
21장 / 상관분석	117
상관분석 작업 정보	117
상관분석: 분석 역할에 변수 할당	118
상관분석: 상관분석 옵션 설정	118
상관분석: 결과 옵션 설정	120
상관분석: 출력 데이터 생성	121
제목 및 각주 지정	121
속성 보기	121
22장 / 출력형식 생성	123
출력형식 생성 작업 정보	123
출력형식 생성: 출력형식 옵션 설정	124
출력형식 생성: 새로운 출력형식 정의	125
속성 보기	127
23장 / 데이터셋에서 출력형식 생성	129
데이터셋에서 출력형식 생성 작업 정보	129
데이터셋에서 출력형식 생성	130
24장 / 지도용 테이블 생성	133
지도용 테이블 생성 작업 정보	133
지도용 테이블 생성: 분석 역할에 변수 할당	134
지도용 테이블 생성: 옵션 설정	135
지도용 테이블 생성: 결과 지정	136
제목 및 각주 지정	136
속성 보기	136
25장 / 시계열 데이터 생성 마법사	139
시계열 데이터 생성 마법사 정보	139
시계열 데이터 생성 마법사: 데이터 선택	139
시계열 데이터 생성 마법사: 역할에 변수 할당	140
시계열 데이터 생성 마법사: 누적, 결측값 및 영(0) 값에 대한 옵션 지정	140
시계열 데이터 생성 마법사: 추가 출력 선택	141
시계열 데이터 생성 마법사: 제목 및 각주 지정	141
26장 / 데이터셋 속성	143
데이터셋 속성 마법사 정보	143
데이터셋 속성: 데이터 선택	143
데이터셋 속성: 출력 지정	144
27장 / 데이터셋 및 출력형식 삭제	145
데이터셋 및 출력형식 삭제 작업 정보	145
데이터셋, 뷰 및 사용자 정의 출력형식 삭제	146
속성 보기	146

28장 / 판별분석	149
판별분석 작업 정보	149
판별분석: 분석 역할에 변수 할당	150
판별분석: 분석 옵션 설정	151
판별분석: 새로운 데이터 분류	152
판별분석: 결과 옵션 설정	154
제목 및 각주 지정	154
속성 보기	155
29장 / 분포분석	157
분포분석 작업 정보	157
분포분석: 분석 역할에 변수 할당	158
분포분석: 분포 선택	158
분포분석: 분포 사용자 정의	159
분포분석: 도표 생성	160
분포분석: 인셋 추가	161
분포분석: 결과 옵션 지정	161
제목 및 각주 지정	162
속성 보기	163
30장 / 도넛 그래프	165
도넛 그래프 작업 정보	165
도넛 그래프: 그래프 유형 선택	166
도넛 그래프: 분석 역할에 변수 할당	166
도넛 그래프: 모양 옵션 설정	167
제목 및 각주 지정	169
속성 보기	169
31장 / 데이터 파일을 PC로 다운로드	171
데이터 파일을 PC로 다운로드 작업 정보	171
데이터 파일을 PC로 다운로드: 다운로드할 SAS 데이터셋 선택	171
데이터 파일을 PC로 다운로드: 다운로드 옵션 지정	172
32장 / 요인분석	173
요인분석 작업 정보	173
요인분석: 분석 역할에 변수 할당	174
요인분석: 요인화 방법 옵션 설정	175
요인분석: 공통성 옵션 설정	176
요인분석: 회전 및 도표 옵션 설정	177
요인분석: 결과 옵션 설정	179
제목 및 각주 지정	180
33장 / Forecast Studio 프로젝트 생성	181
Forecast Studio 프로젝트 생성 마법사 정보	181
Forecast Studio 프로젝트 생성: 데이터 확인	182
Forecast Studio 프로젝트 생성: 예측변수 선택	182
Forecast Studio 프로젝트 생성: 분류변수 선택	183
Forecast Studio 프로젝트 생성: 프로젝트 저장	184
34장 / Forecast Studio 프로젝트 열기	185
Forecast Studio 프로젝트 열기 마법사 정보	185
Forecast Studio 프로젝트 열기: 예측 프로젝트 선택	186

Forecast Studio 프로젝트 열기: 예측 시계열 선택	186
Forecast Studio 프로젝트 열기: 결과 유형 선택	186
Forecast Studio 프로젝트 열기: 제목 및 각주 지정	187
35장 / Forecast Studio 프로젝트 Override	189
Forecast Studio 프로젝트 Override 마법사 정보	189
Forecast Studio 프로젝트 Override: 예측 프로젝트 선택	190
Forecast Studio 프로젝트 Override: 예측 계열 선택 및 Override 실행	190
36장 / 일반화선형모델	195
일반화선형모델 작업 정보	195
일반화선형모델: 분석 역할에 변수 할당	196
일반화선형모델: 모델 생성	197
일반화선형모델: 모델 옵션 설정	197
일반화선형모델: Post Hoc 검정 선택	199
일반화선형모델: 도표 생성	199
일반화선형모델: 예측값 옵션 설정	200
제목 및 각주 지정	201
속성 보기	201
37장 / High-Performance 선형회귀	203
High-Performance 선형회귀 작업 정보	203
High-Performance 선형회귀: 분석 역할에 변수 할당	204
High-Performance 선형회귀: 모델 생성	205
High-Performance 선형회귀: 모델 옵션 설정	208
High-Performance 선형회귀: 파티션 옵션 설정	208
High-Performance 회귀: 성능 옵션 설정	209
High-Performance 회귀: 예측값 옵션 설정	210
제목 및 각주 지정	210
38장 / High-Performance 로지스틱 회귀	213
High-Performance 로지스틱 회귀 작업 정보	213
High-Performance 로지스틱 회귀: 분석 역할에 변수 할당	214
High-Performance 로지스틱 회귀: 모델 생성	215
High-Performance 로지스틱 회귀: 모형 옵션 설정	218
High-Performance 회귀: 성능 옵션 설정	219
High-Performance 회귀: 예측값 옵션 설정	219
제목 및 각주 지정	220
39장 / 히스토그램	221
히스토그램 작업 정보	221
공정능력분석: 분석 역할에 변수 할당	222
히스토그램: 분포 선택	223
히스토그램: 분포 사용자 정의	224
공정능력분석: 도표 모양 옵션 설정	225
공정능력분석: 분석 유형 지정	228
제목 및 각주 지정	229
속성 보기	229
40장 / 개별 측정값관리도	231
개별 측정값관리도 정보	231
개별 측정값관리도: 분석 역할에 변수 할당	232

관리도: 관리한계 지정	233
관리도: 수행할 검정 선택	234
관리도: 도표 모양 옵션 설정	235
관리도: 분석 결과 저장	235
제목 및 각주 지정	236
속성 보기	236
41장 / 생명표 분석	239
생명표 분석 작업 정보	239
생명표 분석: 분석 역할에 변수 할당	240
생명표 분석: 방법 옵션 설정	241
생명표 분석: 도표 생성	242
생명표 분석: 결과 옵션 설정	243
제목 및 각주 지정	244
속성 보기	244
42장 / 선 도표	245
선 도표 작업 정보	245
선 도표: 도표 유형 선택	246
선 도표: 분석 역할에 변수 할당	247
선 도표: 모양 옵션 설정	248
제목 및 각주 지정	249
속성 보기	249
43장 / 선 도표 마법사	251
선 도표 마법사 정보	251
선 도표 마법사: 데이터 선택	251
선 도표 마법사: 역할에 변수 할당	252
선 도표 마법사: 도표 옵션 설정	252
선 도표 마법사: 모양 옵션 설정	253
선 도표 마법사: 제목 및 각주 지정	253
44장 / 선형모델	255
선형모델 작업 정보	255
선형모델: 분석 역할에 변수 할당	256
선형모델: 모델 생성	257
선형모델: 모델 옵션 설정	257
선형모델: 고급 옵션 설정	258
선형모델: Post Hoc 검정 선택	259
선형모델: 도표 생성	262
선형모델: 예측값 옵션 설정	262
제목 및 각주 지정	263
속성 보기	264
45장 / 선형회귀	265
선형회귀 작업 정보	265
선형회귀: 분석 역할에 변수 할당	266
선형회귀: 모델 옵션 설정	266
선형회귀: 통계량 옵션 설정	268
선형회귀: 도표 생성	269
선형회귀: 예측값 옵션 설정	271
제목 및 각주 지정	272

속성 보기	272
46장 / 데이터 리스트	273
데이터 리스트 작업 정보	273
데이터 리스트: 분석 역할에 변수 할당	274
데이터 리스트: 리스트 옵션 설정	274
제목 및 각주 지정	276
속성 보기	277
47장 / 리스트 리포트 마법사	279
리스트 리포트 마법사 정보	279
리스트 리포트 마법사: 데이터 선택	280
리스트 리포트 마법사: 리스트 정의	280
리스트 리포트 마법사: 합계 표시	280
리스트 리포트 마법사: 제목 및 각주 지정	281
48장 / 로지스틱 회귀	283
로지스틱 회귀 작업 정보	283
로지스틱 회귀: 분석 역할에 변수 할당	284
로지스틱 회귀: 모델 생성	285
로지스틱 회귀: 모델 옵션 설정	288
로지스틱 회귀: 도표 생성	289
로지스틱 회귀: 예측값 옵션 설정	290
제목 및 각주 지정	291
속성 보기	292
49장 / 지도 그래프	293
지도 그래프 작업 정보	293
지도 그래프: 지도 유형 선택	294
지도 그래프: 분석 역할에 변수 할당	294
지도 그래프: 모양 옵션 설정	295
제목 및 각주 지정	297
지도 그래프: 결과 미리 보기	297
속성 보기	298
50장 / 평균 및 범위관리도	299
평균 및 범위관리도 정보	299
평균 및 범위관리도: 분석 역할에 변수 할당	300
관리도: 관리한계 지정	301
관리도: 수행할 검정 선택	302
관리도: 도표 모양 옵션 설정	303
관리도: 분석 결과 저장	303
제목 및 각주 지정	304
속성 보기	304
51장 / 평균 및 표준편차관리도	307
평균 및 표준편차관리도 정보	307
평균 및 표준편차관리도: 분석 역할에 변수 할당	308
관리도: 관리한계 지정	309
관리도: 수행할 검정 선택	310
관리도: 도표 모양 옵션 설정	311
관리도: 분석 결과 저장	311

제목 및 각주 지정	312
속성 보기	312
52장 / 혼합모형	315
혼합모형 작업 정보	315
혼합모형: 분석 역할에 변수 할당	316
혼합모형: 고정 효과 모형 생성	317
혼합모형: 고정 효과 모형 옵션 설정	317
혼합모형: 임의효과 지정	318
혼합모형: 반복 효과 지정	320
혼합모형: 최소제곱평균 지정	321
혼합모형: 도표 생성	322
혼합모형: 예측값 옵션 설정	323
제목 및 각주 지정	324
속성 보기	324
53장 / 모델 스코어링	327
모델 스코어링 마법사 정보	327
모델 스코어링: 데이터 선택	328
모델 스코어링: 스코어링 모델 선택	328
모델 스코어링: 필수 모델 입력에 데이터 변수 매핑	328
모델 스코어링: 출력 데이터 칼럼 선택	329
모델 스코어링: 출력 데이터 저장	329
모델 스코어링: 선택 확인	329
모델 스코어링: 결과 해석	330
54장 / 비선형회귀	333
비선형회귀 작업 정보	333
비선형회귀: 분석 역할에 변수 할당	334
비선형회귀: 모델 옵션 설정	334
비선형회귀: 계산 옵션 설정	335
비선형회귀: 도표 생성	336
비선형회귀: 결과 옵션 설정	337
제목 및 각주 지정	338
속성 보기	338
55장 / 비모수 일원분산분석	339
비모수 일원분산분석 작업 정보	339
비모수 일원분산분석: 분석 역할에 변수 할당	340
비모수 일원분산분석: 분석 선택	340
비모수 일원분산분석: 정확 p-값 계산	341
비모수 일원분산분석: 결과 옵션 설정	341
제목 및 각주 지정	342
속성 보기	342
56장 / 불량개수관리도	345
불량개수관리도 작업 정보	345
불량개수관리도: 분석 역할에 변수 할당	346
관리도: 관리한계 지정	347
관리도: 수행할 검정 선택	348
관리도: 도표 모양 옵션 설정	349
관리도: 분석 결과 저장	349

제목 및 각주 지정	350
속성 보기	350
57장 / 일원분산분석	353
일원분산분석 작업 정보	353
일원분산분석: 분석 역할에 변수 할당	354
일원분산분석: 검정 선택	354
일원분산분석: 그룹 평균 비교	355
일원분산분석: 정량적 변수에 대한 통계량 선택	356
일원분산분석: 도표 생성	356
일원분산분석: 결과 옵션 설정	357
제목 및 각주 지정	358
속성 보기	358
58장 / 일원빈도분석	361
일원빈도분석 작업 정보	361
일원빈도분석: 분석 역할에 변수 할당	362
일원빈도분석: 통계량 옵션 설정	362
일원빈도분석: 도표 생성	363
일원빈도분석: 결과 옵션 설정	363
제목 및 각주 지정	364
속성 보기	365
59장 / 불량률관리도	367
불량률관리도 정보	367
불량률관리도: 분석 역할에 변수 할당	368
관리도: 관리한계 지정	369
관리도: 수행할 검정 선택	370
관리도: 도표 모양 옵션 설정	371
관리도: 분석 결과 저장	371
제목 및 각주 지정	372
속성 보기	372
60장 / 파레토 그래프	375
파레토 그래프 작업 정보	375
파레토 그래프: 분석 역할에 변수 할당	376
파레토 그래프: 범주 개수 제한	377
파레토 그래프: 도표 모양 옵션 설정	378
제목 및 각주 지정	380
속성 보기	381
61장 / 원 그래프	383
원 그래프 작업 정보	383
원 그래프: 그래프 유형 선택	384
원 그래프: 분석 역할에 변수 할당	384
원 그래프: 모양 옵션 설정	385
제목 및 각주 지정	387
속성 보기	387
62장 / 원 그래프 마법사	389
원 그래프 마법사 정보	389
원 그래프 마법사: 데이터 선택	389

원 그래프 마법사: 역할에 변수 할당	390
원 그래프 마법사: 모양 옵션 설정	390
원 그래프 마법사: 제목 및 각주 지정	391
63장 / P-P 도표	393
P-P 도표 작업 정보	393
공정능력분석: 분석 역할에 변수 할당	394
P-P 도표: 분포 선택	395
P-P 도표: 분포 사용자 정의	396
공정능력분석: 도표 모양 옵션 설정	396
공정능력분석: 분석 유형 지정	399
제목 및 각주 지정	400
속성 보기	400
64장 / 시계열 데이터 준비	403
시계열 데이터 준비 작업 정보	403
시계열 데이터 준비: 분석 역할에 변수 할당	404
시계열 데이터 준비: 출력 빈도 지정	405
시계열 데이터 준비: 시계열 특성 지정	405
시계열 데이터 준비: 각 시계열 변수에 대한 변환 지정	406
시계열 데이터 준비: 결과 옵션 설정	407
속성 보기	408
데이터셋의 결측값에 대한 관측값 행 생성	408
표본 추출 빈도 간의 시계열 데이터 변환	409
입력 데이터의 관측값 특성 변경	410
비주기적 계열을 주기적 추정값으로 변환	410
결측값 보간	411
데이터에 일반 변환 적용	412
65장 / 주성분분석	413
주성분분석 작업 정보	413
주성분분석: 분석 역할에 변수 할당	414
주성분분석: 분석 옵션 설정	414
주성분분석: 도표 생성	415
주성분분석: 결과 옵션 설정	416
제목 및 각주 지정	417
속성 보기	417
66장 / 확률 도표	419
확률 도표 작업 정보	419
공정능력분석: 분석 역할에 변수 할당	420
확률 도표: 분포 선택	421
확률 도표: 분포 사용자 정의	422
공정능력분석: 도표 모양 옵션 설정	422
공정능력분석: 분석 유형 지정	425
제목 및 각주 지정	426
속성 보기	426
67장 / 비례위험분석	429
비례위험분석 작업 정보	429
비례위험분석: 분석 역할에 변수 할당	430
비례위험분석: 모델 옵션 설정	431

비례위험분석: 방법 옵션 설정	433
비례위험분석: 도표 생성	434
비례위험분석: 결과 옵션 설정	434
제목 및 각주 지정	435
속성 보기	435
68장 / Q-Q 도표	437
Q-Q 도표 작업 정보	437
공정능력분석: 분석 역할에 변수 할당	438
Q-Q 도표: 분포 선택	439
Q-Q 도표: 분포 사용자 정의	440
공정능력분석: 도표 모양 옵션 설정	440
공정능력분석: 분석 유형 지정	443
제목 및 각주 지정	444
속성 보기	444
69장 / 레이더 그래프	447
레이더 그래프 작업 정보	447
레이더 그래프: 분석 역할에 변수 할당	448
레이더 그래프: 모양 옵션 설정	449
제목 및 각주 지정	449
속성 보기	450
70장 / 확률표본	451
확률표본 작업 정보	451
확률표본: 확률표본 생성	451
71장 / 데이터 순위화	455
데이터 순위화 작업 정보	455
데이터 순위화: 분석 역할에 변수 할당	456
데이터 순위화: 분석 옵션 설정	456
순위화: 결과 옵션 설정	457
속성 보기	458
72장 / RFM(Recency, Frequency, Monetary) 분석	459
RFM(Recency, Frequency, Monetary) 분석 작업 정보	459
RFM(Recency, Frequency, Monetary) 분석: 분석할 데이터 유형 선택	460
RFM(Recency, Frequency, Monetary) 분석: 분석 역할에 변수 할당	461
RFM(Recency, Frequency, Monetary) 분석: 데이터를 구간으로 범주화	462
RFM(Recency, Frequency, Monetary) 분석: 결과 저장	463
RFM(Recency, Frequency, Monetary) 분석: 도표 선택	463
73장 / 자기회귀 오차가 있는 회귀분석	465
자기회귀 오차가 있는 회귀분석 작업 정보	465
자기회귀 오차가 있는 회귀분석: 분석 역할에 변수 할당	466
자기회귀 오차가 있는 회귀분석: 모델 옵션 설정	467
자기회귀 오차가 있는 회귀분석: 통계량 옵션 설정	468
자기회귀 오차가 있는 회귀분석: 도표 생성	470
자기회귀 오차가 있는 회귀분석: 결과 옵션 설정	471
제목 및 각주 지정	472
속성 보기	472

74장 / 패널 데이터의 회귀분석	473
패널 데이터의 회귀분석 작업 정보	473
패널 데이터의 회귀분석: 분석 역할에 변수 할당	474
패널 데이터의 회귀분석: 모델 옵션 설정	474
패널 데이터의 회귀분석: 결과 옵션 설정	475
제목 및 각주 지정	476
속성 보기	476
75장 / SAS ODS 통계적 그래프	479
SAS ODS 통계적 그래프 정보	479
ODS Graphics Designer 사용	479
ODS 통계적 그래프 표시	480
76장 / SAS Rapid Predictive Modeler	481
SAS Rapid Predictive Modeler 정보	481
SAS Rapid Predictive Modeler를 위한 표본추출 전략	482
SAS Rapid Predictive Modeler: 데이터 정리	483
SAS Rapid Predictive Modeler: 분석 역할에 변수 할당	484
SAS Rapid Predictive Modeler: 모델 선택	485
SAS Rapid Predictive Modeler: 리포트 옵션 선택	489
SAS Rapid Predictive Modeler: 출력 옵션 설정	490
SAS Rapid Predictive Modeler: 모델 등록	491
77장 / 산점도	493
산점도 작업 정보	493
산점도: 도표 유형 선택	494
산점도: 분석 역할에 변수 할당	494
산점도: 모양 옵션 설정	495
제목 및 각주 지정	497
속성 보기	497
78장 / 산점도 행렬	499
산점도 행렬 작업 정보	499
산점도 행렬: 산점도 행렬 생성	499
79장 / 데이터 정렬	503
데이터 정렬 작업 정보	503
데이터 정렬: 분석 역할에 변수 할당	504
데이터 정렬: 데이터 정렬 옵션 설정	504
데이터 정렬: 결과 옵션 설정	506
속성 보기	506
80장 / 칼럼 분할	507
칼럼 분할 작업 정보	507
칼럼 분할: 분석 역할에 변수 할당	509
칼럼 분할: 결과 옵션 설정	510
속성 보기	510
81장 / 칼럼 세로 배열	511
칼럼 세로 배열 작업 정보	511
칼럼 세로 배열: 분석 역할에 변수 할당	513
칼럼 세로 배열: 칼럼 명명 옵션 설정	513

칼럼 세로 배열: 결과 옵션 설정	514
속성 보기	515
82장 / 데이터 표준화	517
데이터 표준화 작업 정보	517
데이터 표준화: 분석 역할에 변수 할당	518
데이터 표준화: 표준화 옵션 설정	518
데이터 표준화: 결과 옵션 설정	519
속성 보기	519
83장 / 요약통계량	521
요약통계량 작업 정보	521
요약통계량: 분석 역할에 변수 할당	522
요약통계량: 요약통계량 선택	523
요약통계량: 도표 옵션 설정	524
요약통계량: 결과 옵션 설정	525
제목 및 각주 지정	527
속성 보기	527
84장 / 요약통계량 마법사	529
요약통계량 마법사 정보	529
요약통계량 마법사: 데이터 선택	529
요약통계량 마법사: 역할에 변수 할당	529
요약통계량 마법사: 리포트 통계량 및 출력 결과 지정	530
요약통계량 마법사: 제목 및 각주 지정	531
85장 / 요약테이블	533
요약테이블 작업 정보	533
요약테이블: 분석 역할에 변수 할당	534
요약테이블: 요약 테이블 설정	534
요약테이블: 결과 옵션 설정	537
제목 및 각주 지정	537
속성 보기	537
86장 / 요약테이블 마법사	539
요약테이블 마법사 정보	539
요약테이블 마법사: 데이터 선택	539
요약테이블 마법사: 역할에 변수 할당	540
요약테이블 마법사: 분류변수 선택	540
요약테이블 마법사: 합계 지정	541
요약테이블 마법사: 추가 출력 선택	541
요약테이블 마법사: 제목 및 각주 지정	541
87장 / 표면도	543
표면도 작업 정보	543
표면도: 도표 유형 선택	544
표면도: 분석 역할에 변수 할당	544
표면도: 모양 옵션 설정	545
제목 및 각주 지정	546
속성 보기	547

88장 / t 검정	549
t 검정 작업 정보	549
t 검정: t 검정 유형 선택	550
t 검정: 분석 역할에 변수 할당	550
t 검정: 분석 옵션 설정	551
t 검정: 도표 생성	552
제목 및 각주 지정	552
속성 보기	553
89장 / 테이블 분석	555
테이블 분석 작업 정보	555
테이블 분석: 분석 역할에 변수 할당	556
테이블 분석: 테이블 정의	556
테이블 분석: 셀 통계량 선택	557
테이블 분석: 테이블 통계량 지정	558
테이블 분석: 결과 옵션 설정	562
제목 및 각주 지정	563
속성 보기	563
90장 / 타일 그래프	565
타일 그래프 작업 정보	565
타일 그래프: 분석 역할에 변수 할당	566
타일 그래프: 모양 옵션 설정	567
타일 그래프: 고급 옵션 설정	568
제목 및 각주 지정	570
91장 / 전치	571
전치 작업 정보	571
전치: 분석 역할에 변수 할당	572
전치: 분석 옵션 설정	573
전치: 결과 옵션 설정	573
속성 보기	574
92장 / 단위당 결점수관리도	575
단위당 결점수관리도 정보	575
단위당 결점수관리도: 분석 역할에 변수 할당	576
관리도: 관리한계 지정	577
관리도: 수행할 검정 선택	578
관리도: 도표 모양 옵션 설정	579
관리도: 분석 결과 저장	579
제목 및 각주 지정	580
속성 보기	581
93장 / 라이브러리 메타데이터 업데이트	583
라이브러리 메타데이터 업데이트 마법사 정보	583
라이브러리 메타데이터 업데이트: SAS 서버 및 라이브러리 선택	584
라이브러리 메타데이터 업데이트: 메타데이터에서 리포트 실행 또는 메 타데이터 업데이트	584
94장 / CAS로 업로드	587
CAS로 업로드 작업 정보	587
SAS CAS Analytic Server로 데이터 업로드	588

95장 / 데이터 파일을 서버에 업로드	589
데이터 파일을 서버로 업로드 작업 정보	589
데이터 파일을 서버에 업로드: 업로드할 SAS 데이터셋 선택	589
데이터 파일을 서버에 업로드: 업로드 옵션 지정	590
96장 / LASR로 업로드	591
LASR로 업로드 작업 정보	591
SAS LASR Analytic Server로 데이터 업로드	592
부록 1 / 그래프 작업에 대한 공통 모양 옵션	593
축 옵션 설정	593
참조선 추가	595
범례 옵션 설정	596
그래프 영역 사용자 정의	597
막대 모양 지정	598
막대 옵션 설정	599
선 도표 옵션 설정	600
보간법 선택	601
고급 옵션 설정	602
부록 2 / 관리도 작업에 대한 공통 모양 옵션	605
축 옵션 설정	605
참조선 추가	605
인셋 추가	606
기타 도표 옵션 설정	606
블록 변수 옵션 설정	608
부록 3 / 모델 생성	609
모델 정보	609
주효과 지정	610
교차효과(교호작용) 지정	610
지분 효과 지정	610
두 개 이상의 선택 변수에 대한 다중 교차 생성	610
양적변수의 다항 교차 생성	611
분석에서 절편 효과 제거	611
부록 4 / 빌트인 작업	613
빌트인 작업, 관련 SAS 프로시저 및 라이선스	613

작업 사용

빌트인 작업 및 SAS Studio 작업	1
참고 항목	2
작업 데이터 필터링	3
작업에 사용자 코드 추가	4
칼럼 속성 지정	4
그래픽 출력형식 선택	5
작업 템플릿 작업	5
작업 템플릿 정보	5
작업 템플릿 생성	6
작업 템플릿 편집	6
작업 템플릿 실행	6
작업 템플릿 가져오기	7
작업 템플릿 내보내기	7

빌트인 작업 및 SAS Studio 작업

SAS Enterprise Guide 및 **SAS Add-In for Microsoft Office**는 많은 빌트인 SAS 작업과 함께 제공됩니다. **작업** 영역에서 이러한 작업에 액세스할 수 있습니다. 사용자 정의 작업 및 SAS Studio 작업도 이 리스트에 포함될 수 있습니다.

SAS Enterprise Guide 또는 **SAS Add-In for Microsoft Office**의 SAS Studio 작업에 액세스하는 방법:

- 1 사이트에 SAS Studio가 설치되어 있는지 확인하십시오.
- 2 활성화된 프로파일에 대한 기본 서버가 SAS Studio 서버인지 확인하십시오.
- 3 **SAS Enterprise Guide** 또는 **SAS Add-In for Microsoft Office**에 대한 설정에서 **SAS Studio 버전의 SAS Studio 작업 표시** 옵션을 선택합니다.
 - SAS Studio 3.4 이상을 실행 중이면 SAS Studio 기업 버전 또는 SAS Studio 단일 사용자 버전에서 작업에 액세스할 수 있습니다. SAS Studio Single User Edition

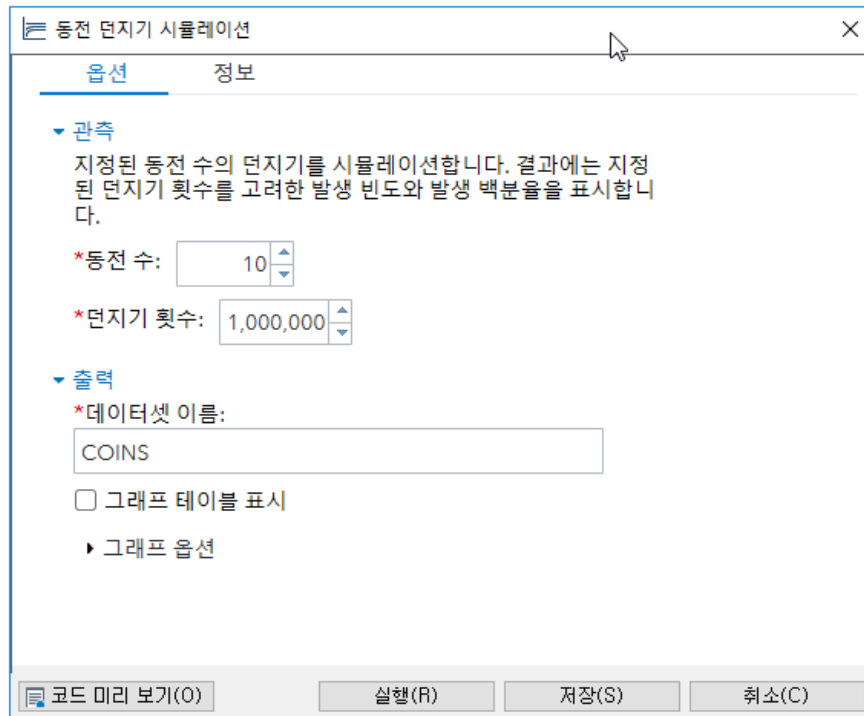
은 **SAS Enterprise Guide** 및 **SAS Add-In for Microsoft Office**와 같은 컴퓨터에 설치되어 있어야 합니다.

- SAS Enterprise Guide에서 SAS Studio 5.2 작업에 액세스할 수도 있습니다. SAS Studio 릴리스에 대해 URL 및 자격 증명을 지정합니다.

주: SAS Studio 5.2 작업은 SAS Add-In for Microsoft Office에서 사용할 수 없습니다.

이제 모든 SAS Studio 작업을 **SAS Enterprise Guide** 및 **SAS Add-In for Microsoft Office**의 **작업** 영역 및 **작업** 창에서 사용할 수 있습니다. **작업** 영역 및 작업 갤러리에서 **SAS Studio** 작업만 보려면 작업 유형으로 **SAS Studio**를 선택합니다.

SAS Studio 작업의 인터페이스가 **SAS Enterprise Guide** 및 **SAS Add-In for Microsoft Office**의 빌트인 작업에 사용되는 대화 상자와 약간 다릅니다. 예를 들어 이 인터페이스에는 동전 던지기 시뮬레이션 작업이 표시됩니다.



특정 SAS Studio 작업 사용은 사이트의 라이선스에 따라 달라질 수 있습니다. 예를 들어 많은 SAS Studio 작업은 SAS Viya 환경으로 작업할 수 있습니다. 그러나, SAS Viya 라이선스가 없거나 설치되어 있지 않으면 이러한 작업은 사용할 수 없습니다.

참고 항목

- 부록 4, “빌트인 작업” (613페이지)
- *SAS Studio: Task Reference Guide*의 “SAS Studio Tasks and Associated Procedures and Licenses”

작업 데이터 필터링

데이터 및 필터 편집 대화 상자의 드롭다운 리스트를 사용하여 현재 데이터 소스에 적용할 필터를 지정할 수 있습니다. 필터를 생성한 후에는 작업의 제목 및 각주에 `_CLIENTTASKFILTER` 매크로 변수를 사용하여 필터 정보가 결과의 제목 또는 각주에 표시되도록 할 수 있습니다.

주: SAS Add-In for Microsoft Office를 사용 중이면 **데이터 소스 수정** 창에서 정의한 필터가 먼저 사용됩니다. 그런 다음 **데이터 및 필터 편집** 대화 상자를 사용하여 생성한 필터가 사용됩니다.

데이터에 대한 필터를 생성하는 방법

- 1 작업의 데이터 패널이나 마법사의 데이터 확인 단계에서 **편집**을 클릭합니다. **데이터 및 필터 편집** 대화 상자가 나타납니다.
- 2 (선택 사항) SAS Enterprise Guide에서 입력 데이터 소스를 선택합니다. 작업을 열면 기본적으로 활성 데이터 소스가 작업의 입력 데이터 소스로 사용됩니다. 활성 데이터 소스의 이름은 데이터 소스 드롭다운 리스트에 표시됩니다.
- 3 첫 번째 드롭다운 리스트에서 필터링할 변수를 선택합니다. 드롭다운 리스트에는 데이터의 모든 변수가 있습니다.

주: 기본적으로 이 드롭다운 리스트에는 변수 이름이 표시됩니다. 이 드롭다운 리스트에서 변수 레이블을 표시하려면 **변수 이름 대신 레이블 표시** 체크 박스를 선택합니다.

- 4 두 번째 드롭다운 리스트에서 연산자를 선택합니다.
- 5 세 번째 상자에서 다음 방법 중 하나로 변수값을 지정합니다.

- 텍스트 상자에 변수값을 입력합니다.
- 값 선택기 버튼을 클릭하여 변수값으로 선택할 수 있는 값 리스트를 표시합니다.

Between 또는 **Not between** 연산자를 선택했으면, 시작 및 끝 값을 지정해야 합니다. **In a list** 또는 **Not in a list** 연산자를 선택했으면, 비교할 값 리스트를 지정해야 합니다. 상자에서 값을 입력하거나 붙여 넣을 수 있습니다. **값 추가**를 클릭하여 데이터 소스에서 값을 선택합니다. 각 값은 새로운 행에 있어야 합니다.

다른 모든 연산자는 데이터 소스에 있는 처음 25개의 개별 값이 리스트에 표시됩니다. 더 많은 값을 보고 싶으면 **다른 값**을 클릭하십시오. **다른 값**을 클릭할 때마다 처음 25개의 개별 값 외에 변수에 대한 다음 25개 값이 표시됩니다.

주: 반면, SAS 데이터 소스가 그 외의 위치에 있으면 오름차순으로 값 리스트가 채워집니다. 데이터 소스가 컴퓨터의 파일 시스템에 없거나 컴퓨터에서 액세스할 수 있는 서버에 없으면 값 리스트가 오름차순으로 채워집니다.

- 6 (선택 사항) 데이터에 다른 필터를 추가하려면 마지막 드롭다운 리스트에서 **AND** 또는 **OR**을 선택합니다. 3-5단계를 반복하여 이러한 필터를 생성합니다.

주: 두 개 이상의 필터를 지정하면 AND 관계로 연결한 필터가 먼저 평가됩니다. 필터는 생성된 순서대로 평가되지 않습니다. 예를 들어, **Sex = 'M' or Name = 'Carol' and Date = '1980'** 필터를 생성했다고 가정합니다.

이 예에서는 먼저 Name이 Carol이고 Date가 1980인 모든 관측값이 필터링됩니다. 그런 다음 OR 관계가 평가됩니다.

- 7 **확인**을 클릭하여 변경 사항을 저장합니다. 이제 입력 데이터 소스에 대한 필터가 작업의 데이터 패널에 표시됩니다.

작업에 사용자 코드 추가

작업 코드를 미리 볼 때 작업을 위해 SAS가 생성한 코드에 자신만의 코드를 삽입하는 옵션을 사용할 수 있습니다. SAS 코드 프로시저 옵션 또는 문장을 프로그램 편집기의 모든 기능과 함께 생성된 코드에 삽입할 수 있습니다. **작업에 대한 코드 미리 보기** 창에서는 자동으로 생성된 코드를 편집하거나 삭제할 수 없으며 새로운 코드를 삽입 있습니다. 작업 창에 자신만의 코드를 삽입하려면 **코드 미리 보기**를 클릭합니다.

- 1 **작업에 대한 코드 미리 보기** 창에서 **사용자 정의 코드 삽입점 표시** 체크 박스를 선택합니다. 다음 행은 사용자 정의 코드를 삽입할 수 있는 각 위치에서 작업 코드를 추가합니다.

```
/* Start of custom user code (Framework_BeforeTaskCode) */
<insert custom code here>
/* End of custom user code (Framework_BeforeTaskCode) */
```

- 2 코드를 삽입할 위치를 클릭하고 코드를 입력합니다.
- 3 변경 사항을 저장하려면 **작업에 대한 코드 미리 보기** 창을 닫습니다.

칼럼 속성 지정

이 옵션에 액세스하려면 작업 창의 선택 영역에서 **데이터**를 클릭합니다.

변수를 마우스 오른쪽 버튼으로 클릭하고 **속성**을 선택하면 설명 레이블 및 출력형식을 임시로 지정할 수 있습니다. 지정하는 모든 속성은 현재 작업에만 적용됩니다.

레이블을 지정하려면 **레이블** 상자에 레이블 텍스트를 입력합니다. 레이블의 문자 수는 256자를 초과할 수 없습니다. 드롭다운 리스트에는 해당 변수와 영구적으로 연결된 레이블 값이 들어 있습니다.

대부분의 작업에서 출력형식을 지정할 수 있습니다. 출력형식을 지정하는 방법:

- 1 **변경**을 클릭합니다. **출력형식** 창이 나타납니다.
- 2 범주 상자에서 출력형식의 **범주**를 선택합니다. 이전에 출력형식 생성 작업에서 정의한 표준 SAS 출력형식 또는 사용자 작성 출력형식을 사용할 수 있습니다.
- 3 **출력형식** 상자에서 출력형식 이름을 선택합니다.
- 4 **속성** 상자에서 속성 변경 사항을 지정한 다음 **확인**을 클릭합니다.

주: 속성 창의 **정렬 여부** 속성이 예로 설정되어 있으면 데이터셋이 해당 속성별로 정렬됩니다.

그래픽 출력형식 선택

SAS Add-In for Microsoft Office에서는 다음 그래픽 출력 형식 중에서 선택할 수 있습니다: ActiveX, ActiveX 이미지, GIF, JPEG, PNG, Java 이미지. **SAS Enterprise Guide**에서는 Java 및 SAS EMF 출력형식도 선택할 수 있습니다.

ActiveX 및 Java 출력형식은 대화식입니다. 이 결과 출력형식으로 생성된 그래픽은 마우스 오른쪽 버튼으로 클릭하여 특정 옵션을 변경할 수 있습니다. 변경할 수 있는 옵션은 Active X 및 Java 출력형식이 서로 다릅니다. 변경한 옵션은 출력에만 반영되고 작업 창에서 그래픽 생성을 위해 선택한 사항에는 반영되지 않습니다.

ActiveX 이미지, GIF, JPEG, PNG, Java 이미지 및 SAS EMF 출력형식은 비대화식입니다. 그러므로 일단 생성한 그래프는 모양을 변경할 수 없습니다.

작업 템플릿 작업

작업 템플릿 정보

막대 그래프 작업과 같은 작업을 실행할 때 동일한 설정을 자주 사용할 수 있습니다. 다른 입력 데이터 소스에 대한 설정을 지정하기 위해 작업을 여러 번 열어보는 대신, 작업 템플릿은 템플릿의 특정 작업을 위한 설정을 저장할 수 있습니다. 그런 다음 입력 데이터 소스로 템플릿을 실행할 수 있습니다.

주: 작업 템플릿을 작업할 때 다음과 같은 제한 사항에 유의하십시오.

- 데이터 소스에 따라 달라지는 값 또는 설정은 작업 템플릿에 저장되지 않습니다. 예를 들어 특정 역할에 변수를 할당하면 변수 할당은 작업 템플릿에 저장되지 않습니다. 데이터 소스의 값에 종속된 옵션이면 옵션 설정은 작업 템플릿에 저장되지 않습니다. 예를 들어 선 도표 작업에서 데이터 값에 따르는 기호 정의는 작업 템플릿에 저장되지 않습니다.

- 다음 작업에서는 작업 템플릿을 생성할 수 없습니다: 테이블 추가, 데이터 비교, 데이터 셋 및 출력형식 삭제, PC에 데이터 파일 다운로드, Forecast Studio 프로젝트 생성, Forecast Studio 프로젝트 열기, Forecast Studio 프로젝트 Override, 리포트 리스트, 모델 스코어링, 및 요약테이블.
- **SAS Enterprise Guide**에서 프롬프트를 사용하는 작업의 템플릿을 생성하면 프롬프트는 작업 템플릿의 일부로 저장되지 않습니다.

주: **SAS Enterprise Guide**에서도 질의 빌더에서 템플릿을 생성할 수 있습니다.

SAS Enterprise Guide에서 작업 템플릿 관리자를 사용하여 템플릿을 관리할 수 있습니다.

작업 템플릿 생성

- 1 작업을 열고 작업에 대한 옵션을 지정합니다.
- 2 **실행** 옆에 있는 아래쪽 화살표를 클릭하고 팝업 메뉴에서 **템플릿 생성**을 선택합니다. **작업 템플릿 생성** 대화 상자가 나타납니다.
- 3 템플릿의 이름 및 설명을 지정합니다.
- 4 그룹에 새로운 작업 템플릿을 생성할지 기존 템플릿을 바꿀지를 지정합니다.
- 5 **생성**을 클릭합니다.

작업 템플릿 편집

생성한 템플릿에 변경 사항을 적용할 수 있습니다. 질의에 기반한 템플릿에 변경 사항을 적용하려면 먼저 데이터 참조가 여전히 유효한지 확인해야 합니다.

- 1 편집하려는 작업 템플릿을 엽니다. 작업 창은 템플릿에 지정된 선택 사항으로 엽니다.
- 2 작업에서 저장하려는 옵션을 지정하고 **실행** ⇒ **템플릿 생성**을 선택합니다. **작업 템플릿 생성** 대화 상자가 나타납니다.
- 3 **기본 작업 템플릿 바꾸기**를 선택하고 바꿀 기존 템플릿을 선택합니다. **생성**을 클릭합니다. 기존 작업 템플릿에 변경 사항이 저장됩니다.

작업 템플릿 실행

- 1 **작업** ⇒ **작업 템플릿** ⇒ **그룹 이름** ⇒ **템플릿 이름**을 선택하여 작업 템플릿을 엽니다.

- 2 (선택 사항) 입력 데이터 소스를 선택합니다. 기본적으로 프로젝트의 활성 데이터 소스가 작업의 입력 데이터 소스입니다. 그러나 대부분의 작업에서는 **데이터** 패널에서 입력 데이터 소스를 변경할 수 있습니다.
- 3 작업 창의 **데이터** 패널에서 작업 역할에서 입력 데이터 소스에 대한 변수를 할당합니다.
- 4 (선택 사항) 작업에서 추가 옵션을 수정합니다.
- 5 **실행**을 클릭합니다.

작업 템플릿 가져오기

- 1 **작업** 영역에서 **템플릿 가져오기**를 클릭합니다.
- 2 가져오려는 템플릿을 하나 이상 선택하려면, **추가**를 클릭합니다.
- 3 **그룹 가져오기** 드롭다운 리스트에서 가져온 템플릿을 배치할 그룹을 선택합니다.
- 4 **가져오기**를 클릭합니다. **작업** 영역에 템플릿이 나타나고 **작업** 메뉴에서 사용할 수 있습니다.

작업 템플릿 내보내기

- 1 **작업** 영역에서 **템플릿 내보내기**를 클릭합니다.
- 2 내보낼 템플릿을 선택하고 **선택한 작업 템플릿** 상자에 추가합니다.
- 3 이러한 템플릿에 대한 위치를 선택하려면 **내보내기**를 클릭합니다.
- 4 **내보내기**를 클릭합니다. 선택한 위치에 템플릿의 복사본이 저장됩니다.

테이블 추가 작업

테이블 추가 작업 정보	9
테이블 추가: 입력 테이블 리스트 작성	10
테이블 추가: 결과 저장	10

테이블 추가 작업 정보

테이블 추가 작업은 두 개 이상의 입력 테이블을 연결하여 하나의 출력 데이터셋 또는 SQL 뷰를 생성합니다. 최대 256개의 테이블을 연결할 수 있습니다. 추가된 데이터셋에는 입력 테이블의 모든 행과 변수가 포함됩니다. 모든 입력 테이블이 아닌 일부 입력 테이블에만 포함된 변수는 해당 변수가 없는 입력 테이블 결과의 각 행에서 결측값을 갖게 됩니다.

예를 들어 2002년 각 분기의 제품 판매량을 표시하는 네 개의 입력 데이터셋이 있을 수 있습니다. 2002년 전체 판매량을 표시하는 하나의 출력 데이터셋을 생성하려면 테이블 추가 작업을 사용하면 됩니다.

주: 테이블 추가 작업은 SAS Enterprise Guide에서만 사용할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	SQL
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	없음

테이블 추가: 입력 테이블 리스트 작성

연결할 테이블을 선택하고 배열하려면 선택 영역에서 **테이블**을 클릭합니다.

추가할 테이블 영역에는 **연결할 테이블**이 연결 순서대로 나열됩니다. 이 영역에서 원하는 테이블 리스트를 작성할 수 있습니다.

- 1 **테이블 추가**를 클릭합니다. **위치 선택** 대화 상자가 열립니다. 대화 상자가 열리면 추가할 테이블의 이름을 선택합니다.
- 2 테이블을 두 개 이상 선택하려면 Ctrl 키를 누른 상태로 추가할 테이블을 선택합니다.
- 3 **열기** 또는 **확인**을 클릭하여 **추가할 테이블** 영역에 테이블을 추가합니다.

주: 테이블 추가 작업을 여는 데 사용한 테이블은 삭제할 수 없습니다.

테이블 추가: 결과 저장

해당 데이터에서 SAS 데이터셋 또는 SQL 뷰를 생성하려면 선택 영역에서 **결과**를 클릭합니다.

추가된 데이터 결과를 SAS 데이터셋으로 저장할지 아니면 SQL 뷰로 저장할지를 **결과 집합 출력형식** 영역에서 선택할 수 있습니다.

- 결과를 SAS 데이터셋으로 저장하려면 **데이터 테이블**을 클릭합니다. 기본값입니다.
- 결과를 SQL 뷰로 저장하려면 **데이터 뷰**를 클릭합니다.

주: SAS Enterprise Guide는 **옵션 창의 작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

영역 도표

영역 도표 작업 정보	11
영역 도표: 도표 유형 선택	12
영역 도표: 분석 역할에 변수 할당	12
영역 도표: 모양 옵션 설정	13
영역 사용자 정의	13
추가 모양 옵션	14
제목 및 각주 지정	14
속성 보기	14

영역 도표 작업 정보

영역 도표 작업을 통해 두 변수 사이의 수학적 관계를 보여 주는 영역, 스플라인, 스텝 또는 오버레이 도표를 생성할 수 있습니다. 곡선 아래의 영역은 시각적 효과를 위해 패턴 또는 색상 조합으로 채워집니다.

자세한 내용은 [“그래픽 출력형식 선택” \(5페이지\)](#)에서 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GPLOT
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

영역 도표: 도표 유형 선택

이 옵션에 액세스하려면 선택 영역에서 **영역 도표**를 클릭합니다.

역할 이름	설명
영역 도표	두 개의 변수값을 사용하여 선 도표를 생성합니다. 도표의 선 아래 영역은 음영 처리됩니다.
스플라인 영역 도표	스플라인 루틴을 사용하여 선 도표를 생성합니다. 스플라인 도표의 선 아래 영역은 음영 처리됩니다.
스텝 영역 도표	스텝함수를 사용하여 선 도표를 생성합니다. 스텝 도표의 선 아래 영역은 음영 처리됩니다.
중첩을 사용한 다중 수직 칼럼 영역 도표	각각의 Y 변수 대 동일한 X 변수로 이루어진 개별 도표를 생성합니다. 중첩 옵션으로 인해 모든 도표가 같은 그래프상에 표시됩니다. 각 도표 아래 영역은 서로 다른 색상으로 음영 처리됩니다.
그룹 칼럼별 다중 영역 도표	그룹 역할에 할당한 각 변수의 고유티값에 대해 별도의 도표를 생성합니다. 모든 도표가 동일한 그래프에 나타납니다. 각 도표 아래 영역은 서로 다른 색상으로 음영 처리됩니다.

영역 도표: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
수평	이 역할에 할당하는 칼럼은 그래프의 가로축 또는 X축 변수입니다.

역할 이름	설명
수직	이 역할에 할당하는 칼럼은 그래프의 세로축 또는 Y축 변수입니다.
세로(오른쪽)	이 역할에 할당하는 칼럼은 그래프의 오른쪽 면에 대한 세로축 또는 Y축 변수입니다. 이 변수는 X축 변수에 대응하여 그려지므로 그래프에 별도의 도표가 생성됩니다.
그룹	이 역할에 할당하는 칼럼 값에 따라 동일한 그래프에 표시되는 도표 선 개수가 결정됩니다.
그룹 분석 칼럼	이 역할에 할당하는 칼럼의 값에 따라 생성되는 그래프의 수가 결정됩니다. 각 고유티값에 대해 별도의 그래프가 생성됩니다. 생성된 모든 그래프에 대해 각 축에 대한 배율을 동일하게 하려면 모든 그래프에 대해 축 길이 같음 체크 박스를 선택합니다. 기본적으로, 각 축에 대한 값의 범위는 도표화된 데이터에서 최소 및 최대값에 기반합니다. 이 옵션을 선택하면 각 축의 범위는 모든 그래프에 대해 동일합니다.

영역 도표: 모양 옵션 설정

영역 사용자 정의

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **영역**을 클릭합니다.

다음에서 영역 채우기 유형을 선택합니다.

- **없음**
- **단색**
- **선**

선을 선택하면 채우기 유형 및 선 밀도의 패턴을 선택합니다.

주: 모든 그래프 출력형식에서 영역 채우기 유형을 모두 지원하지는 않습니다. ActiveX, Java, ActiveX 이미지 또는 Java 이미지 출력형식을 사용하는 경우 영역 채우기 스타일로 **비어 있음** 또는 **선**을 선택해도 영향을 미치지 않습니다. 그래프 출력형식에는 영역 채우기 유형이 단색일 때처럼 영역 도표가 표시됩니다.

추가 모양 옵션

일부 모양 옵션은 여러 그래프 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- “축 옵션 설정” (593페이지)
- “참조선 추가” (595페이지)
- “범례 옵션 설정” (596페이지)
- “그래프 영역 사용자 정의” (597페이지)

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

.....
주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.
.....

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간

- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 속성 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

ARIMA 모델화 및 예측 작업

ARIMA 모델 개발 및 예측 작업 정보	17
ARIMA 모형화 및 예측: 분석 역할에 변수 할당	18
ARIMA 모형화 및 예측: 식별 옵션 설정	19
반응시계열의 차분	19
정상성 검정 지정	19
도표 생성 및 결과 저장	19
ARIMA 모형화 및 예측: 추정 옵션 설정	20
추정 단계 사용	20
ARIMA 모델 정의	20
추정 방법 및 옵션 선택	21
결과 저장	21
ARIMA 모형화 및 예측: 예측 옵션 설정	22
예측 단계 사용	22
예측 구간 지정	22
도표 생성 및 결과 저장	23
제목 및 각주 지정	24
속성 보기	24

ARIMA 모델 개발 및 예측 작업 정보

ARIMA 예측 모델링은 ARIMA(자기회귀 통합이동평균) 또는 ARMA(자기회귀 이동평균) 모델을 사용하여 등간격 일변량 시계열 데이터, 전달함수 데이터 및 개입 데이터를 분석하고 예측합니다. ARIMA 모형은 반응시계열의 값을 해당 시계열의 과거값, 과거 오차(충격 또는 혁신이라고도 함) 및 다른 시계열의 현재 및 과거값의 선형 조합으로 예측합니다.

ARIMA 분석에는 다음과 같은 3가지 기본 수행 단계가 있습니다.

- 1단계: 시계열 식별
- 2단계: 시계열 추정
- 3단계: 시계열의 미래값 예측

ARIMA 모델화 및 예측 작업을 사용하여 다음 분석을 수행할 수 있습니다.

- 계절 요인의 영향도를 결정합니다. 예를 들어 계절 변수(기온, 강우량, 습도 등)를 입력으로 하여 시계열을 모델화할 수 있습니다.
- 과거 패턴을 기반으로 한 제품 수요를 예측합니다.

주: ARIMA 모델 개발 및 예측 작업은 3가지 단계가 각각 별도의 페이지로 구성되어 있습니다. **실행**을 클릭하면 선택한 페이지에 해당하는 단계 및 그 이전 단계만 수행됩니다. 따라서 **3단계: 예측** 페이지로 이동하여 변경하고 **1단계: 식별** 페이지로 이동하여 **실행**을 클릭하면 식별 단계만 수행됩니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	ARIMA
필수 SAS 제품	Base SAS, SAS/ETS
추가 SAS 제품(권장)	SAS/GRAPH

ARIMA 모델화 및 예측: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
시계열 변수	예측할 변수를 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다.
시간 ID 변수	행에 날짜를 할당하는 데 사용될 변수를 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다. SAS 날짜, 시간 또는 날짜/시간 출력형식을 갖는 첫 번째 변수가 자동으로 이 역할에 할당됩니다. 시간 ID 변수를 생성할 수도 있습니다.
예측 그룹화 기준	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다.

역할 이름	설명
	주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다.
	주: 종속변수나 독립변수로 이미 선택한 변수로는 분석을 그룹화할 수 없습니다.

ARIMA 모형화 및 예측: 식별 옵션 설정

반응시계열의 차분

선택 영역의 1단계: 식별 머리글 아래에서 **차분**을 클릭하여 해당 옵션에 액세스합니다.

반응시계열의 차분을 수행하려면 **반응시계열의 차분** 체크 박스를 선택합니다. 이 체크 박스를 선택하면 첫 번째 차분이 기본적으로 수행됩니다. 차분의 차수를 변경하려면 **차분 시차** 상자에 새 값을 입력합니다. 예를 들어 시차 1에서 차분을 두 번 수행하려면 1,1을 지정합니다.

정상성 검정 지정

선택 영역의 1단계: 식별 머리글 아래에서 **정상성 검정**을 클릭하여 해당 옵션에 액세스합니다.

정상성 검정을 지정하려면 원하는 검정의 옵션 버튼을 선택합니다. 기본적으로 검정은 수행되지 않습니다. 정상성 검정을 선택하면 검정에 대한 자기회귀 차수를 지정할 수 있습니다. 기본적으로 자기회귀 차수 0, 1, 2가 지정됩니다. 자기회귀 차수 상자에 원하는 **자기회귀 차수**를 침표로 구분하여 입력하거나 줄임표를 사용하여 차수 범위를 입력합니다. 예를 들어 자기회귀 차수 1, 2 및 3을 지정하려면 **1...3**을 입력합니다.

Augmented Dickey-Fuller 검정 수행을 선택하면 계절성 시차를 지정하여 계절 Dickey-Fuller 검정을 수행할 수 있습니다. **계절성 시차** 상자에 정수값을 지정하십시오. 계절성 시차의 기본값은 비계절성 검정인 1입니다.

도표 생성 및 결과 저장

선택 영역의 1단계: 식별 머리글 아래에서 **도표** 및 결과를 클릭하여 해당 옵션에 액세스합니다.

옵션 이름	설명
실제 데이터 값 도표	결과에 입력 시계열 그래프를 표시합니다. 또한 자기상관 및 교차상관을 계산할 때 고려할 시차의 수를 지정할 수도 있습니다. 기본 시차 수는 24 또는 관측값 수의 1/4이며, 이 중 더 적은 것이 사용됩니다.
자기상관 및 교차공분산 저장	자기상관 및 교차공분산의 출력 테이블을 저장합니다. 주: SAS Enterprise Guide는 옵션 창의 작업> 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기 를 클릭합니다.
식별 결과 출력 안 함	식별 결과를 생략합니다.

ARIMA 모델화 및 예측: 추정 옵션 설정

추정 단계 사용

선택 영역의 2단계: 추정 머리글 아래에서 **추정 단계 사용**을 클릭하여 해당 옵션에 액세스합니다.

추정 단계 옵션을 활성화하려면 **추정 단계 수행** 체크 박스를 선택합니다.

주: ARIMA 모델 개발 및 예측 작업은 3단계로 구성됩니다. 1단계: 식별 옵션은 항상 활성화되어 있습니다. 2단계: 추정 및 3단계: 예측에 대한 옵션을 활성화하려면 적절한 체크 박스를 선택해야 합니다. 예측 단계를 활성화하려면 먼저 추정 단계를 활성화해야 합니다.

ARIMA 모델 정의

선택 영역의 2단계: 추정 머리글 아래에서 **모델 정의**를 클릭하여 해당 옵션에 액세스합니다.

모델 추정에는 자기회귀모수 및 이동평균모수를 모두 추가할 수 있습니다.

모델 추정에 새로운 모수를 추가하는 방법:

- 1 리스트의 각 승법 인수에 대해 추정된 자기회귀모수 또는 이동평균 모수에 대한 시차를 나타내는 정수값, 심표로 구분된 정수 리스트 또는 하이픈이 붙은 정수 리스트를 **AR 모델에 대한 인수** 또는 **MA 모델에 대한 인수** 상자에 입력합니다. 예를 들어, **1...3**, **1 TO 3** 또는 **1,2,3**을 입력하면 (1, 2, 3)을 지정합니다. **3**을 입력하면 (3)을 입력합니다.
- 2 **추가**를 클릭하거나 Enter를 눌러 새로운 파라미터를 리스트에 추가합니다.
- 3 이 절차를 반복하여 승법 항목을 추가합니다. 예를 들어 (1)(12)를 지정하려면 **1**을 입력하고 **추가**를 클릭한 후, **12**를 입력하고 **추가**를 클릭합니다.

모수가 생성된 후 **자기회귀** 또는 **이동평균** 영역에 있는 리스트에서 모수를 편집하거나 삭제할 수 있습니다.

추정 방법 및 옵션 선택

선택 영역의 2단계: 추정 머리글 아래에서 **모델 옵션**을 클릭하여 해당 옵션에 액세스합니다.

옵션 이름	설명
추정 방법	드롭다운 리스트에서 추정 방법을 선택합니다. 다음과 추정 방법을 사용할 수 있습니다. ■ 최대우도법은 최대우도 추정값을 생성합니다. 우도함수는 Marquardt의 방식을 사용하여 비선형 최소제곱을 통해 최대화됩니다. ■ 비조건부 최소제곱 방법은 비조건부 최소제곱 추정값을 생성합니다. ULS 방법은 ELS(Exact Least Squares) 방법이라고도 합니다. ■ 조건부 최소제곱 방법은 조건부 최소제곱 추정값을 생성합니다. 이 추정값은 과거의 비관측 오류가 0이라는 가정을 조건으로 합니다. 기본값입니다.
최대 반복 횟수	추정 단계에 대한 최대 반복 횟수를 지정합니다. 최대 반복 횟수 상자에 정수값을 지정합니다. 기본값은 50 입니다.
모델에서 평균 항 제거	모델에서 평균 항을 제외합니다.

결과 저장

선택 영역의 2단계: 추정 머리글 아래에서 **결과**를 클릭하여 해당 옵션에 액세스합니다.

옵션 이름	설명
출력 데이터 저장	결과의 전부 또는 일부를 출력 데이터셋에 저장할 수 있습니다. 해당 체크 박스를 선택하여 결과를 저장합니다. 주: SAS Enterprise Guide는 옵션 창의 작업 > 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
추정 결과 출력 안 함	추정 단계에서 결과를 생략합니다.

ARIMA 모델화 및 예측: 예측 옵션 설정

예측 단계 사용

선택 영역의 3단계: 추정 머리글 아래에서 **추정 단계 사용**을 클릭하여 해당 옵션에 액세스합니다.

예측 단계 옵션을 활성화하려면 **예측 단계 수행** 체크 박스를 선택합니다.

주: ARIMA 모델 개발 및 예측 작업은 3단계로 구성됩니다. 1단계: 식별 옵션은 항상 활성화되어 있습니다. 2단계: 추정 및 3단계: 예측에 대한 옵션을 활성화하려면 적절한 체크 박스를 선택해야 합니다. 예측 단계를 활성화하려면 먼저 추정 단계를 활성화해야 합니다.

예측 구간 지정

선택 영역의 3단계: 추정 머리글 아래에서 **추정 옵션**을 클릭하여 해당 옵션에 액세스합니다.

옵션 이름	설명
관측값 사이 시간 간격	입력 데이터 테이블의 관측값 간의 시간 간격을 드롭다운 리스트에서 선택합니다. 입력 데이터에 적합한 간격을 선택해야 하며 그렇지 않으면 분석이 실패하거나 부정확한 결과를 얻게 됩니다.

옵션 이름	설명
	단위 수 를 시간 간격으로 선택하면 간격당 시간 단위도 지정할 수 있습니다. 기본값은 1 입니다. 단위 수 이외의 시간 간격을 선택하면 간격당 시간 단위 상자를 사용할 수 없습니다.
예측할 간격 수	예측치를 얻고자 하는 시간 간격 수를 지정합니다. 기본값은 24입니다. 예를 들어 관측값 간의 시간 간격을 월로 지정하고 예측할 시간 간격 수를 6으로 지정하면, 입력 데이터의 끝 부분부터 6개월 동안의 예측을 얻게 됩니다. 신뢰수준 드롭다운 리스트에서 신뢰수준을 선택합니다. 기본값은 95% 신뢰수준입니다.

도표 생성 및 결과 저장

선택 영역의 3단계: 예측 머리글 아래에서 **도표 및 결과**를 선택하여 해당 옵션에 액세스합니다.

옵션 이름	설명
예측 도표 옵션	다음과 같은 도표를 생성합니다. ■ 실제 값, 예측값, 1단계 앞의 예측값, 신뢰한계를 도표를 생성하려면, 예측 체크 박스를 선택합니다. ■ 잔차 도표를 생성하려면 잔차 체크 박스를 선택합니다. 표시할 실제 데이터의 간격 수 제한 및 표시할 간격 수 상자를 사용하여 예측 데이터가 시작되기 전에 도표에 표시할 실제 데이터의 간격 수를 지정할 수 있습니다.
예측값 저장	예측에 대한 출력 테이블을 생성하려면 예측값 저장 체크 박스를 선택합니다. 주: SAS Enterprise Guide는 옵션 창의 작업> 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
예측 결과 출력 안 함	예측 결과를 생략합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

프로젝트 라이브러리 할당

프로젝트 라이브러리 할당 마법사 정보	25
프로젝트 라이브러리 할당: 이름 및 서버 지정	26
프로젝트 라이브러리 할당: 엔진 지정	26
프로젝트 라이브러리 할당: 구성 옵션 지정	28
프로젝트 라이브러리 할당: 요약 정보 검토	29

프로젝트 라이브러리 할당 마법사 정보

프로젝트 라이브러리 할당 마법사로 프로젝트에 대한 라이브러리를 정의할 수 있습니다. 마법사는 라이브러리를 정의하는 데 필요한 LIBNAME 문을 생성합니다. 이 라이브러리는 작업이나 프로젝트에 작업을 포함하는 프로세스 플로우를 실행할지를 정의합니다.

예를 들어 사용자가 분석해야 할 데이터가 있다고 가정합니다. 이 데이터는 사용자의 로컬 디렉터리에 있고 관리자가 정의한 SAS 라이브러리에는 없습니다. 데이터에 액세스할 수 있도록 SAS 라이브러리를 정의해야 합니다. **프로젝트 라이브러리 할당** 마법사를 사용하여 이 LIBNAME 문을 생성할 수 있습니다.

주: 이 작업은 **SAS Enterprise Guide**에서만 사용 가능합니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	없음
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	없음

프로젝트 라이브러리 할당: 이름 및 서버 지정

- 1 **이름** 상자에 라이브러리의 이름을 입력합니다. 이름은 모두 대문자이고 8자 이하여야 합니다. 이름은 라이브러리가 생성된 각 서버에 대해 고유해야 합니다. 이름에는 다음과 같은 문자를 포함할 수 없습니다: \ / : * ? " < > |
- 2 사용 가능한 서버 리스트에서 서버를 선택합니다.
- 3 **다음**을 클릭하여 라이브러리에 대한 엔진을 선택합니다.

프로젝트 라이브러리 할당: 엔진 지정

- 1 엔진을 선택하기 전에 라이브러리에 대한 이름 및 서버를 지정해야 합니다.
- 2 엔진 유형을 선택합니다. 다음 엔진 유형 중에서 선택할 수 있습니다.

엔진 유형	설명
파일 시스템	다음과 같은 옵션을 지정합니다. ■ 라이브러리에 대해 선택한 경로에 있는 파일의 유형을 기반으로 SAS가 엔진을 선택하도록 허용하려면 지정된 경로에 있는 콘텐츠를 기반으로 SAS가 엔진을 선택합니다 체크 박스를 선택합니다. 주: 라이브러리 경로에 여러 가지 파일 유형이 포함되어 있으면 이 체크 박스를 선택하지 마십시오. ■ 라이브러리에 대한 엔진을 선택합니다. 엔진으로 REMOTE-SAS/SHARE를 선택하면 라이브러리 구성 옵션에서 slibref/SAMPSHR의 이름-값 쌍을 지정해야 합니다. 주: 지정된 경로에 있는 콘텐츠를 기반으로 SAS가 엔진을 선택합니다 체크 박스를 선택하면 엔진을 선택할 수 없습니다. ■ 서버에 있는 라이브러리의 실제 경로를 경로 상자에 입력합니다. 찾아보기를 클릭하여 이 위치를 선택합니다. 경로에 와일드카드를 사용할 수 있습니다. <userid> Windows 사용자 ID를 대체합니다.

엔진 유형	설명
	<sample> Enterprise Client 샘플 디렉터리를 대체합니다. <serveruser> 서버 로그인 사용자 ID를 대체합니다. <serverpassword> 서버 로그인 암호를 대체합니다. <libraryuser> 라이브러리 로그인 사용자 ID를 대체합니다. <librarypassword> 라이브러리 로그인 암호를 대체합니다. 선택한 서버에 액세스하기 위해 자격 증명이 필요하면 사용자 ID와 암호를 입력하라는 프롬프트가 나타납니다. 주: BASE, V9, V8, V7, V604 또는 V6을 엔진으로 선택하면 경로를 지정할 필요가 없습니다. 다른 모든 엔진에는 경로가 필요합니다.
데이터베이스 시스템	다음과 같은 옵션을 지정합니다. - 라이브러리의 파일에 액세스할 때 사용할 엔진을 선택합니다. - 라이브러리에 대한 외부 Database 서버를 선택합니다. 주: 활성 SAS 메타데이터 레파지토리에 정의되어 있는 Database 서버가 드롭다운 리스트에 표시됩니다. 사용할 서버가 메타데이터에 정의되어 있지 않으면 해당 서버가 드롭다운 리스트에 표시되지 않습니다. 관리자가 SAS Management Console을 사용하여 추가 서버를 정의할 수 있습니다. - 서버의 데이터에 액세스하는 데 사용되는 데이터베이스 스키마를 지정합니다. 제공한 스키마 이름은 Database 서버에 정의되어 있는 스키마의 이름과 일치해야 합니다. - Database 서버에 대한 사용자 ID 및 암호를 입력합니다.
WebDAV	다음과 같은 옵션을 지정합니다. - 라이브러리 소스로 사용할 WebDAV 서버를 선택합니다. 주: 활성 SAS 메타데이터 레파지토리에 정의되어 있는 WebDAV 서버가 드롭다운 리스트에 표시됩니다. 사용할 서버가 메타데이터에 정의되어 있지 않으면 해당 서버가 드롭다운 리스트에 표시되지 않습니다. 관리자가 SAS Management Console을 사용하여 추가 서버를 정의할 수 있습니다. - 서버에 있는 라이브러리의 실제 경로를 경로 상자에 입력합니다. 서버나 라이브러리 옵션에 있는 라이브러리의 실제 경로를 지정할 때 다음과 같은 와일드카드를 사용할 수 있습니다. <userid> Windows 사용자 ID를 대체합니다.

엔진 유형	설명
	<sample> Enterprise Client 샘플 디렉터리를 대체합니다. <serveruser> 서버 로그인 사용자 ID를 대체합니다. <serverpassword> 서버 로그인 암호를 대체합니다. <libraryuser> 라이브러리 로그인 사용자 ID를 대체합니다. <librarypassword> 라이브러리 로그인 암호를 대체합니다. ■ WebDAV 서버에 대한 사용자 ID 및 암호를 입력합니다.

- 3 다음을 클릭하여 라이브러리에 대한 구성 옵션을 지정합니다.

프로젝트 라이브러리 할당: 구성 옵션 지정

- 구성 옵션을 지정하기 전에 라이브러리에 대한 엔진을 선택해야 합니다.
- 이름-값 쌍을 가진 라이브러리 옵션을 지정합니다. 테이블의 값과 관련된 옵션만 지정할 수 있습니다. 이름-값 쌍이 있는 옵션의 예는 ACCESS=READONLY입니다. 이 옵션은 라이브러리를 읽기 전용으로 지정하여 해당 라이브러리에 출력 데이터를 쓸 수 없도록 합니다. 사용 가능한 옵션 리스트에 대한 자세한 내용은 운영 환경에 대한 문서를 참조하십시오.

주: 엔진으로 **REMOTE-SAS/SHARE**를 선택하면 라이브러리 구성 옵션에서 slibref/SAMPSHR의 이름-값 쌍을 지정해야 합니다.

옵션 이름 또는 값을 입력하는 방법:

- 테이블에서 빈 셀을 클릭합니다. 옵션 이름과 값을 지정할 때 빈 행이 테이블에 자동으로 나타납니다.
- 이름** 칼럼에 옵션의 이름을 입력합니다. 모든 올바른 LIBNAME 문 옵션을 지정할 수 있습니다. 그러나 프롬프트 옵션(예: DBPROMPT=YES)은 사용하지 않아야 합니다.
- 옵션에 해당하는 값을 값 칼럼에 입력합니다. 빈 값은 허용되지 않습니다.

주: 라이브러리가 Sybase 데이터베이스에 연결되면 READLOCK_TYPE=PAGE를 옵션으로 포함해야 합니다.

- 3 값이 필요하지 않은 모든 옵션을 **추가 옵션** 상자에 입력합니다.

- 4 다음을 클릭하여 요약 정보를 검토합니다.

프로젝트 라이브러리 할당: 요약 정보 검토

새로운 라이브러리를 생성하기 전에 지정한 정보를 검토할 수 있습니다.

- 1 요약 정보를 검토하기 전에 라이브러리에 대한 구성 옵션을 지정할 수 있습니다.
- 2 텍스트 상자의 정보를 검토합니다. 값을 변경하려면 **뒤로**를 클릭합니다.
- 3 선택한 서버에서 새로운 라이브러리를 테스트하려면 **라이브러리 테스트**를 클릭합니다. **SAS Enterprise Guide**가 서버로 연결되고 라이브러리를 할당하면 해당 서버 상태로 **확인**이 표시됩니다. 테스트에 대한 로그를 확인하려면 **로그 표시**를 클릭합니다.
- 4 **마침**을 클릭하여 새로운 라이브러리를 생성합니다.

막대 그래프

막대 그래프 작업 정보	31
막대 그래프: 그래프 유형 선택	32
막대 그래프: 분석 역할에 변수 할당	33
막대 그래프: 모양 옵션 설정	34
레이아웃 지정	34
추가 모양 옵션	34
제목 및 각주 지정	35
속성 보기	35

막대 그래프 작업 정보

막대 그래프 작업을 통해 그래프 변수의 숫자값 또는 통계량을 비교하는 수직, 수평 또는 3D 막대 그래프 등이 생성됩니다. 막대 그래프는 여러 가지 높이의 막대를 사용하여 데이터의 상대적 크기를 표시합니다. 각 막대는 하나의 데이터 범주를 나타냅니다.

막대 그래프를 사용하여 각 상점의 총 판매량을 비교할 수 있습니다. 이 그래프 유형에서 각 막대는 각 위치의 총 판매를 나타냅니다.

자세한 내용은 [“그래픽 출력형식 선택” \(5페이지\)](#)에서 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GCHART
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

막대 그래프: 그래프 유형 선택

이 옵션에 액세스하려면 선택 영역에서 **막대 그래프**를 클릭합니다.

옵션 이름	설명
단순	막대로 데이터의 크기를 표시하고 각각의 막대는 데이터의 범주를 나타냅니다. 막대의 길이(또는 높이)는 해당 중간점에 대한 그래프 통계량값을 나타냅니다.
그룹형	그룹화 칼럼 역할에 할당된 변수의 각 고윳값에 대해 개별 막대 그룹을 생성합니다.
누적형	누적화 칼럼 역할에 할당된 모든 변수의 고윳값에 대한 각 막대 안에서 별도의 조각을 생성합니다. 막대 내의 각 조각은 서로 다른 색으로 표시됩니다.
그룹형/누적형	그룹형 막대 그래프와 누적형 막대 그래프를 결합합니다. 이 옵션을 선택하면 그룹화 칼럼 역할에 할당된 변수의 각 고윳값마다 개별 막대 그룹이 생성되고 누적화 칼럼 역할에 할당된 변수의 모든 고윳값에 대하여 각 막대 내에서 별도의 조각이 생성됩니다.
컬러	각 막대를 서로 다른 색으로 표시한 단순 막대 그래프를 생성합니다.
그룹형 컬러	그룹화 칼럼 역할에 할당된 변수의 각 고윳값에 대해 개별 막대 그룹을 생성합니다. 그룹의 각 막대는 서로 다른 색으로 표시됩니다.
컬러 그룹형 그룹	그룹화 칼럼 역할에 할당된 변수의 각 고윳값에 대해 개별 막대 그룹을 생성합니다. 그룹의 막대는 같은 색으로 표시됩니다.
인터리브	그룹화 칼럼 역할에 할당된 변수의 각 고윳값에 대해 개별 막대 그룹을 생성합니다. 각 그룹의 막대는 서로 연결되어 있으며 각 막대는 다른 색으로 표시됩니다.
3D 그룹형	그룹화 칼럼 역할에 할당된 변수의 각 고윳값에 대해 개별 3D 막대 그룹을 생성합니다.
3D 그룹형/누적형	그룹형 막대 그래프와 누적형 막대 그래프를 결합합니다. 이 옵션을 선택하면 그룹화 칼럼 역할에 할당된 변수의 각 고윳값마다 개별 3D 막대 그룹이 생성되고 누적화 칼럼 역할에 할당된 변수의 모든 고윳값에 대하여 각 막대 내에서 별도의 조각이 생성됩니다.
3D 그룹형 컬러	그룹화 칼럼 역할에 할당된 변수의 각 고윳값에 대해 개별 3D 막대 그룹을 생성합니다. 그룹의 각 막대는 서로 다른 색으로 표시됩니다.

옵션 이름	설명
3D 컬러 그룹형 그룹	그룹화 칼럼 역할에 할당한 변수의 각 고윳값에 대해 개별 3D 막대 그룹을 생성합니다. 그룹의 막대는 같은 색으로 표시됩니다.
다중 축도 그룹	합계 칼럼 역할에 여러 분석변수를 할당할 수 있습니다. 생성되는 그래프에는 그래프 칼럼 역할에 할당한 변수의 각 고윳값에 대한 막대 그룹이 포함됩니다. 그룹의 막대는 같은 색으로 표시됩니다.

막대 그래프: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

주: 선택한 그래프 유형에 따라 사용할 수 있는 역할이 다릅니다.

역할 이름	설명
그래프 칼럼	이 역할에 할당하는 칼럼의 값에 따라 그래프로 나타낼 범주가 정의됩니다. 이 역할에는 변수를 하나만 할당해야 합니다.
그룹화 칼럼	이 역할에 할당하는 칼럼의 값에 따라 막대를 나눌 그룹이 결정됩니다. 이 역할에는 변수를 하나만 할당해야 합니다.
누적화 칼럼	이 역할에 할당하는 칼럼 값에 따라 각 막대의 조각 개수가 결정됩니다.
합계 칼럼	이 역할에 할당하는 칼럼에 따라 막대의 길이가 결정됩니다. 막대 길이 결정에 사용되는 통계량을 선택하려면 선택 영역에서 고급 을 선택합니다. 이 역할에 칼럼을 할당하지 않으면 그래프 칼럼 의 각 값의 빈도에 따라 막대의 길이 또는 조각의 크기가 결정됩니다.
그룹 분석 칼럼	이 역할에 할당하는 칼럼의 값에 따라 생성되는 그래프의 수가 결정됩니다. 각 고윳값에 대해 별도의 그래프가 생성됩니다.

막대 그래프: 모양 옵션 설정

레이아웃 지정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **레이아웃**을 클릭합니다.

기본적으로 대부분의 막대 그래프는 2차원이며 **2D** 체크 박스가 선택되어 있습니다. 3차원 그래프를 생성하려면 **2D** 체크 박스를 선택 취소하고 **모양** 드롭다운 리스트에서 막대 모양을 선택합니다.

주: 3차원 막대 그래프를 그래프 유형으로 선택했으면 **2D** 체크 박스를 사용할 수 없습니다.

순서 드롭다운 리스트를 사용하여 막대의 정렬 순서를 지정합니다. 기본적으로 정렬 순서는 자동으로 결정됩니다. 막대를 길이나 높이에 따라 오름차순 또는 내림차순으로 배열할 수 있습니다.

각 막대에 외곽선을 표시하려면 **외곽선 색상** 체크 박스를 선택합니다. 드롭다운 리스트에서 외곽선 색상을 선택할 수 있습니다.

막대 크기 드롭다운 리스트를 사용하여 **막대 크기**를 설정합니다. 기본적으로 막대 크기는 자동으로 결정됩니다. 그러나 텍스트 상자에 막대 너비 또는 막대 간격을 총 그래프 너비의 백분율로 지정할 수 있습니다.

추가 모양 옵션

일부 모양 옵션은 그래프 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- [“막대 모양 지정” \(598페이지\)](#)
- [“축 옵션 설정” \(593페이지\)](#)
- [“참조선 추가” \(595페이지\)](#)
- [“범례 옵션 설정” \(596페이지\)](#)
- [“그래프 영역 사용자 정의” \(597페이지\)](#)
- [“고급 옵션 설정” \(602페이지\)](#)

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

막대 그래프 마법사

막대 그래프 마법사 정보	37
막대 그래프 마법사: 데이터 선택	37
막대 그래프 마법사: 역할에 변수 할당	37
막대 그래프 마법사: 모양 옵션 설정	38
막대 그래프 마법사: 제목 및 각주 지정	39

막대 그래프 마법사 정보

막대 그래프 마법사에서는 그룹형/누적형 막대에 대한 옵션을 사용하여 2차원이나 3차원의 수직 또는 수평 막대 그래프를 생성할 수 있습니다.

막대 그래프 마법사: 데이터 선택

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

다음을 클릭하여 역할에 변수를 할당합니다.

막대 그래프 마법사: 역할에 변수 할당

1 데이터를 선택해야 역할에 변수를 할당할 수 있습니다.

2 다음과 같은 역할에 변수를 할당합니다.

- **막대** - 이 역할에 지정하는 변수의 값에 따라 막대의 개수가 달라집니다. 변수는 문자 또는 숫자가 될 수 있습니다. 이 역할에는 변수를 하나만 할당해야 합니다. 막대를 정렬하거나 변수 유형을 지정할 수 있습니다.
 - **정렬** - 막대의 정렬 순서를 지정할 수 있습니다. 기본적으로 막대는 중간값에 대해 오름차순으로 자동으로 정렬됩니다. 오름차순 또는 내림차순 막대 길이로 막대를 정렬할 수 있습니다.
 - **변수 유형** - 기본적으로 막대 수는 자동으로 결정됩니다. 숫자 칼럼이면 값이 여러 범위로 나뉘며 각 범위마다 막대가 한 개씩 표시됩니다. 숫자 칼럼이 이산 값을 나타내면 **이산** 옵션을 선택할 수 있습니다.
- **막대 길이** - 이 역할에 지정하는 변수나 통계량에 따라 막대 길이가 결정됩니다.
 - 이 역할에 변수를 할당하면 각 막대의 합계 또는 평균을 계산할 수 있습니다. 기본적으로 합이 계산됩니다.
 - 이 역할에 변수를 할당하지 않으면, 각 막대의 빈도, 누적빈도, 백분율 또는 누적 백분율을 계산할 수 있습니다. 기본적으로 빈도가 계산됩니다.

주: 수평 막대 그래프를 생성할 때 이 옵션은 **막대 길이**입니다.

3 (선택 사항) 다음과 같은 역할에 변수를 할당합니다.

- **그룹화** - 이 역할에 할당한 변수의 값에 따라 막대를 나눌 그룹이 결정됩니다. 수직 막대 그래프에서는 막대 그룹이 X 축을 따라 표시됩니다. 수평 막대 그래프에서는 막대 그룹이 Y 축을 따라 표시됩니다.

주: **깊이** 역할에 변수를 할당하면 이 옵션을 사용할 수 없습니다.

- **깊이** - 이 역할에 할당한 변수의 값에 따라 막대를 나눌 그룹이 결정됩니다. 막대 그룹은 Z 축을 따라 표시됩니다.

주: 수평 막대 그래프를 생성하거나 **그룹화** 역할에 변수를 할당하면 이 옵션은 사용할 수 없습니다.

- **누적화** - 이 역할에 지정하는 변수의 값에 따라 막대 그래프에서 각 막대의 조각 개수가 결정됩니다.
- **그래프 분류** - 각 그룹에 대해 생성된 그래프를 구별합니다. 이 역할에 지정하는 변수의 값에 따라 그룹이 결정됩니다.

4 다음을 클릭하여 모양 옵션을 설정합니다.

막대 그래프 마법사: 모양 옵션 설정

1 그래프 모양을 설정하려면 먼저 역할에 변수를 할당해야 합니다.

- 2 **3D 그래프** 체크 박스를 선택하여 3D 막대 그래프를 생성합니다. 기본적으로 막대는 2차원입니다.
- 3 **막대 색상 분류** 드롭다운 리스트에서 그래프의 막대에 색상을 할당하는 방법을 지정합니다. 기본적으로 **모든 막대 동일** 및 **막대 범주** 옵션을 사용할 수 있습니다. 변수를 **그룹화** 기준 또는 **깊이** 역할에 할당하면 **그룹 변수** 또는 **깊이 변수** 옵션도 사용할 수 있습니다.

주: **누적화** 역할에 변수를 할당하면 이 옵션을 사용할 수 없습니다.

모든 막대의 색상이 같으면 **막대 색상** 드롭다운 리스트에서 색상을 선택합니다. 각 막대 범주의 색상이 다르면 막대 그래프 마법사에서 색상 구성이 자동으로 선택됩니다.

- 4 범례를 사용할 수 있으면 그래프에 포함하려면 **범례** 체크 박스를 선택합니다. 드롭다운 리스트에서 범례 위치를 선택합니다.
- 5 각 막대의 레이블로 표시할 통계량을 지정하려면 **데이터 레이블** 체크 박스를 선택한 다음, 드롭다운 리스트에서 표시할 통계량을 선택합니다. **막대 길이** 역할에 변수를 할당하지 않으면 빈도, 누적빈도, 백분율 및 누적백분율 통계량을 사용할 수 있습니다. **막대 길이** 역할에 변수를 할당하면 합 및 평균 통계량도 사용할 수 있습니다.
- 6 할당된 각 변수의 레이블을 편집하려면 **축 레이블**을 클릭합니다. **축 레이블** 대화 상자가 나타납니다.
- 7 그래프에 참조선을 포함하려면 **참조선 사용** 체크 박스를 선택합니다.
- 8 그래프에 눈금 표식을 포함하려면 **눈금 표식** 체크 박스를 선택합니다.
- 9 **다음**을 클릭하여 제목 및 각주를 지정합니다.

막대 그래프 마법사: 제목 및 각주 지정

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

- 1 모양 옵션을 설정해야 제목과 주석을 지정할 수 있습니다.
- 2 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다. 제목과 각주에는 매크로 변수를 사용할 수 있습니다.
- 3 **마침**을 클릭하여 작업을 실행합니다.

막대-선 그래프

막대-선 그래프 작업 정보	41
막대-선 그래프: 분석 역할에 변수 할당	42
막대-선 그래프: 모양 옵션 설정	43
보간법 선택	43
추가 모양 옵션	43
제목 및 각주 지정	44
속성 보기	44

막대-선 그래프 작업 정보

막대-선 그래프 작업을 통해 선 도표를 중첩시킨 수직 막대 그래프가 생성됩니다. 선 도표는 입력 데이터셋의 변수 중 하나에 대해 계산된 통계량값을 나타냅니다.

이 작업을 사용하여 다음과 같은 작업을 수행할 수 있습니다.

- 정확하고 상대적인 크기 표시 및 비교
- 전체에 대한 각 부분의 기여도 검사
- 데이터의 추세 및 패턴 결정

예를 들어 이 작업을 사용하여 제품이 팔리고 있는 각 지역에 대하여 제품 판매와 가격 사이의 관계를 표시할 수 있습니다. 가로축은 중간점축이며 각 지역에 대한 막대를 포함합니다. 세로축은 반응축입니다. 각 막대는 각 지역에 대한 판매량을 나타냅니다. 오른쪽 세로축에는 각 지역에서의 제품 생산 및 마케팅 가격에 대한 선 도표가 막대 위에 포개져 있습니다. 이 그래프를 사용하여 각 지역 및 전체 지역의 가격 대 판매를 비교할 수 있습니다.

자세한 내용은 [“그래픽 출력형식 선택” \(5페이지\)](#)에서 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
----------	--------------

사용되는 SAS 프로시저 GBARLINE

요구 사항 이름	프로시저 및 제품 이름
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

막대-선 그래프: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
그래프 칼럼	이 역할에 할당하는 칼럼의 값에 따라 그래프로 나타낼 범주가 정의됩니다. 이 역할에는 변수를 하나만 할당해야 합니다.
막대 합계	이 역할에 할당하는 칼럼에 따라 막대의 길이가 결정됩니다. 기본적으로 이 역할에 칼럼이 할당되지 않으면 총 관측값 수(빈도통계량)에 따라 막대 길이가 결정됩니다. 이 역할에 칼럼이 할당되면 각 범주에 대한 값의 합계에 따라 막대 길이가 결정됩니다. 막대 길이 결정에 사용되는 통계량을 선택하려면 선택 영역에서 고급 을 선택합니다. 자세한 내용은 고급 옵션 설정을 참조하십시오.
선 합계	이 역할에 할당하는 칼럼에 따라 선 도표에 사용할 변수가 지정됩니다. 이 역할에는 둘 이상의 변수를 할당할 수 있습니다. 각 선의 도표를 생성하는 데 사용되는 통계량을 선택하려면 선택 영역에서 도표 를 선택합니다. 기본적으로 각 범주의 값 합계를 사용하여 각 선의 도표를 생성합니다. 주: 이 역할에 칼럼을 할당하지 않으면 각 범주의 총 관측값 수(빈도통계량)가 선의 도표를 생성하는 데 사용됩니다. 선의 도표를 생성하는 데 사용되는 통계량을 선택하려면 선택 영역에서 고급 을 선택합니다.
누적화 칼럼	이 역할에 할당하는 칼럼 값에 따라 각 막대의 조각 개수가 결정됩니다.
그룹 분석 칼럼	이 역할에 할당하는 칼럼의 값에 따라 생성되는 그래프의 수가 결정됩니다. 각 고유티 값에 대해 별도의 그래프가 생성됩니다.

막대-선 그래프: 모양 옵션 설정

보간법 선택

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **보간법**을 클릭합니다.

- 1 리스트에서 변수를 선택합니다. 이러한 변수는 **선 함께** 역할에 할당된 변수입니다.
- 2 드롭다운 리스트에서 보간법을 선택합니다. 선택한 보간법에 따라 추가 설정이 필요할 수도 있습니다. 다음과 같은 보간법을 사용할 수 있습니다.
 - **선** - 데이터 포인트를 직선으로 연결합니다. 포인트는 입력 데이터에 나타난 순서대로 연결됩니다.
 - **산점도** - 보간을 생략합니다. 데이터 포인트 도표가 생성됩니다.
 - **스텝** - 스텝함수를 사용하여 데이터를 도표로 그립니다.
- 3 (선택 사항) 동일한 보간법 및 속성 설정을 리스트의 모든 변수에 사용하려면 **모두 적용** 체크 박스를 선택합니다. 이 옵션은 리스트에 둘 이상의 변수가 있을 때만 사용할 수 있습니다.

추가 모양 옵션

일부 모양 옵션은 그래프 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- ["막대 모양 지정" \(598페이지\)](#)
- ["막대 옵션 설정" \(599페이지\)](#)
- ["선 도표 옵션 설정" \(600페이지\)](#)
- ["축 옵션 설정" \(593페이지\)](#)
- ["참조선 추가" \(595페이지\)](#)
- ["범례 옵션 설정" \(596페이지\)](#)
- ["그래프 영역 사용자 정의" \(597페이지\)](#)
- ["고급 옵션 설정" \(602페이지\)](#)

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

기본 예측 모형

기본 예측 모형 작업 정보	45
기본 예측: 분석 역할에 변수 할당	46
기본 예측: 예측 옵션 설정	47
예측 방법 및 간격 선택	47
입력 데이터 옵션 선택	47
예측 방법 옵션 설정	48
기본 예측: 도표 생성	49
기본 예측: 결과 옵션 설정	49
제목 및 각주 지정	50
속성 보기	50

기본 예측 모형 작업 정보

기본 예측 모형 작업에서는 빠른 자동화 절차를 통해 여러 시계열에 대한 예측을 한 번에 생성할 수 있습니다. 이 작업은 외삽적 예측 방법을 사용합니다. 여기서 시계열에 대한 예측은 다른 변수가 아닌 시간 및 과거 값의 함수입니다.

제품의 판매량을 예측할 때 이 작업을 사용할 수 있습니다. 예를 들어 데이터셋에 일부 제품의 판매량에 대한 월 데이터를 저장하였고 다음 10개월 동안의 판매량을 예측한다고 가정합니다. 이 작업은 예측값의 날짜와 예측된 판매량의 신뢰한계를 계산하고 사용자는 SAS 데이터셋에 결과를 저장할 수 있습니다. 또한 예측 모델이 얼마나 데이터에 적합한지 측정한 통계량과 사용된 예측 모델의 모수도 결과 SAS 데이터셋에 저장할 수 있습니다. 사용되는 예측 모델의 파라미터와 예측 모델이 데이터를 얼마나 잘 적합시키는지를 측정하는 통계량을 출력 SAS 데이터셋에 저장할 수도 있습니다. 마지막으로 예측, 신뢰한계, 잔차 및 실제 값을 도표화할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
----------	--------------

사용되는 SAS 프로시저	FORECAST
---------------	----------

요구 사항 이름	프로시저 및 제품 이름
필수 SAS 제품	Base SAS, SAS/ETS
추가 SAS 제품(권장)	SAS/GRAPH

기본 예측: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
예측 변수	예측할 변수를 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다.
시간 ID 변수	행에 날짜를 할당하는 데 사용될 변수를 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다. SAS 날짜, 시간 또는 날짜/시간 출력형식을 갖는 첫 번째 변수가 자동으로 이 역할에 할당됩니다. 시간 ID 변수를 생성할 수도 있습니다.
예측 그룹화 기준	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 다른 분석 역할에 이미 할당한 변수로는 그룹화할 수 없습니다.

기본 예측: 예측 옵션 설정

예측 방법 및 간격 선택

이 옵션에 액세스하려면 선택 영역에서 **예측 옵션**을 클릭합니다.

옵션 이름	설명
예측 방법	예측 방법을 지정하려면 드롭다운 리스트에서 원하는 방법을 선택합니다. 노트 ■ 단계별 자기회귀(자기회귀 과정에 적합)는 시간 추세 모델을 계열에 적용하고 각 값과 추세 추정값 간의 차이를 가져옵니다. 그런 다음, 남은 분산에 자기회귀모형을 적용합니다. 기본값입니다. ■ 지수평활(이동평균과정에 적합)은 이전 시계열의 데이터보다 최근 데이터에 가중값을 두는 추세 모델을 적용합니다. 관측값의 가중값은 관측값이 현재 기간과 연관된 과거까지 연장되는 기간 수에 대한 기하(지수)함수입니다. ■ Winters(승법) 방법(승법 계절 과정)은 모델에 대한 파라미터를 적합시키기 위해 지수 평활과 유사한 업데이트 방정식을 사용합니다. 이 모형에는 두 개의 추세 모수와 시간에 해당하는 계절모수를 선택하는 함수를 사용합니다. ■ 계절모수를 추세에 곱하는 것이 아니라 더한다는 것만 제외하면 Winters(가법) 방법(가법계절과정에 적합)은 Winters(승법) 방법과 유사합니다. 주: 선택하는 예측 방법에 따라 이 창의 다른 컨트롤 옵션이 활성화 또는 비활성화됩니다.
예측할 간격 수	예측할 간격 수 상자에 예측할 시간 간격 수를 지정합니다. 기본값은 12입니다. 예를 들어 관측값 간의 시간 간격을 달로 지정하고 예측할 시간 간격 수를 6으로 지정하면, 입력 데이터의 끝 부분부터 6개월 동안의 예측을 얻게 됩니다.

입력 데이터 옵션 선택

이 옵션에 액세스하려면 선택 영역에서 **예측 옵션**을 클릭합니다.

옵션 이름	설명
관측값 사이의 시간 간격	입력 데이터 테이블의 행 사이의 시간 간격을 지정합니다. 기본 시간 간격은 단위 수입니다. 단위 수 를 시간 간격으로 선택하면 간격당 시간 단위도 선택할 수 있습니다. 다른 시간 간격에 대해서는 간격당 시간 단위 가 비활성화됩니다. 입력 데이터에 적합한 간격을 선택해야 하며 그렇지 않으면 분석이 실패하거나 부정확한 결과를 얻게 됩니다.
계절 순환 길이	계절 예측 방법(Winters 승법 또는 Winters 가법)에 대한 계절 순환 길이와 계절 순환당 간격을 지정합니다. 사용 가능한 계절 순환 길이는 위의 상자에서 선택한 시간 간격에 따라 결정됩니다. 기본 계절 순환 길이는 간격 수 입니다.

예측 방법 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **예측 옵션**을 클릭합니다.

도표 이름	설명
시간 추세 모델의 차수	시간 추세 모델의 차수를 지정하려면 드롭다운 리스트에서 값을 선택합니다. 상수 , 1차 및 2차 모형을 선택할 수 있습니다. 기본적으로 지수평활법은 2차 모형을, 다른 예측 방법은 1차 시간 추세 모형을 사용합니다.
신뢰수준	드롭다운 리스트에서 신뢰수준을 선택합니다. 기본값은 95% 신뢰수준입니다.
요소 평활 가중값	지수평활, Winters 승법 및 Winters 가법에 대한 평활 가중값을 지정할 수 있습니다. 지수평활 방법은 상수항에 대해서만 평활 가중값을 지정할 수 있습니다. Winters 승법 및 Winters 가법에 대해서는 상수 요소, 1차 및 2차 요소, 계절 요소에 대한 평활 가중값을 지정할 수 있습니다. 1차 및 2차 요소는 상수 평활 가중값을 지정할 때만 사용할 수 있습니다. 마찬가지로 계절 평활 가중값은 1차 및 2차 요소에 대한 가중값을 지정할 때만 사용할 수 있습니다. 기본 평활 가중값은 지정된 시간 추세 모델의 차수와 지정된 다른 평활 가중값에 따라 결정됩니다. 평활 가중값을 변경하려면 해당 체크 박스를 선택하고 필드에 새로운 값을 입력합니다.

기본 예측: 도표 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

종류 영역에 있는 체크 박스를 사용하여 실제값, 예측 데이터 및 잔차의 도표를 생성합니다.

예측 데이터 도표에는 다음과 같은 데이터를 포함할 수 있습니다.

- **실제 데이터값과 1-단계 앞의 예측값**에는 도표의 실제 데이터 값과 1단계 앞의 예측이 포함됩니다. 1단계 앞의 예측값은 미래의 특정 시간 기간에 대한 이전 시간 기간($t-1$) 동안 생성된 예측값을 말합니다.
- **예측의 신뢰한계**에는 도표의 각 예측값에 대한 신뢰한계가 포함됩니다.

모든 도표 유형에서 **표시할 실제 데이터 제한** 체크 박스 및 **표시할 간격 수** 상자를 사용하여 예측 데이터를 시작하기 전에 표시할 실제 데이터의 간격 수를 선택합니다.

주: **예측 그룹화 기준** 역할에 변수를 할당하면 **표시할 실제 데이터 제한** 체크 박스가 비활성화됩니다.

기본 예측: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

예측값이 포함된 데이터셋은 항상 저장됩니다. 영구적으로 저장되는 모수 추정값 테이블을 생성하려면 **모수 추정값** 체크 박스를 선택합니다. 추가로 포함할 통계량을 선택합니다.

주: SAS Enterprise Guide는 **옵션 창**의 **작업 > 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

예측값과 모수 추정값을 결과에 포함할지를 선택할 수 있습니다. **모수 추정값** 영역의 **모수 추정값** 체크 박스를 선택해야만 **출력 표시** 영역에서 **모수 추정값** 체크 박스를 선택할 수 있습니다.

주: 도표를 선택하지 않고 예측값 또는 추정값의 표시도 요청하지 않으면 어떤 결과도 생성되지 않습니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

상자관리도

상자관리도 정보	51
관리도: 분석 역할에 변수 할당	52
관리도: 관리한계 지정	53
관리도: 수행할 검정 선택	54
관리도: 도표 모양 옵션 설정	55
상자관리도 옵션 설정	55
추가 모양 옵션	56
관리도: 분석 결과 저장	56
제목 및 각주 지정	57
속성 보기	57

상자관리도 정보

Shewhart 관리도는 공정이 통계적 관리 상태에 있는지 여부를 결정하기 위한 그래픽 및 분석 도구입니다. 상자관리도 작업은 하위그룹 평균에 대한 평균 관리도를 생성합니다. 각 하위그룹에 대한 측정값의 상자 관리도를 평균 관리도에 추가합니다. 상자관리도는 하위그룹 분포에 대한 자세한 분석을 제공하며 표본 크기가 10개 이상인 큰 하위그룹에 사용하는 것이 좋습니다.

상자 관리도를 사용하여 출력 분포를 확인하고 프로세스가 통계적 관리 상태에 있는지 여부를 확인할 수 있습니다. 예를 들어 석유 회사가 터빈으로 물을 가열한다고 가정합니다. 가열한 물은 증기가 되어 땅속으로 주입됨으로써 석유의 점성이 낮아지고 추출이 쉬워집니다. 이러한 프로세스는 매일 20회 수행되면 물을 희망 온도까지 가열하는 데 사용되는 전력량(단위: 킬로와트)을 기록합니다. 상자 관리도를 사용하여 일별 출력 분포를 조사하고 가열 과정의 평균 레벨이 통계적 관리 내에 있는지 여부를 확인할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	SHEWHART
필수 SAS 제품	Base SAS, SAS/QC, SAS/GRAPH
추가 SAS 제품(권장)	SAS/STAT

관리도: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
공정 측정 변수	분석할 공정 변수를 지정합니다. 이 변수에는 개별 측정값이 포함됩니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
하위그룹 식별자	데이터의 하위그룹을 식별하는 변수를 지정합니다. 변수값은 입력 테이블의 관측값이 논리적 하위그룹으로 배열되는 방법을 나타냅니다. 일반적으로 값은 다음과 같습니다. ■ 하위그룹 표본의 수집 순서를 지정하는 지수 ■ 하위그룹 표본이 수집된 날짜 또는 시간 ■ 하위그룹 표본을 고유하게 식별하는 레이블 하위그룹 식별자는 필수 변수이며 이 역할에는 변수를 하나만 할당할 수 있습니다. 식별된 하위그룹으로 데이터를 정렬하려면 하위그룹별 정렬 체크 박스를 선택합니다.
하위그룹 표본 크기	하위그룹 표본 크기를 변수값으로 지정합니다. 하위그룹 표본 크기는 반드시 입력해야 하며 이 역할에는 하나의 숫자 변수만 할당할 수 있습니다. 하위그룹 표본 크기가 고정되어 있고 크기에 상수값을 지정해야 할 때 하위그룹 표본 크기로 (변수 대신) 숫자값 사용 체크 박스를 선택합니다. 하위그룹 표본 크기 상자에 사용할 값을 입력합니다. 하위 그룹 표본 크기는 프로세스 변수의 최댓값과 99999(포함) 사이의 값이어야 합니다.

역할 이름	설명
	주: 옵션 영역에서 하위그룹 표본 크기로 (변수 대신) 숫자값 사용 체크 박스를 선택하면 이 역할은 필요하지 않습니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
블록 변수	데이터를 연속된 하위그룹 블록으로 그룹화하는 변수를 최대 두 개까지 지정합니다. 이 블록은 범례에서 레이블로 표시되며 각 블록 변수는 범례에서 한 종류의 레이블이 됩니다.

관리도: 관리한계 지정

이 옵션에 액세스하려면 선택 영역에서 **관리한계**를 클릭합니다.

옵션 이름	설명
Sigma 한계	관리도에 그려지는 요약 하위그룹 통계량의 표준오차(sigma)의 배수로 관리한계의 너비를 지정합니다. 너비는 양수여야 합니다. 기본 배수는 3입니다.
계산 방법 선택	관리한계의 계산 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 활성 데이터에서 관리한계 계산 ■ 선택한 데이터에서 관리한계 계산 관리한계 또는 관리한계를 계산할 수 있는 모수를 포함하는 데이터셋을 지정하려면 찾아보기를 클릭합니다. 사용할 데이터셋을 선택하고 열기를 클릭합니다. 주: 선택한 데이터셋은 LIMITS= 데이터셋이어야 합니다. LIMITS= 데이터셋에 필요한 변수에 대한 자세한 내용은 생성 중인 그래프의 프로시저 문서를 참조하십시오. 모든 하위 그룹에 대한 요약통계량을 도표화하려면 하위그룹 표본 크기가 명목 관리한계 표본 크기와 동일한지 관계없이 모든 하위그룹에 대한 요약통계량 도표를 생성합니다. 체크 박스를 선택합니다. 거의 모든 하위 그룹의 표본 크기가 공통이고 Nominal 표본 크기 n에 대응하는(가변이 아닌)

옵션 이름	설명
	고정된 관리한계를 표시하려고 할 때 이 통계량을 도표화할 수 있습니다. 이 옵션을 선택하고 하위 그룹에 대한 표본 크기가 유의하게 다르면 관리한계 해석은 표본 크기가 n인 하위 그룹에 대해서만 의미가 있습니다. 표본 크기가 n과 같지 않은 데이터 포인트에 대해 특수 표식을 추가하려면 표본 크기에 해당하는 포인트에 대해 특수 표식을 추가하며, 표본 크기는 고정 관리한계의 명목표본 크기와 동일하지 않습니다. 체크 박스를 선택합니다. ■ 관리한계 지정 상한, 중심 및 하한 필드에 관리상한, 중심선 및 관리하한을 각각 지정합니다. 하위그룹의 표본 크기가 다양하면, 명목표본 크기를 지정해야 합니다. 주: 그룹 분석변수 역할에 변수를 할당하면 관리한계 값을 입력할 수 없습니다.

관리도: 수행할 검정 선택

이 옵션에 액세스하려면 선택 영역에서 **검정**을 클릭합니다.

주: 이 검정은 3 Sigma 한계가 아니거나 하위그룹 표본 크기에 따라 달라지는 관리한계에는 적용되지 않습니다.

옵션 이름	설명
검정 선택	체크 박스를 사용하여 런 검정, 패턴 검정 및 Western Electric 규칙이라고도 하는 특수 원인에 대한 검정을 하나 이상 요청합니다. 이러한 검정은 주관리도에 그려진 포인트에서 특정 비임의 패턴을 찾아냅니다. 이러한 비임의 패턴의 발생을 신호라고 합니다. 이러한 검정을 통해 특수한 변동의 원인이 존재하는지를 알 수 있습니다. 각 검정을 선택하면 해당 검정에 대한 설명이 설명 상자에 나타납니다.
레이블	텍스트 상자를 사용하여 특수 원인에 대한 검정이 양성으로 나타난 포인트에 대한 레이블을 입력합니다. 레이블의 문자 수는 16자를 초과할 수 없습니다. 각 레이블은 검정에서 신호가 발생한 모든 점에 표시됩니다.
레이블 색상	■ 드롭다운 리스트를 사용하여 지정한 레이블의 색상을 선택합니다.

옵션 이름	설명
	■ 영역 선 표시 체크 박스를 사용하여 주관리도에 A, B, C 영역을 구분하는 선을 표시합니다. 이 영역은 관리한계 사이의 간격을 여섯 개의 동일한 간격을 지닌 하위 간격으로 분할함으로써 구성됩니다. ■ 3 sigma 한계 무시 체크 박스를 사용하여 기본 배수 3 이외의 sigma 관리한계를 지정했을 때 특수한 원인을 찾는 검정을 수행할 수 있습니다. sigma 한계를 변경하려면 선택 영역에서 관리한계를 클릭합니다. ■ 모든 점을 검정의 시작점으로 사용 체크 박스를 사용하여 중복되는 점 패턴에 특수 원인 검정을 적용합니다.

관리도: 도표 모양 옵션 설정

상자관리도 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **옵션**을 클릭합니다.

옵션 이름	설명
관리한계 내부, 관리 한계 및 프레임	이러한 그래프 요소들의 색상을 지정합니다. ■ 관리한계 내부 - 관리상한 및 관리하한 내부 색상을 지정합니다. ■ 관리한계 - 관리한계 및 중심선의 색상을 지정합니다. 이들 선에 대한 레이블 색상도 지정할 수 있습니다. ■ 프레임 - 축과 프레임으로 둘러싸인 직사각형을 채울 색상을 지정합니다. 주: 관리한계 표시를 생략하면 상자관리도에서 관리한계 내부 및 관리한계 옵션을 사용할 수 없습니다. 관리한계를 활성화하려면 관리한계 표시 생략 체크 박스의 선택을 해제합니다.
상자 색상	상자의 외곽선 색상 및 채우기 색상을 지정합니다.
관리한계 표시 생략	하위그룹 표본크기에 대한 기본 범례를 숨깁니다. 이 옵션은 관리한계 표본크기가 자동으로 그래프의 오른쪽 상단에 나타나기 때문에 하위그룹 표본크기가 일정하고 관리한계 표본크기와 동일할 때 유용합니다.
하위그룹 크기에 따라 상자 너비 변경	하위그룹 크기에 맞게 상자 도표를 변경하려면 이 옵션을 지정합니다.

옵션 이름	설명
상자에 V자형 홈 만들기	V자형의 상자 도표를 생성합니다. V자형의 끝점은 중위수 + 및 - 수량 $1.58(IQR/\sqrt{n})$ 에 있습니다. 여기서 IQR은 사분위범위이고 $\sqrt{n}$ 는 하위 그룹 표본 크기의 제곱근입니다. 두 상자 도표의 중위수(중심선)는 대응하는 V자형이 중첩되지 않으면 약 0.05 레벨에서 유의하게 다릅니다.

추가 모양 옵션

일부 모양 옵션은 여러 관리도 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- “축 옵션 설정” (605페이지)
- “참조선 추가” (605페이지)
- “블록 변수 옵션 설정” (608페이지)

관리도: 분석 결과 저장

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

분석 결과를 저장하기 위해 다양한 통계량을 포함하는 출력 데이터셋을 생성할 수 있습니다. 영구적으로 저장하려는 출력 데이터의 체크 박스를 선택합니다.

역할 이름	설명
하위그룹 통계량과 관리 한계 데이터셋	관리도에 그려지는 정보를 포함하는 출력 데이터셋을 생성합니다. 여기에는 하위그룹 식별자값 및 해당 요약통계량과 관리한계가 포함됩니다. 주: 이 옵션은 OUTTABLE= 데이터셋을 생성합니다. SHEWHART 프로시저에서 TABLE= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.
하위그룹 통계량 출력 데 이터셋	하위그룹 요약통계량을 포함하는 출력 데이터셋을 생성합니다. 주: 이 옵션은 OUTHISTORY= 데이터셋을 생성합니다. SHEWHART 프로시저에서 HISTORY= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.

역할 이름	설명
관리한계 출력 데이터셋	관리한계를 저장하는 출력 데이터셋을 생성합니다. 추후에 이 작업 사용 시 관리한계의 계산 방법을 선택할 때 이 데이터셋을 사용할 수 있습니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버

- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

상자 도표

상자 도표 작업 정보	59
상자 도표: 도표 유형 선택	60
상자 도표: 분석 역할에 변수 할당	60
상자 도표: 모양 옵션 설정	61
상자 도표 옵션 설정	61
Hi-Lo 도표 옵션 설정	62
추가 모양 옵션	62
제목 및 각주 지정	63
속성 보기	63

상자 도표 작업 정보

상자 도표 작업은 그래프 변수의 여러 값에 대해 일부 숫자 변수의 다중 요약통계량을 표시하는 상자 도표, Hi-Lo 도표, Hi-Lo-Close 도표를 생성합니다. 예를 들어 상자 도표는 변수의 범위, 사분위범위, 평균 등을 보여 줍니다. Hi-Lo-Close 도표는 여러 날짜의 주식 고가, 저가 및 종가를 표시할 수 있습니다.

자세한 내용은 [“그래픽 출력형식 선택” \(5페이지\)](#)에서 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GPLOT
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

상자 도표: 도표 유형 선택

이 옵션에 액세스하려면 선택 영역에서 **상자 도표**를 클릭합니다.

옵션 이름	설명
상자 도표	표본의 25번째 및 75번째 백분위수에 위치한 상자의 하단 및 상단 모서리입니다. 중앙의 수평선은 50번째 백분위수(중위선)에 그려집니다. 수직선, 즉 수염은 상자에서 1.5 사분위 범위 안의 극 값까지 그려집니다. 이 범위 외의 값은 기호로 표시됩니다.
Hi-Lo 도표	각 X 값에 대한 Y의 low(최소), high(최대) 값을 실수직선으로 연결하도록 지정합니다. 눈금 표식은 각 X 값에 대한 Y 값 평균을 나타냅니다. 이 유형의 도표는 주식 시장 데이터에 유용합니다.
Hi-Lo-Close 도표	각 X 값에 대한 Y의 low(최소), high(최대) 값을 실수직선으로 연결하도록 지정합니다. 눈금 표식은 close 값을 나타냅니다. 이 유형의 도표는 주식 시장 데이터에 유용합니다.

상자 도표: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
수평	이 역할에 할당하는 칼럼은 그래프의 가로축 또는 X축 변수입니다.
수직	이 역할에 할당하는 칼럼은 그래프의 세로축 또는 Y축 변수입니다.
세로(오른쪽)	이 역할에 할당하는 칼럼은 그래프의 오른쪽 면에 대한 세로축 또는 Y축 변수입니다. 이 변수는 X축 변수에 대응하여 그려지므로 그래프에 별도의 도표가 생성됩니다.

역할 이름	설명
그룹 분석 칼럼	이 역할에 할당하는 칼럼의 값에 따라 생성되는 그래프의 수가 결정됩니다. 각 고윳값에 대해 별도의 그래프가 생성됩니다. 생성된 모든 그래프에 대해 각 축에 대한 배율을 동일하게 하려면 모든 그래프에 대해 축 길이 같음 체크 박스를 선택합니다. 기본적으로, 각 축에 대한 값의 범위는 도표화된 데이터에서 최소 및 최대값에 기반합니다. 이 옵션을 선택하면 각 축의 범위는 모든 그래프에 대해 동일합니다.

상자 도표: 모양 옵션 설정

상자 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **상자 도표**를 클릭합니다.

주: 이 옵션은 상자 도표를 선택할 때만 사용할 수 있습니다.

역할 이름	설명
중위수 연결 선 및 수염 위/아래에 선 그리기	중위수를 연결하는 선을 그리거나 수염의 위와 아래에 선을 그릴 수도 있습니다.
상자 및 이상치 색상 지정	채우기 및 윤곽선 색상 조합을 직접 선택하려면 색상 지정 체크 박스를 선택하고 상자 내부 및 이상치 와 윤곽선 드롭다운 리스트에서 색상을 선택합니다. 외곽선 색상에 따라 Hi-Lo 도표와 상자 도표의 상자 외곽선 및 수염 선 색상이 지정됩니다.
축 범위 밖의 값 포함	축 범위 밖의 값이 보간 계산에 포함되도록 지정하려면 축 범위 밖의 값 포함 체크 박스를 선택합니다. 이 옵션을 선택하면 축 범위 밖에 있는 값이 보간 계산에는 포함되지만 도표에서는 제외됩니다. 기본적으로 이 체크 박스는 선택되지 않으며 축 범위 밖의 값이 보간 계산에서 생략됩니다.
수염 길이 지정	수염에 대한 길이 백분위수를 선택할 수 있습니다. 다음과 같은 옵션을 사용할 수 있습니다. ■ 사분위범위의 +/- 1.5배 - 상자에서 사분위범위의 +/-1.5배 내에 있는 극 값까지 수직선이나 수염을 그립니다. ■ 최솟값/최댓값 - 극단최댓값에서 극단최솟값까지 수염을 그립니다.

역할 이름	설명
	주: 이 옵션을 선택하면 수염 길이 백분위수 범위 외 값(이상치)은 기호로 표시되지 않습니다. 길이 백분위수의 다른 옵션을 선택하면 이상치가 기호로 표시됩니다. ■ 1번째/99번째 백분위수 - 1번째 백분위수값에서 99번째 백분위수값까지 수염을 그립니다. ■ 5번째/95번째 백분위수 - 5번째 백분위수값에서 95번째 백분위수값까지 수염을 그립니다. ■ 10번째/90번째 백분위수 - 10번째 백분위수값에서 90번째 백분위수값까지 수염을 그립니다. ■ 25번째/75번째 백분위수 - 25번째 백분위수값에서 75번째 백분위수값까지 수염을 그립니다.
이상치 사용자 정의	사용할 글꼴 유형, 기호 및 높이를 지정할 수 있습니다.

Hi-Lo 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **Hi-Lo**를 클릭합니다.

주: 이 옵션은 Hi-Lo 도표 또는 Hi-Lo-Close 도표를 선택할 때만 사용할 수 있습니다.

역할 이름	설명
중위수 연결선 및 수염 위/아래에 선 그리기	중위수를 연결하는 선을 그리거나 수염의 위와 아래에 선을 그릴 수도 있습니다.
Y의 하한 값과 상한 값을 상자로 연결	Y의 최솟값과 최댓값 사이에 선 대신 막대를 그리도록 선택할 수도 있습니다.
평균 대신 Close 값에 눈금 표시 그리기	평균값이 아닌 Close값에 눈금 표시를 그리도록 선택할 수 있습니다. 이 옵션을 사용하려면 각 X값에 대해 정확히 3개의 Y값(High, Low, Close)이 있어야 합니다.

추가 모양 옵션

일부 모양 옵션은 여러 그래프 작업에 사용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- “축 옵션 설정” (593페이지)
- “참조선 추가” (595페이지)
- “그래프 영역 사용자 정의” (597페이지)

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

버블 도표

버블 도표 작업 정보	65
버블 도표: 분석 역할에 변수 할당	66
버블 도표: 모양 옵션 설정	67
버블 도표 옵션 설정	67
추가 모양 옵션	67
제목 및 각주 지정	67
속성 보기	68

버블 도표 작업 정보

버블 도표 작업으로 세로축 및 가로축에 그려진 데이터 포인트에 다양한 비율의 원을 그려서 단순 버블 도표를 생성합니다. 변수 중 두 개는 데이터 포인트의 위치를 결정하며, 세 번째 변수는 원의 크기를 제어합니다.

버블 도표를 사용하여 임금 정보를 비교할 수 있습니다. 예를 들어 각 버블이 항공 우주, 화학, 도시, 전기, 기계 등의 엔지니어 범주를 나타낸다고 가정합니다. 버블의 위치는 해당 버블이 나타내는 범주의 평균 임금에 의해 결정됩니다. 버블의 크기는 데이터의 총 엔지니어 수와 관련된 범주의 엔지니어 수를 나타냅니다.

자세한 내용은 [“그래픽 출력형식 선택” \(5페이지\)](#)에서 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GPLOT
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

버블 도표: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
수평	이 역할에 할당하는 칼럼은 그래프의 가로축 또는 X축 변수입니다.
수직	이 역할에 할당하는 칼럼은 그래프의 세로축 또는 Y축 변수입니다.
세로(오른쪽)	이 역할에 할당하는 칼럼은 그래프의 오른쪽 면에 대한 세로축 또는 Y축 변수입니다. 이 변수는 X축 변수에 대응하여 그려지므로 그래프에 별도의 도표가 생성됩니다.
버블 크기	이 역할에 할당한 칼럼은 가로 및 세로 변수의 도표에 대해 상대적인 버블 크기를 결정합니다. 주: 오른쪽 세로 역할에 할당한 변수를 가졌지만 버블2 크기 역할에 변수를 할당하지 않으면, 이 역할이 오른쪽 수평 및 수직 변수의 도표에 대해 상대적인 버블 크기도 결정합니다.
그룹 분석 칼럼	이 역할에 할당하는 칼럼의 값에 따라 생성되는 그래프의 수가 결정됩니다. 각 고유티 값에 대해 별도의 그래프가 생성됩니다. 생성된 모든 그래프에 대해 각 축에 대한 배율을 동일하게 하려면 모든 그래프에 대해 축 길이 같음 체크 박스를 선택합니다. 기본적으로, 각 축에 대한 값의 범위는 도표화된 데이터에서 최소 및 최대값에 기반합니다. 이 옵션을 선택하면 각 축의 범위는 모든 그래프에 대해 동일합니다.
버블2 크기	이 역할에 할당한 칼럼은 오른쪽 가로 및 세로 변수의 도표에 대해 상대적인 버블 크기를 결정합니다. 주: 오른쪽 세로 역할에 변수를 할당할 때만 이 옵션을 사용할 수 있습니다.

버블 도표: 모양 옵션 설정

버블 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **버블 옵션**을 클릭합니다.

드롭다운 리스트에서 버블 색상을 선택합니다.

버블에 레이블을 지정할 수 있습니다. 레이블은 가로 및 세로 변수값 조합에 대해 **버블 크기** 역할에 지정된 칼럼의 값입니다. 이 변수에 출력형식이 있으면 서식화된 값이 사용됩니다. 기본적으로 버블에는 레이블이 지정되지 않습니다.

슬라이더를 사용하여 모든 버블의 크기를 확대하거나 축소할 수 있습니다. 이 값을 조정하려면 먼저 출력을 보고 확인하는 것이 좋습니다.

주: 오른쪽 세로 역할에 변수를 할당하면 가로 및 오른쪽 세로 변수 조합을 위한 버블에 버블 2 옵션을 설정할 수 있습니다.

추가 모양 옵션

일부 옵션은 여러 그래프 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- ["축 옵션 설정" \(593페이지\)](#)
- ["참조선 추가" \(595페이지\)](#)
- ["그래프 영역 사용자 정의" \(597페이지\)](#)

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.

- 2 기본 텍스트 사용 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

.....

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

.....

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

.....

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

.....

정준상관분석

정준상관분석 작업 정보	69
정준상관분석: 분석 역할에 변수 할당	70
정준상관분석: 통계량 옵션 설정	71
정준상관분석: 도표 생성	72
정준상관분석: 결과 옵션 설정	72
제목 및 각주 지정	73
속성 보기	73

정준상관분석 작업 정보

정준상관분석은 X 변수 집합의 선형 결합과 Y 변수 집합의 선형 결합 간의 관계를 검사합니다. 각 정준상관분석 및 모든 소형 정준상관분석을 검증하는 일련의 가설은 모집단이 0입니다. 단순 정준상관분석과 다중 정준상관분석은 한쪽 혹은 양쪽 집합 모두에 단일 변수가 포함된 특별한 사례의 정준상관분석입니다.

정준상관분석 작업을 사용하여 직업 특성 집합과 직업 만족도 집합 간의 일치 정도를 결정하는 데 사용할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	CANCORR
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

정준상관분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
Set 1 변수	이 역할에 할당된 변수를 사용하여 분석이 수행됩니다. 이 역할에는 변수를 하나 이상 할당해야 합니다. Set 1 정준변량 영역을 사용하여 Set 1 변수의 접두어와 레이블을 지정합니다.
Set 2 변수	이 역할에 할당된 변수를 사용하여 분석이 수행됩니다. 이 역할에는 변수를 하나 이상 할당해야 합니다. Set 2 정준변량 영역을 사용하여 Set 2 변수의 접두어와 레이블을 지정합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: Set 1 변수 또는 Set 2 변수에서 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
편상관변수	편상관계수를 기준으로 한 정준상관분석에 사용할 변수를 지정합니다. 이 역할의 변수는 Set 1 변수와 Set 2 변수의 편상관변수입니다.
상대 가중값 변수	가중 곱적률상관계수를 계산하기 위한 각 행의 변수값을 지정합니다. 행은 이 역할의 변수값이 0보다 클 때만 사용됩니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.

정준상관분석: 통계량 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **통계량**을 클릭합니다.

옵션 이름	설명
수행할 회귀분 석	수행할 회귀분석을 설정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 회귀분석 실행 안 함은 회귀분석을 수행하지 않습니다. ■ 종속: Set 2, 설명: Set 1 및 종속: Set 1, 설명: Set 2는 Set 1 변수가 종속변수가 되고 다른 분석에서는 회귀변수가 되도록 두 개의 다중 회귀분석을 요청합니다. ■ 종속: Set 2, 설명: Set 1은 Set 2 변수를 종속변수로 사용하고 Set 1 변수를 회귀변수로 사용하여 다중 회귀분석을 요청합니다. ■ 종속: Set 1, 설명: Set 2는 Set 1 변수를 종속변수로 사용하고 Set 2 변수를 회귀변수로 사용하여 다중 회귀분석을 요청합니다.
회귀 통계량	다음 통계량 중에서 하나 이상을 계산합니다. ■ 회귀계수는 회귀분석에서 원시 회귀계수를 생성합니다. ■ 표준화회귀계수 - 표준화회귀계수를 생성합니다. ■ 계수의 표준오차 - 회귀계수의 표준오차를 생성합니다. ■ t 통계량 및 확률 > t 는 회귀계수에 대한 t 통계량 및 t 통계량에 대한 확률 레벨을 생성합니다. ■ 다중상관계수의 제곱합 - 회귀분석에 대한 다중상관계수의 제곱합과 F 검정을 생성합니다.
상관통 계량	다음 통계량 중에서 하나 이상을 계산합니다. ■ 회귀계수의 상관계수 - 회귀계수 추정값 간의 상관계수를 생성합니다. ■ 편상관계수 - 각 종속변수와 회귀변수에서 다른 모든 회귀변수의 영향을 제외한 회귀변수와 종속변수 간의 편상관계수를 생성합니다. ■ 편상관계수의 제곱합 - 각 종속변수와 회귀변수에서 다른 모든 회귀변수의 영향을 제외한 회귀변수와 종속변수 간의 편상관계수의 제곱합을 생성합니다. ■ 준-편상관계수 - 각 회귀변수에서 다른 모든 회귀변수의 영향을 제외한 회귀변수와 종속변수 간의 준-편상관계수를 생성합니다. ■ 준-편상관계수의 제곱합 - 각 회귀변수에서 다른 모든 회귀변수의 영향을 제외한 회귀변수와 종속변수 간의 준-편상관계수의 제곱합을 생성합니다.

정준상관분석: 도표 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

주: 과잉통계량과 계수를 표시하는 정준변수의 개수가 0개이면, 도표 옵션을 사용할 수 없습니다. 자세한 내용은 결과의 정준변수 개수 옵션을 참조하십시오.

도표를 표시하려면 **정준변수의 도표 표시** 체크 박스를 선택합니다. **시작 변수** 및 **마지막 변수** 상자에서 도표로 그릴 변수를 입력합니다. 다음은 몇 가지 예입니다.

- 첫 번째 정준변수만 도표에 그리려면(V1 대 W1) 두 입력 필드에 **1**을 입력합니다.
- 첫 번째, 두 번째 및 세 번째 정준변수를 도표에 그리려면(V1 대 W1, V2 대 W2, V3 대 W3) 첫 번째 입력 필드에 **1**을, 두 번째 입력 필드에 **3**을 입력합니다.

정준상관분석: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
출력 데이터셋에 결과 저장	결과를 출력 데이터셋에 저장합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 상관계수 저장 - 정준상관분석 및 계수, 다중회귀 통계량 등의 다양한 통계량을 포함한 데이터셋을 생성합니다. ■ 점수 저장 - 모든 원 데이터와 정준변수의 점수를 포함한 데이터셋을 생성합니다. 새로운 변수의 개수는 정준변수 개수 상자에 지정한 값의 두 배입니다. 주: SAS Enterprise Guide는 옵션 창의 작업> 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
결과 표시	출력을 표시합니다. 다음 옵션을 지정할 수 있습니다. ■ 정준변수 개수 - 결과의 전체 출력이 포함될 정준변수의 개수를 지정합니다. ■ 정준과잉분석 포함 - 표준화되지 않은 원 분산 및 표준화된 분산과 각 변수 집합으로 이루어진 분산의 누적 비율을 포함한 정준과잉분석을 수행하며 해당 정준변수 및 역정준변수에 의해 설명됩니다.

옵션 이름	설명
	■ 점수 포함 - 모든 원 데이터와 정준변수의 점수를 포함한 출력 데이터셋을 표시합니다. 추후에도 사용할 수 있도록 영구적인 데이터셋을 생성하려면 점수 저장 체크 박스를 선택합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트

- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

결점수관리도

결점수관리도 작업 정보	75
결점수관리도: 분석 역할에 변수 할당	76
관리도: 관리한계 지정	77
관리도: 수행할 검정 선택	78
관리도: 도표 모양 옵션 설정	79
관리도: 분석 결과 저장	79
제목 및 각주 지정	80
속성 보기	80

결점수관리도 작업 정보

Shewhart 관리도는 공정이 통계적 관리 상태에 있는지 여부를 결정하기 위한 그래픽 및 분석 도구입니다. 결점수관리도는 하위그룹 표본의 결점 개수에 대한 결점수관리도를 생성합니다.

결점수관리도를 사용하여 새로운 제품에서 발견된 결점수를 모니터링할 수 있습니다. 예를 들어 자동차 회사에서 새로운 트럭에 대한 페인트 결점수를 모니터링하고자 한다고 가정합니다. 같은 모델에서 20대의 트럭을 검사하였고 트럭당 페인트 결점수를 기록하였습니다. 결점수관리도에서의 각 포인트는 트럭에 대한 결점수를 나타냅니다. 관리상한 및 하한을 표시하여 포인트가 이 한계의 위 또는 아래에 위치하는지 확인할 수 있습니다. 지정된 관리한계를 넘는 포인트가 없으면 페인팅 프로세스가 통계적 관리 상태에 있다고 할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	SHEWHART
필수 SAS 제품	Base SAS, SAS/QC, SAS/GRAPH

요구 사항 이름	프로시저 및 제품 이름
추가 SAS 제품(권장)	SAS/STAT

결점수관리도: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
결점 수	결점 수를 갖는 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
하위그룹 식별자	데이터의 하위그룹을 식별하는 변수를 지정합니다. 변수값은 입력 테이블의 관측값이 논리적 하위그룹으로 배열되는 방법을 나타냅니다. 일반적으로 값은 다음과 같습니다. ■ 하위그룹 표본의 수집 순서를 지정하는 지수 ■ 하위그룹 표본이 수집된 날짜 또는 시간 ■ 하위그룹 표본을 고유하게 식별하는 레이블 하위그룹 식별자는 필수 변수이며 이 역할에는 변수를 하나만 할당할 수 있습니다. 식별된 하위그룹으로 데이터를 정렬하려면 하위그룹별 정렬 체크 박스를 선택합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
블록 변수	데이터를 연속된 하위그룹 블록으로 그룹화하는 변수를 최대 두 개까지 지정합니다. 이 블록은 범례에서 레이블로 표시되며 각 블록 변수는 범례에서 한 종류의 레이블이 됩니다.

관리도: 관리한계 지정

이 옵션에 액세스하려면 선택 영역에서 **관리한계**를 클릭합니다.

옵션 이름	설명
Sigma 한계	관리도에 그려지는 요약 하위그룹 통계량의 표준오차(sigma)의 배수로 관리한계의 너비를 지정합니다. 너비는 양수여야 합니다. 기본 배수는 3 입니다.
계산 방법 선택	관리한계의 계산 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 활성 데이터에서 관리한계 계산 ■ 선택한 데이터에서 관리한계 계산 관리한계 또는 관리한계를 계산할 수 있는 모수를 포함하는 데이터셋을 지정하려면 찾아보기를 클릭합니다. 사용할 데이터셋을 선택하고 열기를 클릭합니다. 주: 선택한 데이터셋은 LIMITS= 데이터셋이어야 합니다. LIMITS= 데이터셋에 필요한 변수에 대한 자세한 내용은 생성 중인 그래프의 프로시저 문서를 참조하십시오. 모든 하위 그룹에 대한 요약통계량을 도표화하려면 하위 그룹 표본 크기가 명목 관리한계 표본 크기와 동일한지 관계없이 모든 하위그룹에 대한 요약통계량 도표를 생성합니다. 체크 박스를 선택합니다. 거의 모든 하위 그룹의 표본 크기가 공통이고 Nominal 표본 크기 <i>n</i>에 대응하는(가변이 아닌) 고정된 관리한계를 표시하려고 할 때 이 통계량을 도표화할 수 있습니다. 이 옵션을 선택하고 하위 그룹에 대한 표본 크기가 유의하게 다르면 관리한계 해석은 표본 크기가 <i>n</i>인 하위 그룹에 대해서만 의미가 있습니다. 표본 크기가 <i>n</i>과 같지 않은 데이터 포인트에 대해 특수 표식을 추가하려면 표본 크기에 해당하는 포인트에 대해 특수 표식을 추가하며, 표본 크기는 고정 관리한계의 명목표본 크기와 동일하지 않습니다. 체크 박스를 선택합니다. ■ 관리한계 지정 상한, 중심 및 하한 필드에 관리상한, 중심선 및 관리하한을 각각 지정합니다. 하위그룹의 표본 크기가 다양하면, 명목표본 크기를 지정해야 합니다. 주: 그룹 분석변수 역할에 변수를 할당하면 관리한계 값을 입력할 수 없습니다.

관리도: 수행할 검정 선택

이 옵션에 액세스하려면 선택 영역에서 **검정**을 클릭합니다.

주: 이 검정은 3 Sigma 한계가 아니거나 하위그룹 표본 크기에 따라 달라지는 관리한계에는 적용되지 않습니다.

옵션 이름	설명
검정 선택	체크 박스를 사용하여 런 검정, 패턴 검정 및 Western Electric 규칙이라고도 하는 특수 원인에 대한 검정을 하나 이상 요청합니다. 이러한 검정은 주관리도에 그려진 포인트에서 특정 비임의 패턴을 찾아냅니다. 이러한 비임의 패턴의 발생을 신호라고 합니다. 이러한 검정을 통해 특수한 변동의 원인이 존재하는지를 알 수 있습니다. 각 검정을 선택하면 해당 검정에 대한 설명이 설명 상자에 나타납니다.
레이블	텍스트 상자를 사용하여 특수 원인에 대한 검정이 양성으로 나타난 포인트에 대한 레이블을 입력합니다. 레이블의 문자 수는 16자를 초과할 수 없습니다. 각 레이블은 검정에서 신호가 발생한 모든 점에 표시됩니다.
레이블 색상	■ 드롭다운 리스트를 사용하여 지정한 레이블의 색상을 선택합니다. ■ 영역 선 표시 체크 박스를 사용하여 주관리도에 A, B, C 영역을 구분하는 선을 표시합니다. 이 영역은 관리한계 사이의 간격을 여섯 개의 동일한 간격을 지닌 하위 간격으로 분할함으로써 구성됩니다. ■ 3 sigma 한계 무시 체크 박스를 사용하여 기본 배수 3 이외의 sigma 관리한계를 지정했을 때 특수한 원인을 찾는 검정을 수행할 수 있습니다. sigma 한계를 변경하려면 선택 영역에서 관리한계를 클릭합니다. ■ 모든 점을 검정의 시작점으로 사용 체크 박스를 사용하여 중복되는 점 패턴에 특수 원인 검정을 적용합니다.

관리도: 도표 모양 옵션 설정

일부 모양 옵션은 여러 관리도 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- “축 옵션 설정” (605페이지)
- “참조선 추가” (605페이지)
- “기타 도표 옵션 설정” (606페이지)
- “블록 변수 옵션 설정” (608페이지)

관리도: 분석 결과 저장

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

분석 결과를 저장하기 위해 다양한 통계량을 포함하는 출력 데이터셋을 생성할 수 있습니다. 영구적으로 저장하려는 출력 데이터의 체크 박스를 선택합니다.

역할 이름	설명
하위그룹 통계량과 관리한계 데이터셋	관리도에 그려지는 정보를 포함하는 출력 데이터셋을 생성합니다. 여기에는 하위그룹 식별자값 및 해당 요약통계량과 관리한계가 포함됩니다. 주: 이 옵션은 OUTTABLE= 데이터셋을 생성합니다. SHEWHART 프로시저에서 TABLE= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.
하위그룹 통계량 출력 데이터셋	하위그룹 요약통계량을 포함하는 출력 데이터셋을 생성합니다. 주: 이 옵션은 OUTHISTORY= 데이터셋을 생성합니다. SHEWHART 프로시저에서 HISTORY= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.
관리한계 출력 데이터셋	관리한계를 저장하는 출력 데이터셋을 생성합니다. 추후에 이 작업 사용 시 관리한계의 계산 방법을 선택할 때 이 데이터셋을 사용할 수 있습니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식

■ 결과의 그래픽 출력에 대한 출력형식

.....
주: 기존 작업의 레이블을 변경하는 경우 속성 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.
.....

CDF 도표

CDF 도표 작업 정보	83
공정능력분석: 분석 역할에 변수 할당	84
CDF 도표: 분포 선택	85
CDF 도표: 분포 사용자 정의	86
공정능력분석: 도표 모양 옵션 설정	86
축 옵션 설정.....	86
참조선 추가.....	87
인셋 추가.....	87
기타 도표 옵션 설정.....	87
규격 한계 옵션 설정.....	89
공정능력분석: 분석 유형 지정	89
제목 및 각주 지정	90
속성 보기	90

CDF 도표 작업 정보

공정능력분석은 통계적 관리 상태 공정의 출력분포를 해당 규격한계와 비교하여 규격과 일치하는지를 파악합니다. 공정능력분석에는 CDF 도표가 사용되어 변수의 관측된 누적분포 함수의 도표를 그립니다. CDF는 경험적 누적분포함수(ECDF)라고도 하며 증가 계단함수입니다.

CDF 도표를 사용하여 규격 하한보다 작거나 규격 상한보다 큰 관측값의 백분율을 확인할 수 있습니다. 예를 들어 Fiber-Optic 코드를 생산하는 회사가 코드의 절단 강도에 관심이 있다고 가정합니다. 예를 들어 광섬유 코드를 생산하는 회사에서 코드의 강도에 관심이 있다고 가정하면 50개의 강도(평당 인치당 파운드) 관측값을 포함한 데이터셋을 생성합니다. CDF 도표 작업은 데이터를 사용하여 규격하한에 있는 표본의 관측값 백분율을 표시하는 CDF 도표를 생성합니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	CAPABILITY
필수 SAS 제품	Base SAS, SAS/QC, SAS/GRAPH
추가 SAS 제품(권장)	SAS/STAT

공정능력분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분석변수	분석할 숫자 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다. P-P 도표를 제외한 모든 도표에서는 변수 이름의 규격한계 영역에서 각 분석변수에 대한 상한, 목표 값 및 하한을 지정할 수 있습니다. 이 한계를 지정하면 정규성 검정을 보여 주는 테이블, 규격한계를 보여 주는 테이블 및 공정능력지수를 보여 주는 테이블이 자동으로 출력에 포함됩니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n은 해당 행의 빈도변수값입니다.
상대 가중값 변수	상대 가중값으로 사용할 변수를 지정합니다. 이 역할에 변수를 할당하면 각 통계량의 변수값이 가중 통계량 계산에 사용됩니다. 이 역할에는 변수를 하나만 할당할 수 있습니다. 주: 히스토그램, 확률 도표, P-P 도표, Q-Q 도표 또는 CDF 도표를 생성할 때는 상대 가중값 변수 역할에 변수를 할당하지 않는 것이 좋습니다. 이 역할에 변수를 할당했으면 출력을 해석할 때 주의해야 합니다.

역할 이름	설명
분류변수 (히스토그램 작업에 서만 사용 가능)	분류변수로 사용할 하나 이상의 변수를 지정합니다. 이 역할에 변수를 할당하면 전체 분류 레벨에서 분석변수의 분포를 비교할 수 있는 비교 히스토그램이 생성됩니다. 주: 비교 확률 도표 또는 비교 Q-Q 도표를 생성하려면 분포분석 작업을 사용하십시오.

CDF 도표: 분포 선택

이 옵션에 액세스하려면 선택 영역의 분포 머리글 아래에서 **요약**을 클릭합니다.

CDF 도표: 분포를 선택한 후 해당 분포의 모수값 및 선 속성을 지정하려면 선택 영역에서 분포 이름을 선택합니다.

주: 여러 분포를 선택하면 각 분포에 대해 별도의 CDF 도표가 생성됩니다.

옵션 이름	설명
분포	체크 박스를 사용하여 다음 추정된 밀도 곡선을 하나 이상 선택합니다. ■ 정규 - 평균(μ)과 표준편차(σ)를 가지는 정규밀도함수를 적합시킵니다. ■ 로그 정규 - 척도모수(ζ), 임계모수(θ) 및 형상모수(σ)를 가지는 로그 정규 밀도함수를 적합시킵니다. ■ 지수 - 척도모수(σ) 및 임계모수(θ)를 가지는 지수밀도함수를 적합시킵니다. ■ 와이블 - 척도모수(σ), 임계모수(θ) 및 형상모수(c)를 가지는 3-모수 와이블 밀도함수를 적합시킵니다. ■ 베타 - 척도모수(σ), 임계모수(θ) 및 형상모수(α 및 β)를 가지는 베타 밀도함수를 적합시킵니다. ■ 감마 - 척도모수(σ), 임계모수(θ) 및 형상모수(α)를 가지는 감마 밀도함수를 적합시킵니다.
그래프 스타일	출력에 사용할 그래프 스타일을 선택합니다. 기존 SAS 그래프나 ODS 통계 그래프 중에 선택할 수 있습니다. ODS 그래프에 대한 자세한 내용은 SAS Output Delivery System: User's Guide를 참조하십시오.

CDF 도표: 분포 사용자 정의

이 옵션에 액세스하려면 선택 영역의 분포 머리글 아래에서 사용자 정의할 분포 이름을 클릭합니다.

주: 이러한 분포 옵션은 **요약** 페이지에서 분포를 선택했거나 해당 분포 페이지에서 **분포 이름** 체크 박스를 선택할 때만 사용할 수 있습니다.

각 분석변수 또는 모든 분석변수에 대해 분포 모수를 지정할 수 있습니다. 기본적으로 작업에서는 모든 분석변수의 모수 값을 추정합니다.

각 분석변수에 대한 모수를 지정하는 방법:

- 1 **모든 변수에 분포 적용** 체크 박스의 선택을 해제합니다.
- 2 **분석변수** 상자에서 변수에 해당되는 체크 박스를 선택합니다.
- 3 모수값을 지정합니다.

주: 추정 모수값을 생성하려면 **추정값 사용**을 클릭합니다.

공정능력분석: 도표 모양 옵션 설정

축 옵션 설정

이 옵션에 액세스하려면 선택 영역의 **모양** ⇨ **축** 머리글 아래에서 **축**을 클릭합니다.

다음 옵션을 설정할 수 있습니다.

- 축 색상
- 눈금 표시 색상
- 가로축 및 세로축 레이블

참조선 추가

이 옵션에 액세스하려면 선택 영역의 **모양** ⇨ **축** 머리글 아래에서 **가로** 및 **세로**를 클릭합니다.

참조선을 사용하려면 **참조선 사용**을 선택합니다. 참조선의 출력형식을 지정하고 각 참조선의 위치를 지정할 수 있습니다.

인셋 추가

이 옵션에 액세스하려면 선택 영역의 **모양** 머리글 아래에서 **인셋**을 클릭합니다.

요약통계량 상자 또는 테이블(인셋)을 그래프에 직접 추가하여 도표를 보기 좋게 만들 수 있습니다. 비교 히스토그램을 생성하면 각각의 작은 히스토그램에 인셋이 표시됩니다.

인셋 포함을 선택하여 도표에 요약통계량을 추가합니다. 인셋에 포함할 통계량을 선택합니다.

주: 공정능력지수를 계산하기 위한 통계량은 **분석변수** 역할에 할당한 변수에 대한 규격한계를 생성할 때만 사용할 수 있습니다. 히스토그램, CDF 도표, 확률 도표 및 Q-Q 도표에 대한 규격한계를 생성할 수 있습니다. 공정능력지수 계산에 대한 자세한 내용은 CAPABILITY 프로시저에 대한 도움말을 참조하십시오.

기타 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **옵션**을 클릭합니다.

사용할 수 있는 도표 모양 옵션은 작업에서 생성되는 도표의 유형에 따라 다릅니다.

역할 이름	설명
배경	도표의 배경 색상을 지정합니다.
막대 외곽선 및 막대 채우기	히스토그램의 막대 색상을 지정합니다. 주: 이 옵션은 비교 히스토그램에서만 사용할 수 있습니다.
축 바꾸기 : 백분위수를 세로축으로 사용	도표에서 축을 반대로 바꿉니다. 이 옵션을 선택하면 도표의 방향이 바뀌므로 선택한 축 옵션 및 참조선에서 지정한 내용은 적용되지 않습니다. 주: 이 옵션은 확률 도표 및 Q-Q 도표에서만 사용할 수 있습니다.

역할 이름	설명
기본 범례 생략	도표에 기본 범례를 표시하지 않습니다.
중간점 사용	히스토그램에 대한 중간점을 지정합니다. 드롭다운 리스트에서 사용할 수 있는 옵션은 히스토그램을 생성하는지 또는 비교 히스토그램을 생성하는지에 따라 달라집니다. 분류변수 역할에 변수를 할당할 때에는 비교 히스토그램을 생성합니다. 비교 히스토그램에서는 다음 옵션을 사용할 수 있습니다. ■ 중간점 지정 - 시작 중간점과 마지막 중간점을 지정할 수 있습니다. 다른 중간점을 계산하는 데 사용할 단계 간격도 지정할 수 있습니다. 중간점은 각 막대로 표시되는 값 범위의 중간 값입니다. ■ 키 - 키 셀의 데이터에 대해 중간점을 결정합니다. 기본적으로 분류변수의 레벨은 첫 번째 분류변수의 내부(서식화되지 않은) 값의 증가 순으로 위에서 아래로(왼쪽에서 오른쪽으로) 표시됩니다. 분류변수를 하나만 지정하면 키 셀은 이러한 순서로 첫 번째에 나타나는 레벨이 됩니다. 분류변수를 두 개 지정하면 키 셀은 이러한 순서로 첫 번째에 나타나는 변수 1과 변수 2의 레벨 결합이 됩니다. 따라서 키 셀의 선택에 따라 모든 셀에 사용되는 균등 수평 축이 결정됩니다. 그런 다음 키 셀에 대한 중간점 리스트에 나머지 셀의 데이터가 포함될 때까지 리스트가 어느 방향이든지 필요한 만큼 확장됩니다. ■ 균등 - 총 표본 크기를 기반으로 중간점 개수를 결정합니다. 분류변수 역할에 변수를 할당하지 않을 때에는 히스토그램을 생성합니다. 히스토그램에서는 다음 옵션을 사용할 수 있습니다. ■ 중간점 지정 - 시작 중간점과 마지막 중간점을 지정할 수 있습니다. 다른 중간점을 계산하는 데 사용할 단계 간격도 지정할 수 있습니다. 중간점은 각 막대로 표시되는 값 범위의 중간 값입니다. ■ 그룹 수 지정 - 데이터에 대한 막대 수(히스토그램 구간이라고도 함)를 지정할 수 있습니다. 기본적으로 표시되는 구간 수는 데이터 범위에 따라 달라집니다. 그러나 표시할 최대 구간 수는 지정할 수 있습니다. 주: 히스토그램 작업을 사용하는 경우에만 중간점을 지정할 수 있습니다.
사용자 정의 그래프 크기 지정(픽셀)	이미지 크기를 지정합니다. 참조할 수 있도록 전체 화면 크기가 너비 및 높이 상자에 제공됩니다.

주: 비교 옵션은 비교 히스토그램에서만 사용할 수 있습니다. 히스토그램 작업에서 **분류변수** 역할에 변수를 할당하여 비교 히스토그램을 생성할 수 있습니다.

역할 이름	설명
행 및 칼럼	분석 역할에 하나 이상의 분류변수를 할당하면, 페이지당 행 수 및 페이지당 칼럼 수 상자를 사용하여 비교 히스토그램에서의 칼럼과 행 배열 방식을 변경할 수 있습니다. 한 개의 분류변수를 할당하면 페이

역할 이름	설명
	지당 한 개의 칼럼과 두 개의 행이 기본 정렬 방식이 됩니다. 두 개의 분류변수를 할당하면 페이지당 두 개의 행과 두 개의 칼럼이 기본 정렬 방식이 됩니다. 옆 프레임 및 위 프레임 드롭다운 리스트를 사용하여 행 레이블과 칼럼 레이블의 프레임 영역을 채울 색상을 지정할 수도 있습니다.
막대 수	기본적으로 표시되는 구간 수는 데이터 범위에 따라 달라집니다. 그러나 표시할 최대 구간 수는 지정할 수 있습니다. 또한 표시되는 구간 수가 데이터 평균 이상 및 이하의 표준편차 값 범위를 기반으로 하도록 표준편차를 지정할 수 있습니다.

규격 한계 옵션 설정

주: 히스토그램, 확률 도표, Q-Q 도표 및 CDF 도표에서 이러한 옵션을 사용할 수 있습니다.

선택 영역의 **모양** 머리글 아래에서 **규격한계**를 클릭하여 규격하한, 목표 값 및 규격상한을 지정합니다.

주: **규격한계** 옵션은 적어도 하나의 분석변수에 대해 하나 이상의 한계를 지정할 때만 사용할 수 있습니다. 분석변수에 대한 규격한계를 지정하려면 선택 영역에서 **데이터**를 클릭하고 분석변수를 선택합니다.

공정능력분석: 분석 유형 지정

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

공정능력분석 리포트에는 기본적으로 기본 신뢰구간 테이블, 기본 측도 테이블, 위치모수에 대한 검정 테이블, 적률 테이블 등이 포함됩니다.

리포트에 테이블을 추가하려면 **테이블** 상자에서 테이블을 선택합니다.

분석 옵션을 지정하려면 **테이블** 상자에서 테이블을 선택합니다. 옵션을 사용할 수 있으면, 옵션을 지정하는 데 사용한 오른쪽 영역에 상자가 나타납니다.

리포트에서 모든 기술통계량 테이블을 숨기려면 **기술통계량 및 공정능력지수 테이블 생략** 체크 박스를 선택합니다. 이 옵션을 선택해도 다음 테이블은 숨겨지지 않습니다.

- 선택한 분포에 의해 생성된 테이블. 해당 분포에 대한 패널에서 **분포테이블 생략** 체크 박스를 선택하여 이러한 테이블을 숨길 수 있습니다.

- INTERVALS 문에 의해 생성된 테이블. NOPRINT 옵션을 지정하면 이러한 테이블을 생략할 수 있습니다. INTERVALS 문에 대한 자세한 내용은 CAPABILITY 프로시저에 대한 도움말을 참조하십시오.

결과를 출력 데이터셋에 저장하려면 **데이터셋에 출력 통계량 저장** 체크 박스를 선택합니다. 데이터셋에는 일반량 통계량 및 공정능력지수가 포함됩니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간

- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 속성 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

데이터 특성화

데이터 특성화 마법사 정보	93
데이터 특성화: 데이터 선택	93
데이터 특성화: 리포트 옵션 선택	94
데이터 특성화: 출력 한계 지정	94

데이터 특성화 마법사 정보

데이터 특성화 마법사를 사용하여 데이터의 주요 특성을 기술하는 요약 리포트, 그래프, 빈도 및 일변량 SAS 데이터셋을 생성합니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	CONTENTS, FREQ, UNIVARIATE
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	SAS/GRAPH

데이터 특성화: 데이터 선택

- 1 **추가**를 클릭합니다. **열기** 대화 상자가 나타납니다.
- 2 사용할 데이터셋을 선택합니다. 여러 데이터셋을 선택하려면 Ctrl 키를 누릅니다. **열기**를 클릭합니다. 선택한 데이터셋이 **데이터** 상자에 나타납니다.

- 3 다음을 클릭하여 리포트 옵션을 선택합니다.

데이터 특성화: 리포트 옵션 선택

- 1 리포트 옵션 중 하나를 선택하려면 먼저 데이터를 선택해야 합니다.
- 2 출력 유형을 선택합니다. 기본적으로 빈도 데이터 및 일변량 데이터에 대한 요약 리포트, 그래프 및 SAS 데이터셋이 생성됩니다.

SAS 데이터셋을 출력 유형으로 선택하면 **SAS Enterprise Guide**에서는 **옵션** 대화 상자의 **작업 > 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 쓰기 가능한 첫 번째 라이브러리에 출력 데이터를 저장합니다. 위치를 변경하려면 **찾아보기**를 클릭합니다. 결과 데이터셋도 프로젝트에 추가됩니다.

- 3 다음을 클릭하여 출력 한계를 지정합니다.

데이터 특성화: 출력 한계 지정

- 1 출력 한계를 지정하려면 먼저 리포트 옵션 중 하나를 선택해야 합니다.
- 2 **범주형 값 제한** 체크 박스를 선택하여 변수당 리포팅되는 최대 범주형 값 개수를 지정합니다. 기본적으로는 **30**개 값이 리포팅됩니다.
- 3 **마침**을 클릭하여 출력을 생성합니다.

군집분석

군집분석 작업 정보	95
군집분석: 분석 역할에 변수 할당	96
군집분석: 군집 옵션 설정	97
군집분석: 도표 및 다이어그램 생성	98
군집분석: 결과 옵션 설정	99
제목 및 각주 지정	100
속성 보기	101

군집분석 작업 정보

군집분석 작업에서는 좌표 데이터 또는 거리 데이터 등의 SAS 데이터셋 관측값을 계층적으로 군집화합니다. 데이터셋에 좌표 데이터가 포함되어 있으면, 작업은 군집화 방법을 적용하기 전에 유클리드 거리를 계산합니다. 이 작업에서는 계층적 군집화 결과를 그래프로 표시하여, 덴드로그램이라는 트리 다이어그램을 생성할 수 있습니다.

군집분석 작업에서는 K-평균 방법을 사용하여 좌표 데이터를 비계층적으로 군집화할 수도 있습니다. K-평균 방법은 거리 데이터에는 사용할 수 없습니다. 이 작업으로 모집단 데이터를 분석할 수 있습니다.

예를 들어 국가의 특정 유형 또는 범주를 결정하는 데 출생률, 사망률 및 영아 사망률의 국가별 수치를 사용할 수 있는지를 결정하려고 합니다. 군집분석을 수행하여 관측값이 데이터에서 제시한 그룹을 형성할 수 있는지를 결정할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	CLUSTER, FASTCLUS, TREE
필수 SAS 제품	Base SAS, SAS/STAT

요구 사항 이름	프로시저 및 제품 이름
추가 SAS 제품(권장)	SAS/GRAPH

군집분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분석변수	이 역할에 할당된 변수를 사용하여 분석이 수행됩니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크박스의 선택을 취소합니다. 주: Set 1 변수 또는 Set 2 변수에서 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
ID 레이블	군집 히스토리 및 출력 트리 테이블에서 행을 식별하는 데 사용할 값을 지정합니다. 이 역할에 할당된 변수가 없으면, 각 행은 $OBSn$ 으로 표시됩니다(n :
COPY 변수	이러한 변수의 값을 출력 트리 데이터셋으로 복사합니다. 입력 데이터셋의 관측값 군집을 나타내는 출력 트리 데이터셋의 관측값은 COPY 변수의 결측값을 가집니다.

군집분석: 군집 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **군집**을 클릭합니다.

옵션 이름	설명
군집법	분석에 사용할 군집법을 지정합니다. ■ 평균연결법 - 산술평균을 사용한 비가중 쌍그룹법을 사용하여 그룹 평균을 계산합니다. ■ 중심성 방법 - 중심성 방법을 사용합니다(중심성을 사용한 비가중 쌍-그룹 방법, UPGMC, 중심 정렬, 가중 그룹 방법). ■ K-평균 알고리즘 - 유클리드 거리를 사용하므로 군집 중심이 최소제곱추정을 기반으로 합니다. 이 방법에서는 알고리즘을 실행하여 수렴이 완료되면 군집 중심이 각 군집에 할당된 관측값 평균임을 가정합니다. 수렴에 도달할 때까지 각 반복 시마다 최소제곱 기준이 낮아 집니다. ■ Ward 최소분산법은 추적 W를 사용하여 오차제곱합을 계산하는 Ward의 최소분산법을 사용합니다.
K-평균 군집 옵션	K-평균 알고리즘에 대한 추가 옵션을 지정합니다. ■ 최대 군집 개수 - 군집의 최대 개수를 지정합니다. ■ 최대 반복 횟수 - 군집 시드를 재계산하는 최대 반복 횟수를 지정합니다. ■ 시드 바꾸기 - 시드 대체 방법을 지정합니다. □ full - 시드와 가장 가까운 시드 간의 거리가 시드 사이의 최소 거리보다 크거나 관측값에서 가장 가까운 시드를 제외한 모든 시드까지의 최소 거리가 최근접 시드에서 다른 모든 시드까지의 최단 거리보다 클 때 시드를 대체합니다. □ partial - 관측값과 최근접 시드 간의 거리가 시드 사이의 최소 거리보다 클 때에만 시드를 바꿉니다. □ none - 시드 교체를 억제합니다. □ random - 전체 관측값의 단순 난수 표본을 초기 군집 시드로 선택합니다. ■ 랜덤 시드 지정 - 양의 정수를 의사 랜덤 생성기의 시작 값으로 지정합니다. 주: 이 옵션은 시드 바꾸기 옵션으로 random을 선택할 때만 사용할 수 있습니다.

군집분석: 도표 및 다이어그램 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

주: K-평균 알고리즘 군집법에서는 도표와 다이어그램을 사용할 수 없습니다.

군집화 프로세스를 그래프로 표시하면 군집을 이해하는 데 도움이 됩니다. 입력 데이터 소스의 서브셋을 사용하여 도표가 생성됩니다. 데이터의 서브셋을 지정하는 방법은 CCC(3차원 군집화 기준) 통계량을 계산할 수 있는지 여부에 따라 달라집니다. CCC 통계량을 계산할 수 있으면 CCC 통계량의 비결측값을 기준으로 데이터의 서브셋이 지정됩니다. CCC 통계량을 계산할 수 없으면 의사 F 통계량의 비결측값을 기준으로 데이터의 서브셋이 지정됩니다.

옵션 이름	설명
의사 F 통계량	군집 수에 대한 의사 F 통계량의 산점도를 표시합니다. 주: 결과 패널에서 의사 F 및 t-제공 통계량 옵션을 선택하면 의사 F 통계량 그래프가 항상 생성됩니다.
의사 t-제공 통계량	군집 수에 대한 의사 t^2 통계량의 산점도를 표시합니다. 주: 결과 패널에서 의사 F 및 t-제공 통계량 옵션을 선택하면 의사 t-제공 통계량 그래프가 항상 생성됩니다.
CCC 통계량	군집의 개수에 대한 CCC(3차원 군집화 기준) 통계량의 산점도를 표시합니다. 주: 입력 데이터셋이 DISTANCE 유형이면 이 옵션을 사용할 수 없습니다.
트리 다이어그램 방향	군집에 대한 트리 다이어그램을 표시합니다. 군집된 개체들을 리프라고 하고, 모든 개체를 포함하는 군집이 뿌리입니다. 그 중 2개 이상의 개체를 포함하는 군집을 가지라고 합니다. 두 개 이상의 개체를 포함하지만 일부만 포함하는 군집은 가지입니다. 리프, 가지, 루트의 일반적인 용어는 노드입니다. 군집 A가 군집 B와 C의 합집합이면 A는 B와 C의 상위이고 B와 C는 A의 하위입니다. 따라서 A 리프는 하위가 없는 노드이고 루트는 부모가 없는 노드입니다. 모든 군집에 두 개의 하위가 있으면 트리 다이어그램은 이진 트리입니다. 높이 축이 수평이고 루트가 왼쪽에 있는 트리 다이어그램의 방향을 지정하려면 수평을 선택합니다. 높이 축이 수직이고 루트가 맨 위에 오도록(기본값) 트리 다이어그램의 방향을 지정하려면 수직을 선택합니다.

군집분석: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
-------	----

- 출력 데이터 저장** 출력 데이터셋을 생성합니다. 다음 옵션 중에서 선택할 수 있습니다.
- **군집 트리 데이터** - 군집 계층의 트리 다이어그램을 그리는 데 사용할 수 있는 출력 데이터셋을 생성합니다. 데이터셋에는 입력 데이터셋의 각 관측값에 대한 한 개의 관측값 및 둘 이상의 관측값에 대한 한 개의 관측값이 포함됩니다(즉, 군집 트리의 각 노드에 대해 하나의 관측값). 일반적으로 출력 관측값의 총 개수는 $2n-1$ 입니다(n :
 - **트리 출력** - 평균 연결 방법, 중심 방법 및 Ward의 최소분산법에 대한 출력 데이터셋을 생성합니다. 데이터셋에는 트리의 각 개체에 대해 하나의 관측값이 있으며 트리의 지정된 레벨에서 군집 멤버십을 나타내는 CLUSTER 및 CLUSNAME 변수가 있습니다.
- 또한 테이블에 포함하려는 군집 또는 레벨의 개수를 지정할 수도 있습니다.
- 출력 데이터셋에 포함하려는 군집의 개수를 지정하려면 **군집** 체크박스를 선택합니다. 군집의 기본 개수는 10개입니다. 그러나 다음과 같은 조건에 해당되면 지정한 개수와 군집의 개수가 동일하지 않을 수 있습니다.
 - 트리에 있는 리프 개수가 원하는 군집의 개수보다 적을 때
 - 데이터셋에 연결되지 않은 트리의 개수가 원하는 군집의 개수보다 클 때
 - 다원 트리에 지정한 개수의 군집이 있는 레벨이 포함되지 않을 때
 - 출력 데이터셋에 포함할 레벨의 개수를 지정하려면 **가지치기** 체크박스를 선택한 다음 출력 데이터셋의 배반군집을 정의하는 트리 레벨을 지정합니다. 뿌리와 지정한 레벨의 높이 사이에 해당하는 군집만 표시됩니다.
 - **K-평균 군집** - K-평균 알고리즘을 위한 출력 데이터셋을 생성합니다. 출력 데이터셋에는 모든 원 데이터 및 군집분석에서 생성된 CLUSTER, DISTANCE 변수가 포함됩니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

옵션 이름	설명
출력 표시	출력을 표시합니다. 다음 옵션을 지정할 수 있습니다. ■ 출력 표시 - 군집화 프로세스의 히스토리 정보를 표시합니다. 군집의 개수, 연결된 군집 이름, 새로운 군집에 있는 관측값의 수 등이 포함됩니다. 표시할 군집 생성 히스토리의 수를 군집 생성 상자에 지정합니다. ■ 단순 요약통계량 - Bimodality의 평균, 표준편차, 왜도, 첨도, 계수를 표시합니다. ■ CCC(Cubic clustering criterion) - 균등 귀무가설 아래의 삼차 군집화 기준 및 군사 기대 R 제곱을 표시합니다. 이 옵션은 좌표 데이터에만 적용됩니다. 이러한 통계량은 데이터 군집 개수를 결정하는 데 유용합니다. CCC 값이 2 또는 3보다 크면 군집의 상태가 양호한 것이며, 0 - 2 사이의 값은 군집에 문제가 있을 수 있음을 나타내므로 주의를 기울여야 합니다. 주: 입력 데이터셋이 DISTANCE 유형이면 이 옵션을 사용할 수 없습니다. ■ 의사 F 및 t-제곱 통계량 - 의사 F 및 t-제곱 통계량을 표시합니다. 이 통계량은 군집 수에 대한 유용한 지표로 사용될 수 있습니다. 의사 F 통계량의 값이 상대적으로 큰 것은 중단점을 나타냅니다. 의사 t-제곱 통계량을 해석하는 일반적인 규칙은 이전 값보다 크다고 표시된 첫 번째 값을 찾을 때까지 칼럼을 아래로 이동한 다음, 한 군집씩 다시 위로 돌아오는 것입니다. 주: 이 옵션을 선택하면 도표 패널에서 의사 F 통계량 및 의사 t 제곱 통계량 옵션을 선택하더라도 의사 F 및 t 제곱 통계량 도표가 생성됩니다. 주: 출력 표시는 K-평균 알고리즘 군집법에서 사용할 수 있는 유일한 표시 옵션입니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

데이터 비교

데이터 비교 마법사 정보	103
데이터 비교: 데이터 선택	104
데이터 비교: 비교할 변수 선택	104
데이터 비교: 비교 기준 지정	104
데이터 비교: 출력 옵션 지정	105

데이터 비교 마법사 정보

데이터 비교 마법사를 사용하여 데이터셋 내 또는 데이터셋 간의 두 데이터셋이나 변수를 비교하는 리포트를 생성합니다.

데이터 비교 마법사를 사용하여 SAS 데이터셋의 변경 사항을 비교할 수 있습니다. 예를 들어 컴퓨터에 데이터셋을 복사하여 로컬 컴퓨터와 서버에 있는 데이터셋 간의 차이를 비교한다고 가정합니다. 데이터 비교 마법사를 사용하면 원 데이터셋에 대한 변경 사항의 요약 확인할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	COMPARE
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	없음

데이터 비교: 데이터 선택

- 1 두 데이터셋 간의 변수를 비교할 것인지 아니면 동일한 데이터셋 내의 변수를 비교할 것 인지를 지정합니다.
- 2 사용하려는 기본 및 비교 데이터셋을 선택합니다.
- 3 다음을 클릭하여 비교할 변수를 선택합니다.

데이터 비교: 비교할 변수 선택

기본적으로 테이블에서 이름과 유형이 같은 모든 변수는 일치합니다. 추가 변수를 포함하고 비교할 변수를 편집하며 변수를 삭제할 수 있습니다.

비교할 변수를 지정하는 방법

- 1 비교할 변수를 선택하려면 먼저 데이터를 선택해야 합니다.
- 2 **데이터 일치 기준** 드롭다운 리스트에서 변수를 비교할 기준을 선택합니다. 관측값 또는 ID 변수 중에서 데이터를 일치시킬 기준을 선택할 수 있습니다. ID 변수에 따라 데이터를 일치시키도록 선택하면 변수는 기본 및 비교 데이터셋에 모두 존재해야 합니다. 동일한 데이터셋 내에서 변수를 비교하는 중이라면 데이터 일치 기준으로 관측값만 선택할 수 있습니다.
- 3 새로운 변수 쌍을 테이블에 추가하려면 **새로 만들기**를 클릭합니다. 새로운 행이 테이블에 추가됩니다. 각 셀의 드롭다운 리스트를 사용하여 비교할 변수를 선택합니다.
변수 쌍 중 하나가 올바르지 않으면 해당 셀의 드롭다운 리스트를 사용하여 올바른 변수 이름을 선택합니다. **편집**을 클릭할 수도 있습니다.
변수 쌍을 삭제하려면 테이블에서 행을 선택하고 **삭제**를 클릭합니다.
- 4 다음을 클릭하여 비교 기준을 지정합니다.

데이터 비교: 비교 기준 지정

- 1 비교 기준을 지정하려면 먼저 비교할 변수를 선택해야 합니다.
- 2 **동등 기준** 상자에서 동등 기준값을 지정합니다.

- 3 **동등 판단법** 드롭다운 리스트에서 동등 결정 방법을 선택합니다. 선택한 방법에 따라 결정되는 숫자 값의 차이 크기가 동등 기준의 값보다 크면 숫자 값이 같지 않은 것으로 결정됩니다. 다음과 같은 방법을 사용할 수 있습니다.
 - **절대** - 값의 절대 차이를 동등 기준 값과 비교합니다. Y에서 X를 뺀 절대 편차가 동등 기준보다 크면 값이 다르게 결정됩니다.
 - **정확** - 정확 동등 여부를 검정합니다. Y값과 X값이 다르면 값은 다르게 결정됩니다.
 - **백분율** - 절대 백분율 차이를 동등 기준의 값과 비교합니다.
 - **상대** - 절대 상대 차이를 동등 기준 값과 비교합니다.
- 4 결측값 처리 방법을 지정합니다. **기본 데이터셋의 결측값을 같은 범주의 임의 값과 동등한 것으로 처리** 또는 **비교 데이터셋의 결측값을 임의 값과 동등한 것으로 처리** 옵션 중에서 선택할 수 있습니다.
- 5 다음을 클릭하여 출력 옵션을 지정합니다.

데이터 비교: 출력 옵션 지정

기본적으로 데이터 비교 작업은 데이터셋 요약, 변수 요약, 관측값 요약, 값 비교 요약 및 값 비교 결과가 포함된 HTML 리포트를 생성합니다. 또한 관측 출력이 포함된 출력 데이터셋과 요약통계량이 포함된 출력 데이터셋을 생성하도록 선택할 수 있습니다.

- 1 출력 옵션을 지정하려면 먼저 비교 기준을 지정해야 합니다.
- 2 **출력 데이터셋 포함**을 선택하여 일치하는 각 관측값에 대한 행이 포함된 출력 데이터셋을 생성합니다. 이 데이터셋에는 관측값의 각 변수, 관측값 유형(_TYPE_)에 대한 칼럼 및 관측값 번호에 대한 칼럼(_OBS_)에 대한 칼럼이 포함되어 있습니다. _TYPE_ 칼럼의 값은 다음 유형 중 하나입니다.
 - **BASE** - 이 관측값의 값은 기본 데이터셋의 관측값에서 나온 것입니다.
 - **COMPARE** - 이 관측값의 값은 비교 데이터셋의 관측값에서 나온 것입니다.
 - **DIF** - 이 관측값의 값은 기본 및 비교 데이터셋값 간의 차이입니다.
 - **PERCENT** - 이 관측값의 값은 기본 및 비교 데이터셋값 간의 백분율 차이입니다.

다음 관측값이 출력 데이터셋에 포함되도록 선택할 수 있습니다.

- **출력 데이터셋에 기본 데이터의 각 관측값 쓰기** - 출력 데이터셋에 기본 데이터셋의 관측값을 기록합니다. 출력 데이터셋의 _TYPE_ 칼럼값은 BASE로 설정되어 있습니다.
- **출력 데이터셋에 비교 데이터의 각 관측값 쓰기** - 출력 데이터셋에 비교 데이터셋의 관측값을 기록합니다. 출력 데이터셋의 _TYPE_ 칼럼값은 COMP로 설정되어 있습니다.
- **차이 값 포함** - 출력 데이터셋에 기본 및 비교 데이터셋값 간의 차이를 기록합니다. 출력 데이터셋의 _TYPE_ 칼럼값은 DIF로 설정되어 있습니다.
- **백분율 차이 값 포함** - 출력 데이터셋에 기본 및 비교 데이터셋값 간의 백분율 차이를 기록합니다. 출력 데이터셋의 _TYPE_ 칼럼값은 PERCENT로 설정되어 있습니다.

- **모든 값이 동일하면 관측값 생략** - 출력 데이터셋에 모든 변수값이 같은 관측값이 포함되지 않습니다.
- 3 **요약통계량에 대한 출력 데이터셋 포함**을 선택하여 각 변수 쌍의 요약통계량에 대한 행이 포함된 출력 데이터셋을 생성합니다. 출력 데이터셋에는 다음 칼럼이 포함됩니다.
 - **_VAR_** - 기본 데이터셋의 변수 이름이 포함됩니다.
 - **_WITH_** - 비교 데이터셋의 변수 이름이 포함됩니다.
 - **_TYPE_** - 관측값의 통계량 이름이 포함됩니다.
 - **_BASE_** - 일치하는 비교 데이터셋의 관측값이 있는 기본 데이터셋의 변수값에서 계산된 통계량값이 포함됩니다.
 - **_COMP_** - 일치하는 기본 데이터셋의 관측값이 있는 비교 데이터셋의 변수값에서 계산된 통계량값이 포함됩니다.
 - **_DIF_** - 일치하는 비교 데이터셋의 변수와 기본 데이터셋의 변수값 차이에서 계산된 통계량값이 포함됩니다.
 - **_PCTDIF_** - 일치하는 비교 데이터셋의 변수와 기본 데이터셋의 변수값 간의 백분율 차이에서 계산된 통계량값이 포함됩니다.
 - 4 출력할 최대 차이 수를 지정합니다.
 - 5 **마침**을 클릭하여 출력을 생성합니다.

등고선 도표

등고선 도표 작업 정보	107
등고선 도표: 도표 유형 선택	108
등고선 도표: 분석 역할에 변수 할당	108
등고선 도표: 모양 옵션 설정	109
등고선 도표 옵션 설정	109
추가 모양 옵션	110
제목 및 각주 지정	110
속성 보기	110

등고선 도표 작업 정보

등고선 도표 작업은 3개의 숫자 변수 간의 수학적 관계를 보여 주는 선 도표, 색칠된 도표, 패턴 도표 또는 평할 도표를 생성합니다. 이 도표는 3차원적 관계를 2차원으로 나타냅니다. 등고선 도표의 선이나 영역은 평면에서의 위치(x, y)에 대응하는 크기(z) 레벨을 나타냅니다.

이 작업을 사용하여 자연 현상을 나타낼 수 있습니다. 예를 들어 여러 테스트 현장의 토양 샘플에서 찾을 수 있는 진흙의 비율을 보여 주는 등고선 도표를 생성할 수 있습니다. 도표의 x 및 y 축은 x-y 위치에서의 표면 높이에 대한 그래프를 나타냅니다. 등고선은 지정된 진흙의 백분율을 포함하는 평면에서의 위치를 나타냅니다.

자세한 내용은 [“그래픽 출력형식 선택” \(5페이지\)](#)에서 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GCONTOUR, G3GRID
필수 SAS 제품	Base SAS, SAS/GRAPH
추가 SAS 제품(권장)	없음

등고선 도표: 도표 유형 선택

이 옵션에 액세스하려면 선택 영역에서 **등고선 도표**를 클릭합니다.

옵션 이름	설명
등고선 도표	선을 사용하여 Z(또는 등고선) 변수를 표현합니다.
색칠된 등고선 도표	각각의 등고선을 여러 다른 색상으로 채웁니다. 각 등고선 레벨마다 한 가지 색상이 지정됩니다.
색칠된 외곽선이 있는 등고선 도표	각 등고선을 여러 다른 색으로 채우고, 등고선의 위치를 뚜렷하게 표시하기 위해 각 영역의 윤곽을 그립니다.
외곽선이 색칠된 등고선 도표	패턴으로 채워진 직사각형으로 등고선 레벨을 표시합니다. 각 직사각형 패턴은 직사각형의 네 모서리에 대한 Z 변수값의 평균을 계산하여 평균에 가까운 수준의 패턴을 할당하여 결정됩니다.
평할 등고선 도표	인접한 그리드 셀을 단일 패턴 영역을 이루는 같은 패턴으로 결합합니다.

등고선 도표: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
수평	이 역할에 할당하는 칼럼은 그래프의 가로축 또는 X축 변수입니다.
수직	이 역할에 할당하는 칼럼은 그래프의 세로축 또는 Y축 변수입니다.
깊이	이 역할에 할당하는 칼럼은 도표의 세 번째 차원 또는 Z축 변수입니다.

역할 이름	설명
그룹 분석 칼럼	이 역할에 할당하는 칼럼의 값에 따라 생성되는 그래프의 수가 결정됩니다. 각 고유티값에 대해 별도의 그래프가 생성됩니다.

등고선 도표: 모양 옵션 설정

등고선 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **등고선**을 클릭합니다.

역할 이름	설명
등고수준의 개수 지정	등고수준의 개수 즉, 깊이 역할(또는 Z축 변수)에 할당된 변수의 값을 분할하는 영역 수를 지정합니다.
등고수준에 자동적으로 레이블 생성	등고선에 자동으로 레이블을 생성하려면 이 체크 박스를 선택합니다. 각 등고선의 레이블은 해당 등고선의 깊이 값입니다. 이때에도 범례는 나타납니다. 주: 이 옵션은 ActiveX, ActiveX 이미지, Java 및 Java 이미지 그래프 출력형식에서 지원되지 않습니다.
보간 옵션을 통한 데이터 사전 처리	등고선 작업 실행 중 로그에 "Less than half of the grid cells have data values"라는 오류가 나타나면 사전 처리 체크 박스를 선택한 후 작업을 재실행합니다. 수평(x 및 y) 변수값들로 완벽한 그리드를 이루는 데이터셋이 생성되고 X-Y 평면 위의 각 포인트에 대해 수직(Z) 변수값이 보간됩니다. 사용자는 다음 보간 옵션 중 하나를 선택하여 보간법을 제어할 수 있습니다. ■ 조인은 삼각형 영역 집합 내에서 선형 보간법을 사용합니다. 이 보간법은 수직 변수의 초기 값 범위 내의 값을 생성하지만 생성되는 보간 표면은 평활하지 않습니다. ■ 척도 없음은 보간 전 X 및 Y 변수가 동일한 범위로 척도화되지 않도록 합니다. ■ 부분은 보간에 대한 도함수를 추정하기 위해 스플라인을 사용합니다. 이변량 스플라인은 가장 인접한 데이터에 적합하여 필요한 미분계수를 추정하는 데 사용됩니다. 등고선은 Spline 옵션처럼 매끄럽지는 않지만 컴퓨팅 리소스가 적게 사용됩니다. ■ Spline - 보간을 위해 이변량 스플라인을 사용합니다. 이 방법은 데이터 포인트가 100개 이상이면 특히 시간이 오래 걸릴 수 있습니다.

추가 모양 옵션

일부 옵션은 여러 그래프 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- “축 옵션 설정” (593페이지)
- “참조선 추가” (595페이지)
- “범례 옵션 설정” (596페이지)
- “그래프 영역 사용자 정의” (597페이지)
- “고급 옵션 설정” (602페이지)

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블

- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 속성 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

파일 복사

파일 복사 작업 정보	113
로컬 컴퓨터에서 SAS 서버로 파일 복사	113

파일 복사 작업 정보

파일 복사 작업에서는 로컬 컴퓨터에서 SAS Server로 파일을 전송할 수 있습니다. SAS Server에서 로컬 컴퓨터로도 파일을 전송할 수 있습니다. 파일 복사 작업은 FTP 응용 프로그램과 유사하게 작동합니다. 그러나, 이 작업은 파일 전송을 완료하는 데 SAS 프로토콜에 의존하며 FTP 서버는 필요하지 않습니다.

이 작업에서는 **SAS Integration Technologies**를 사용하여 네트워크를 통해 파일 데이터를 복사합니다. 작업은 **SAS/CONNECT**의 PROC DOWNLOAD 또는 PROC UPLOAD에 의존하지 않습니다.

.....
 주: 이 작업은 **SAS Enterprise Guide**에서만 사용 가능합니다. **SAS Enterprise Guide** 이외의 작업에서 생성된 코드를 실행할 수 없습니다. **SAS Enterprise Guide** 프로젝트 또는 프로세스 플로우를 SAS 프로그램으로 내보내려면, 파일 복사 작업 단계는 포함되지 않습니다.

로컬 컴퓨터에서 SAS 서버로 파일 복사

- 1 복사할 파일의 소스 또는 대상으로 사용할 서버를 선택합니다.
- 2 전송 유형을 선택합니다. 로컬 컴퓨터에서 SAS 서버로 파일을 업로드하거나 SAS 서버에서 로컬 컴퓨터로 파일을 다운로드할 수 있습니다.

- 3 복사할 파일의 전체 경로를 지정합니다. **복사할 소스 파일** 상자에서 여러 파일과 일치시킬 표준 와일드카드 문자를 사용할 수 있습니다. 별표(*)는 몇 개든지 와일드카드가 아닌 다음 문자 앞의 모든 문자와 일치합니다. 물음표(?)는 파일 이름의 해당 위치를 점유하고 있는 단일 문자와 일치합니다.

일치하는 와일드카드와 패턴 유형은 다음과 같습니다.

/u/dept/finance/*.xls

/u/dept/finance 폴더 내 모든 XLS 파일과 일치합니다.

C:\Data\??Sep2012.csv

이름이 2자로 시작하고 "Sep2012.csv"로 끝나는 모든 파일과 일치합니다.

"01Sep2012.csv" 및 "15Sep2012.csv"를 예로 들 수 있습니다. 파일 이름

"199Sep2011.csv"는 이 패턴과 일치하지 않습니다. 접두어가 두 자가 아닌 세 자이기 때문입니다.

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

파일 이름이 "15"와 그 다음 3자, 그 다음 "2012", 그 다음 파일 확장자로 시작하는

모든 파일과 일치합니다. "15Sep2012.csv" 및 "15Oct2012.xls"를 예로 들 수 있습니다.

주: 와일드카드 문자는 파일 이름에서만 허용되며 폴더 경로에서는 허용되지 않습니다. 예를 들어 **c:\data\Sept*\report.xls**는 이 작업에 대한 올바른 소스 파일 지정 사항이 아닙니다.

- 4 로컬 컴퓨터 또는 SAS 서버에서 파일을 복사할 위치를 지정합니다. 대상 폴더를 지정할 때 다음 규칙을 유념해두십시오.

- 폴더 이름은 타겟 시스템에 이미 존재하고 있어야 합니다. 파일 복사 작업은 새로운 폴더를 생성하지 않습니다.
- 파일에 콘텐츠를 쓸 알맞은 파일 권한이 있어야 합니다.
- 타겟 폴더의 절대 경로를 지정해야 합니다(파일 업로드, SAS Workspace가 SASUSER의 시작 루트 경로 또는 비 시스템 루트 경로로 구성될 때).

- 5 파일 또는 폴더 지정 사항에서 매크로 변수를 사용할지를 지정합니다. **소스 및 대상 경로에서 SAS 매크로 변수 찾기** 체크 박스를 선택하여 파일 이름 또는 대상 폴더 경로의 SAS 매크로 변수 및 표현식을 사용해서 파일 복사 작업 좀 더 동적으로 작동하도록 합니다. 이 옵션이 선택되면, 파일 복사 작업은 시작하기 전 매크로 변수 및 매크로 함수 평가를 시도합니다.

매크로 참조를 포함하는 예와 확인 방법은 아래와 같습니다.

/u/&SYSUSERID/Data

SAS Workspace Server 내에서 SAS 사용자의 사용자 ID를 포함하는 폴더를 확인합니다.

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

SAS Workspace Server 내에서 SAS 사용자 ID의 부분만 포함하는 폴더를 확인합니다.

/u/Data/&outputFile.xls

파일 이름의 루트가 &OUTPUTFILE 매크로 변수에 저장되는 XLS 파일 이름으로 확인됩니다. 이 매크로 변수를 사용하면 이전 프로세스(SAS 프로그램)가 기존 파일 이름과 일치하도록 설정됐다고 가정됩니다.

주의

매크로 변수는 SAS 세션(SAS 빌트인 매크로 변수 또는 이전 프로세스에서 정의된 세션) 내에서 존재해야 합니다. 매크로 변수가 올바르게 지정되지 않았거나 매크로 함수에 올바르게 않은 구문을 사용하면, 파일 복사 작업에서 오류 또는 예상하지 못한 결과가 생성될 수 있습니다.

매크로 표현식 확인을 올바르게 확인하려면, **매크로 값 테스트**를 클릭합니다. 간단한 SAS 프로그램을 실행하여 매크로 표현식을 확인하면 결과는 새로운 창에 표시됩니다. 이러한 값을 테스트할 때에는 매크로 표현식이 매크로 변수의 현재 상태 또는 사용하고 있는 매크로 함수 결과를 반영하는 값을 확인함에 유념하십시오. 이러한 값은 작업을 나중에 실제 실행할 때 달라질 수 있습니다.

- 6 대상 폴더의 기존 파일을 덮어쓸지를 지정합니다. 기본적으로, 파일 복사 작업은 기존 파일을 덮어쓰지 않습니다. 소스 파일이 기존 파일의 이름과 일치하면, 작업은 해당 파일을 복사하지 않고 로그에 노트를 추가합니다.
- 7 작업이 알맞은 행 종료 스타일을 사용하여 전송된 텍스트 파일을 자동 감지 및 다시 쓸지를 지정합니다. UNIX 및 Windows 환경은 텍스트 파일의 행 종료 문자에 대해 다른 규칙을 사용합니다. UNIX 환경은 LF(라인피드) 문자를 사용하고 Windows 환경은 LF(라인피드)와 CR(캐리지 리턴)을 차례로 사용합니다. CR 문자가 없으면 UNIX 텍스트 파일을 노트패드와 같은 표준 텍스트 편집 응용 프로그램에서 읽기 어렵습니다.

상관분석

상관분석 작업 정보	117
상관분석: 분석 역할에 변수 할당	118
상관분석: 상관분석 옵션 설정	118
상관분석: 결과 옵션 설정	120
상관분석: 출력 데이터 생성	121
제목 및 각주 지정	121
속성 보기	121

상관분석 작업 정보

데이터 상관분석은 숫자 변수 사이의 관계를 설명하기 위한 통계적 프로시저입니다. 이러한 관계는 변수의 상관계수를 계산하여 설명됩니다.

상관분석 작업을 사용하여 제품 판매에 대한 추가 정보를 얻을 수 있습니다. 예를 들어 이 작업을 사용하여 고객의 규모(판매량으로 측정)와 고객이 회사에서 구매하는 제품량 간에 상관관계가 있는지 결정할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	CORR
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

상관분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분석변수	이 역할에 할당된 변수를 사용하여 분석이 수행됩니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
상관 분석변수	이 역할에 할당된 각 변수와 분석변수의 상관관계를 분석합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석변수 또는 상관 분석변수로 이미 할당한 변수로는 그룹 분석을 수행할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
편상관변수	사용자가 선택하는 상관분석의 종류에 따라 변수가 Pearson의 편상관계수, Spearman의 편순위상관계수 또는 Kendall의 부분 Tau-b 계수에서 사용할 변수를 지정합니다.
상대 가중값 변수	곱적률상관계수를 계산하기 위한 각 행의 변수값을 지정합니다. 이 역할에는 최대 한 개의 칼럼을 지정할 수 있습니다.

상관분석: 상관분석 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

옵션 이름	설명
상관계수 종류	상관분석을 지정합니다. ■ Pearson - Pearson의 곱적률상관을 계산합니다. 이것은 두 연속 확률변수의 모수적 연관성측도입니다. 상관계수 범위는 -1에서 1 사이입니다. ■ Hoeffding - Hoeffding의 종속 측도 D를 계산합니다. 이는 보다 일반적인 독립성으로부터의 이탈을 탐지하는 비모수적 연관성 측도입니다. 이 D 통계량은 일반적인 정의의 30배이며 큰 양의 값만 종속 관계를 나타내도록 -0.5와 1 사이의 범위를 조정합니다. ■ Kendall - Kendall tau-b를 계산합니다. 쌍을 이루는 관측값의 일치도 및 불일치 수를 기반으로 하는 비모수적 연관성 측도입니다. 일치도는 쌍을 이루는 관측값이 서로 다를 때 발생하고 불일치도는 쌍을 이루는 관측값이 서로 다를 때 발생합니다. Kendall의 tau-b 범위는 -1에서 1 사이입니다. ■ Spearman - Spearman의 순위상관계수를 계산합니다. 이것은 데이터 값의 순위를 기반으로 하는 비모수적 연관성측도입니다. 상관계수 범위는 -1에서 1 사이입니다. 주: 상대 가중값 역할에 변수를 할당하면 Hoeffding, Kendall 및 Spearman 상관계수 종류를 사용할 수 없습니다. Hoeffding 상관계수 유형은 편상관변수 역할에 변수를 지정할 때도 사용할 수 없습니다.
Pearson 상관계수 옵션	Pearson 상관계수에 대한 옵션을 지정합니다. ■ Cronbach의 alpha 계수 - Cronbach의 alpha 계수를 계산합니다. 원시 및 표준화된 값(변수를 단위 분산 1로 척도화하여)을 사용하는 별도의 계수가 계산됩니다. 각 상관변수에 대해 변수와 나머지 변수의 합계 간의 상관관계가 계산됩니다. Cronbach의 alpha 계수는 나머지 변수를 사용해서만 계산됩니다. 주: 상관 분석변수 역할에 변수를 지정하면 이 옵션을 사용할 수 없습니다. ■ 공분산 - 공분산행렬을 계산합니다. 부분 변수를 지정하면 부분공분산행렬이 계산됩니다. ■ 제공합과 교차곱 - 제공합과 교차곱을 계산합니다. 부분 변수를 할당하면 비부분 SSCP 행렬이 계산됩니다. ■ 수정제공합과 교차곱 - 수정제공합과 교차곱을 계산합니다. 부분 변수를 할당하면 비부분 및 부분 CSSCP 행렬이 모두 계산됩니다. ■ 결과에 Pearson 상관계수 생략 - 결과에서 Pearson 상관분석을 생략합니다.
Fisher 옵션	Fisher 옵션을 사용하여 Fisher의 z 변환을 사용하는 상관계수에 대해 지정된 귀무가설 하에서 신뢰한계 및 p-값을 요청합니다. 주: 이러한 옵션은 Pearson 또는 Spearman을 상관계수 종류로 선택할 때만 사용할 수 있습니다. ■ alpha 값 지정 - 상관분석의 신뢰한계 레벨을 지정합니다. 이 옵션의 값은 0에서 1 사이여야 합니다. 기본값은 0.05입니다.

옵션 이름	설명
	■ 신뢰한계 유형 - 신뢰한계 유형을 지정합니다. 기본값은 양측 대체에서 양측 신뢰한계입니다. ■ 귀무가설의 rho0 값 지정 - 귀무가설 하에서 rho0 값을 지정합니다. 기본값은 0입니다. ■ 신뢰레벨 생성에 편향(bias) 보정 사용 - 신뢰한계를 생성할 때 편향(bias) 보정을 사용할지를 지정합니다. 기본적으로 이 체크 박스가 선택되며 편향(bias) 보정을 사용하여 새로운 상관계수 추정값이 생성됩니다.
분산 계산에 사용할 분모	분산, 표준편차 및 공분산 계산에 사용할 분모를 지정합니다. 분산 계산에 사용되는 기본 분모는 자유도입니다.
상관 변수에 대해 결측값이 있는 행 제외	분석변수에 결측값이 포함되지 않은 행만 사용합니다. 기본적으로 상관분석 작업은 분석변수의 각 쌍에 대해 모든 비결측값 쌍을 사용합니다. 이것은 일부 상관계수는 다른 상관계수보다 더 많은 관측값을 사용하여 계산된다는 것을 의미합니다.

상관분석: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
도표	각 상관관계 쌍에 대한 산점도를 생성합니다.
표시할 결과	상관분석에 대한 옵션을 지정합니다. 결과에 다음과 같은 값을 포함할 수 있습니다. ■ 각 분석변수에 대해 평균, 표준편차, 합, 최솟값 및 최댓값 등의 요약계량 표시 ■ 각 상관관계와 연관된 유의수준 표시 ■ 상관계수 순서화(큰 값부터 작은 값 순서). 각 분석변수에 대해 상관 분석변수는 상관계수의 절댓값이 내림차순으로 나열됩니다. 기본적으로 상관계수는 순위화되지 않습니다. 상관계수를 순위화하도록 선택하면, 각 변수에 대해 표시할 상관계수의 수를 지정할 수도 있습니다. 최대 수보다 적은 수의 상관관계를 지정하면, 상관관계 분석 작업은 절대값이 가장 높은 것부터 가장 낮은 것까지 상관관계의 순위를 매기고 최량의 n개 상관관계를 보고합니다. 여기서 n은 지정한 상관관계 수입니다.

상관분석: 출력 데이터 생성

이 옵션에 액세스하려면 선택 영역에서 **출력 데이터**를 클릭합니다.

출력 데이터 저장 체크 박스를 선택합니다. **상관계수 종류** 영역에서 선택한 상관계수 종류에 따라 최종 출력에 포함되는 통계량이 결정됩니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

출력형식 생성

출력형식 생성 작업 정보	123
출력형식 생성: 출력형식 옵션 설정	124
출력형식 생성: 새로운 출력형식 정의	125
새로운 출력형식을 정의하는 방법	125
예 1: 문자 출력형식	125
예 2: 숫자 출력형식	126
속성 보기	127

출력형식 생성 작업 정보

출력형식 생성 작업에서는 숫자 또는 문자 출력형식이 생성됩니다. 숫자 출력형식은 숫자 변수의 저장값을 표시하거나 인쇄할 때 사용되는 텍스트를 지정합니다. 문자 출력형식은 문자 변수에 대해 동일한 작업을 수행합니다. 이 작업을 사용하여 출력의 형식을 지정할 수 있습니다.

이 작업을 사용하여 다음과 같은 작업을 수행할 수 있습니다.

- 숫자 값을 문자 값으로 인쇄(예: 1을 MALE로, 2를 FEMALE로 인쇄)
- 하나의 문자열을 다른 문자열로 인쇄(예: YES를 OUI로 인쇄)
- 템플릿을 사용하여 숫자값을 인쇄(예: 9458763450을 945-876-3450으로 인쇄).

주: 이 작업은 **SAS Enterprise Guide**에서만 사용 가능합니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	FORMAT
필수 SAS 제품	Base SAS

요구 사항 이름	프로시저 및 제품 이름
추가 SAS 제품(권장)	없음

출력형식 생성: 출력형식 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

옵션 이름	설명
출력형식 이름	생성할 출력형식의 이름을 지정합니다. 이 이름은 적용할 때마다 출력형식을 참조하는 데 사용됩니다. 출력형식 이름에는 영문자, 숫자 및 밑줄(_)만 포함될 수 있습니다. ■ 다른 문자는 사용할 수 없습니다. GENDER 및 M_OR_F는 올바른 출력형식 이름이지만 M-OR-F 또는 M#F는 올바르지 않은 출력형식 이름입니다. ■ 출력형식 이름은 숫자로 시작하거나 끝날 수 없습니다. HEX2DEC는 올바른 출력형식 이름이지만 4MAT는 올바르지 않은 출력형식 이름입니다. ■ 출력형식 유형에 따라 출력형식 이름의 최대 길이가 정해져 있습니다. 문자 출력형식의 이름은 31자를 넘을 수 없습니다. 숫자 출력형식의 이름은 32자를 넘을 수 없습니다. 주: SAS 9 이전의 SAS 버전을 실행 중이라면 문자 출력형식의 이름은 7자, 숫자 출력형식의 이름은 8자를 넘을 수 없습니다. 출력형식 이름이 더 길면 예기치 못한 결과를 초래할 수 있습니다. MALEORFEMALE은 너무 길어서 두 유형의 올바른 출력형식 이름이 될 수 없습니다. STATCODE는 숫자 출력형식의 올바른 이름이지만 문자 출력형식의 이름이 되기에는 너무 길습니다.
출력형식 너비 지정	출력형식의 너비를 지정합니다. 기본적으로 출력형식 너비는 서식화된 가장 긴 값의 길이가 됩니다.
출력형식 저장 위치	새로운 출력형식을 저장할 서버 및 라이브러리를 지정합니다. 기본적으로 출력형식은 선택한 서버에 EGTASK 라이브러리가 있으면 그 곳에 저장됩니다. EGTASK 라이브러리가 없으면 출력형식은 SASUSER 라이브러리에 저장됩니다. 어느 때나 출력형식은 종료 시에도 남아 다음에 사용할 수 있습니다. 주: WORK 라이브러리에 출력형식을 저장하도록 선택하면 종료 시 출력형식이 삭제됩니다. 출력형식이 존재하더라도 사용하지 못할 때가 있습니다. 출력형식이 저장된 라이브러리를 출력형식 검색 경로에 지정해야만 해당 출력형식을 사용할 수 있습니다.

옵션 이름	설명
출력형식 유형	생성할 출력형식이 문자 출력형식인지 숫자 출력형식인지를 지정합니다. 부동 소수를 저장할 때 발생하는 부정확성을 방지하기 위해 숫자 출력형식은 퍼지 인수 상자에 있는 퍼지 인수와 함께 적용됩니다. 퍼지 인수 기본값은 1E-12 입니다. 퍼지 인수를 변경하려면 퍼지 인수 지정 체크 박스를 선택하고 상자에 새로운 값을 지정합니다.

출력형식 생성: 새로운 출력형식 정의

새로운 출력형식을 정의하는 방법

이 페이지에 액세스하려면 선택 영역에서 **출력형식 정의**를 클릭합니다.

이 페이지를 사용하여 변수의 저장된 값을 해당 값을 표시하거나 인쇄할 때 나타날 텍스트로 매핑합니다. 출력형식 정의에는 두 부분이 있습니다. **레이블** 칼럼은 표시되거나 인쇄될 텍스트를 지정하고 **범위** 칼럼은 해당 텍스트로 번역될 하나 이상의 저장된 값 범위를 지정합니다. 특정 범위에 대해 빨간색 박스가 표시되면 해당 범위와 다른 범위 간에 충돌이 있음을 나타냅니다. 예를 들어 0과 5 사이의 모든 숫자를 LOW 레이블에 매핑하고 3과 7 사이의 모든 숫자를 레이블 MEDIUM에 매핑하면 충돌이 발생합니다.

새로운 출력형식 정의 방법

- 1 **출력형식 정의** 상자 옆의 **새로 만들기**를 클릭합니다. 새로운 행은 **출력형식 정의** 및 **범위 정의** 상자에 추가됩니다.
- 2 **레이블** 칼럼에서 표시되거나 인쇄될 텍스트를 입력합니다.
- 3 **범위 정의** 상자에서 해당 텍스트로 매핑될 저장된 값을 지정합니다. 범위 정의 지정 방법:
 - a 범위 정의에 대한 유형을 선택합니다(**이산** 또는 **범위**).
 - b 값을 입력하여 출력형식 정의의 레이블로 매핑합니다. 각 이산 값 또는 범위를 별도 항목으로 입력해야 합니다.

예 1: 문자 출력형식

다음 네 개의 저장된 값이 있는 GENDER 변수에 대한 문자 출력형식을 정의하려고 합니다. f, F, m, M m 및 M 값을 저장했습니다. **출력형식 정의** 페이지는 다음 표시와 유사해야 합니다.

출력형식 생성

출력형식 정의

출력형식 정의:

레이블	범위
여성	"f" "F"
남성	"m" "M"

새로 만들기(N)
삭제(D)

"남성"에 대한 범위 정의:

유형	값
이산	"m"
이산	M

새로 만들기(W)
삭제(L)

코드 미리 보기(C) 실행(R) 저장(S) 취소 도움말

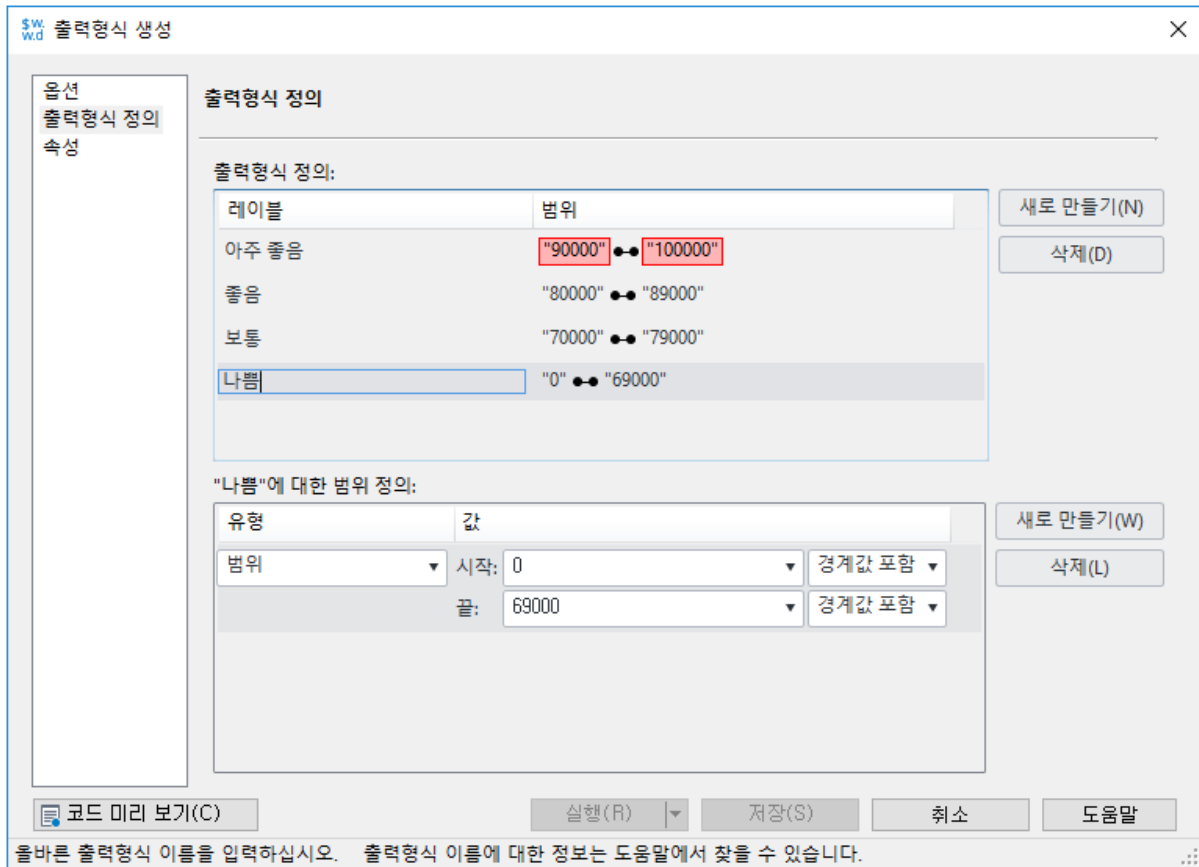
올바른 출력형식 이름을 입력하십시오. 출력형식 이름에 대한 정보는 도움말에서 찾을 수 있습니다.

예 2: 숫자 출력형식

이름이 SALES인 변수에 대해 숫자 출력형식을 정의하고 있습니다. 판매 값의 범위에 기반하여 소프트웨어에서 등급을 표시하려고 합니다.

- Excellent = 90,000–100,000 (inclusive)
- Good = 80,000–89,000 (inclusive)
- Fair = 70,000–79,000 (inclusive)
- Poor = 0–69,000 (inclusive)

출력형식 정의 페이지는 다음과 유사한 모양이어야 합니다.



속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

데이터셋에서 출력형식 생성

데이터셋에서 출력형식 생성 작업 정보	129
데이터셋에서 출력형식 생성	130

데이터셋에서 출력형식 생성 작업 정보

데이터셋에서 출력형식 생성 작업을 수행하면 SAS 데이터셋에 저장된 데이터를 사용하여 SAS 출력형식을 만들 수 있습니다. SAS 출력형식을 사용하면 원시 데이터 값을 출력형식이 지정된 값에 매핑할 수 있습니다. 문자 또는 숫자 데이터에 대한 SAS 출력형식을 생성할 수 있습니다.

다음은 데이터셋에서 출력형식 생성 작업을 사용해야 하는 경우의 예입니다.

- 다음 해에 조직은 미국의 각 지역(북부, 남부, 동부, 서부)에서 고객 포커스 그룹을 수행할 계획입니다. 2009년 판매 데이터에 각 고객의 주소와 시/도(외국: 주)가 나열됩니다. 그러나 사용자는 이러한 판매 데이터를 지역별로 분류하려고 합니다. 지역 데이터셋에서는 사용자가 각 주를 해당하는 지역에 이미 매핑해 놓았습니다.

주	지역
아리조나	서부
캘리포니아	서부
플로리다	남부
버몬트	북부
버지니아	동부

데이터셋에서 출력형식 생성 작업에서 지역 데이터셋을 입력값으로 사용하면, 나중에 판매 데이터의 주 값에 적용할 수 있는 SAS 출력형식을 만들 수 있습니다.

- 최근에 회사에서 여러 임상 시험을 수행했습니다. 환자 데이터에는 각 시험 참여자의 연령이 포함됩니다. AGES 데이터셋에서 연령 범위를 이미 이산 범주로 나누었습니다:

Age	낮음	높음
영아	0	2
유아	3	4
어린이	5	9
아동	10	12
10대	13	17
성인	18	100

주: 이 작업은 **SAS Enterprise Guide**에서만 사용 가능합니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	FORMAT
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	없음

데이터셋에서 출력형식 생성

- 1 입력 데이터셋을 선택합니다. 기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름이 **소스 데이터셋** 필드에 나타납니다.
- 2 출력형식 이름은 다음 규칙을 따라야 합니다.
 - 출력형식 이름에는 영문자, 숫자 및 밑줄(_)만 포함될 수 있습니다.
 - 다른 문자는 사용할 수 없습니다. GENDER 및 M_OR_F는 올바른 출력형식 이름이지만 M-OR-F 또는 M#F는 올바르지 않은 출력형식 이름입니다.
 - 출력형식 이름은 숫자로 시작하거나 끝날 수 없습니다. HEX2DEC는 올바른 출력형식 이름이지만 4MAT는 올바르지 않은 출력형식 이름입니다.

- 출력형식 유형에 따라 출력형식 이름의 최대 길이가 정해져 있습니다. 문자 출력형식의 이름은 31자를 넘을 수 없습니다. 숫자 출력형식의 이름은 32자를 넘을 수 없습니다.

주: SAS 9 이전의 SAS 버전을 실행 중이라면 문자 출력형식의 이름은 7자, 숫자 출력형식의 이름은 8자를 넘을 수 없습니다. 출력형식 이름이 더 길면 예기치 못한 결과를 초래할 수 있습니다. MALEORFEMALE은 너무 길어서 두 유형의 올바른 출력형식 이름이 될 수 없습니다. STATCODE는 숫자 출력형식의 올바른 이름이지만 문자 출력형식의 이름이 되기에는 너무 깁니다.

- 3 생성할 출력형식이 문자 출력형식인지 숫자 출력형식인지를 지정합니다. 부동 소수를 저장할 때 발생하는 부정확성을 방지하기 위해 숫자 출력형식은 퍼지 인수 상자에 있는 퍼지 인수와 함께 적용됩니다. 퍼지 인수 기본값은 **1E-12**입니다.
- 4 새로운 출력형식을 저장할 라이브러리를 선택합니다. 기본적으로 출력형식은 선택한 서버의 Work 라이브러리에 저장됩니다. 그러나 Work 라이브러리에 출력형식을 저장하면 **SAS Enterprise Guide** 종료 시 출력형식이 삭제됩니다. 나중에 출력형식을 사용할 수 있도록 하려면 다른 출력 라이브러리를 선택합니다.

주: 출력형식이 존재하더라도 사용하지 못할 때가 있습니다. 출력형식이 저장된 라이브러리를 출력형식 검색 경로에 지정해야만 해당 출력형식을 사용할 수 있습니다.

- 5 입력 데이터셋의 값이 이산 값인지 값 범위인지 지정합니다.

이산 값에서 출력형식을 생성하는 방법:

- a **값 유형** 상자에서 **이산/검색**을 선택합니다.
- b **이산 값** 드롭다운 리스트의 **변수** 영역에서 출력형식을 지정하려는 값이 포함된 데이터셋 내에서 변수를 선택합니다. 선택한 변수의 값은 미리 보기 테이블의 **검색 값** 칼럼에 표시됩니다.
- c **레이블** 드롭다운 리스트에서 값에 대한 레이블이 포함된 데이터셋 내에서 변수를 선택합니다. 선택한 변수의 값은 미리 보기 테이블의 **레이블** 칼럼에 표시됩니다.

값 범위에서 출력형식을 생성하는 방법:

- a **값 유형** 상자에서 **범위**를 선택합니다.
 - b **하한 값** 드롭다운 리스트의 **변수** 영역에서 범위에 대한 초깃값이 포함된 데이터셋 내에서 변수를 선택합니다. 선택한 변수의 값은 미리 보기 테이블의 **시작** 칼럼에 표시됩니다.
 - c **상한 값** 드롭다운 리스트에서 범위에 대한 마지막 값이 포함된 데이터셋 내에서 변수를 선택합니다. 선택한 변수의 값은 미리 보기 테이블의 **끝** 칼럼에 표시됩니다.
 - d **레이블** 드롭다운 리스트에서 값에 대한 레이블이 포함된 데이터셋 내에서 변수를 선택합니다. 선택한 변수의 값은 미리 보기 테이블의 **레이블** 칼럼에 표시됩니다.
- 6 레이블의 최대 길이를 지정합니다. 레이블의 최대 길이 지정 기본적으로 레이블의 길이는 8자까지 사용할 수 있습니다.
 - 7 다른 값에 대한 레이블을 지정합니다. 이 출력형식을 데이터 소스에 적용하면 **SAS Enterprise Guide**가 데이터의 일부 값이 출력형식에서 설명되지 않음을 발견할 수 있

습니다. 데이터 소스의 모든 값에 대해 고유한 레이블을 생성할 수 있으며 여기에는 해당하는 서식화된 값이 없습니다. 기본적으로 이러한 값의 레이블은 **기타**입니다.

- 8 결과에 요약 리포트를 포함할지를 지정합니다. 기본적으로 요약 리포트는 작업 실행 시에는 생성되지 않습니다.

지도용 테이블 생성

지도용 테이블 생성 작업 정보	133
지도용 테이블 생성: 분석 역할에 변수 할당	134
지도용 테이블 생성: 옵션 설정	135
지도 데이터 투영 옵션 설정	135
지도용 테이블 옵션 설정	135
지도용 테이블 생성: 결과 지정	136
제목 및 각주 지정	136
속성 보기	136

지도용 테이블 생성 작업 정보

지도용 테이블 생성 작업은 지도 그래프 작업에 사용하는 사용자 정의 또는 기타 지도 데이터셋에 대한 지도용 테이블을 생성합니다. 또한 지도 데이터셋을 투영할 수 있습니다. 즉, 지도 데이터셋의 구면 좌표를 데카르트 좌표로 변환합니다. 이 작업에서는 여러 가지 지도 투영 기법 중 하나를 사용하여 왜곡된 부분, 거리, 방향 및 원래 구형의 모양 속성을 최소화하면서 전통적인 지도 데이터셋의 좌표를 2차원 평면으로 투영할 수 있습니다.

주: 이 작업에는 SAS 릴리스 8.2 이상 릴리스와 함께 제공되는 %ADDFEAT 매크로가 필요합니다. 이 매크로를 구하는 방법은 SAS 기술 지원부로 문의하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GPROJECT
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	MAPS

지도용 테이블 생성: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

주: 이 작업에 대한 입력 데이터 소스에는 지도에 대한 공간 데이터 좌표를 포함하는 X 및 Y 칼럼이 필요합니다. MAPS 라이브러리에는 연관된 지도용 테이블이 있는 지도 데이터셋이 포함되어 있습니다. 예를 들어 US2 데이터셋은 미국의 지도용 테이블입니다.

MAPS 라이브러리는 보호되므로 이 라이브러리에 있는 데이터셋은 지도용 테이블 생성 작업의 입력 데이터로 사용할 수 없습니다. 그러나 MAPS 라이브러리의 데이터셋을 보호되지 않는 위치로 복사하면 이 작업에서 이 데이터셋을 사용할 수 있습니다.

역할 이름	설명
표시 변수	이 역할에 지정한 변수는 지도 내의 섹션을 표시하는 데 사용됩니다. 데이터를 투영하고 지도용 테이블을 생성하려면 최소한 한 개의 표시 변수가 필요합니다.
유지할 속성 변수	생성할 지도용 테이블에 포함할 입력 지도(geometry) 데이터셋의 변수입니다. 주: 이 데이터 역할은 지도용 테이블 생성 옵션을 선택할 때만 나타납니다.
투영할 경도 변수	투영되지 않은(구형) 경도 데이터 포인트를 포함하는 입력 데이터셋의 변수입니다. 주: 이 데이터 역할은 지도 데이터 투영(구형 좌표를 평면 좌표로 바꾸기) 옵션을 선택할 때만 나타납니다.
투영할 위도 변수	투영되지 않은(구형) 위도 데이터 포인트를 포함하는 입력 데이터셋의 변수입니다. 주: 이 데이터 역할은 지도 데이터 투영(구형 좌표를 평면 좌표로 바꾸기) 옵션을 선택할 때만 나타납니다.

지도용 테이블 생성: 옵션 설정

지도 데이터 투영 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

- 1 **지도 데이터 투영(구형 좌표를 평면 좌표로 바꾸기)** 체크 박스를 선택합니다.
- 2 **투영 방법** 드롭다운 리스트에서 적절한 투영 방법을 선택합니다.
 - **Albers의 동등 영역 투영** - 구 표면에서 구의 원추 시컨트까지의 원추 투영으로, 위도의 두 개의 표준 평행선에서 자릅니다. 원뿔의 축은 구의 극축을 확장한 선과 일치합니다. 생성된 지도의 각 섹션은 구의 영역에 대해 일정비를 유지합니다. 일반적으로 두 표준위선을 벗어난 위도 상의 극으로 갈수록 모양의 왜곡이 심해지는 경향이 있습니다. Albers의 투영은 동서의 범위가 좁은 영역을 표시하는 데 적합합니다. 두 개의 표준위선은 동일한 적도선상에 있어야 합니다.
 - **Lambert의 정각 투영** - 경도의 자오선이 시컨트 원추의 꼭짓점에서 방사되는 직선이고 위도의 평행이 동심원인 투영입니다. 동서 방향이 비교적 작은 지도 및 항법 지도에 이상적입니다. 두 개의 표준위선은 동일한 적도선상에 있어야 합니다.
 - **Gnomic 투영(방위각)** - 지도 투영 극에서 구에 접하는 가상의 평면에 구 표면을 직접 투영하는 평면 투영입니다. 기본적으로 투영축은 투영할 지도 데이터셋의 중심에 놓여집니다. 지도축에서의 거리가 멀어질수록 왜곡이 심해집니다. 이 투영법은 동서의 범위가 좁은 영역을 지도화하는 데 가장 적합합니다.
- 3 입력 데이터셋의 비투영 경도 및 위도 값에 대한 **좌표 단위**를 선택합니다. 기본값은 라디안입니다.
- 4 **경도의 증가 방향**을 선택합니다. 기본값은 서쪽입니다.
- 5 X 및 Y 변수의 투영 값이 이전 관측값의 해당 값과 동일한 모든 관측값을 유지하려면 **연속하는 동일 관측값 유지** 체크 박스를 선택합니다. 기본적으로 연속하는 동일 관측값은 삭제됩니다.

지도용 테이블 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

입력 데이터셋과 관련된 새 지도용 테이블을 생성하는 방법:

- 1 **지도용 테이블 생성** 체크 박스를 선택합니다.

- 2 map geometry 변수의 이름을 변경하려면 **새로운 map geometry 변수** 이름 상자에 이름을 지정합니다. 이름은 숫자로 시작할 수 없으며 공백, 구두점 또는 특수 문자가 포함될 수 없습니다.

지도용 테이블 생성: 결과 지정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

기본적으로 이 작업에서는 사용자의 SASUSER 라이브러리에 프로젝트 지도와 지도용 테이블 출력 데이터셋을 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **다른 이름으로 저장** 대화 상자를 사용합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버

- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

시계열 데이터 생성 마법사

시계열 데이터 생성 마법사 정보	139
시계열 데이터 생성 마법사: 데이터 선택	139
시계열 데이터 생성 마법사: 역할에 변수 할당	140
시계열 데이터 생성 마법사: 누적, 결측값 및 영(0) 값에 대한 옵션 지정	140
시계열 데이터 생성 마법사: 추가 출력 선택	141
시계열 데이터 생성 마법사: 제목 및 각주 지정	141

시계열 데이터 생성 마법사 정보

시계열 데이터 생성 마법사를 사용하여 트랜잭션 데이터를 고정-간격 시계열로 변환할 수 있습니다. 트랜잭션 데이터는 비정기적이거나 다양한 빈도를 가진 시간에 대해 수집된 타임스탬프 데이터입니다.

예를 들어 회사 웹사이트의 판매 데이터를 가지고 있다고 가정합니다. 이러한 판매는 타임스탬프이지만 특정 구간에서 수집되지는 않았습니다. 이 작업을 사용하여 시계열 데이터셋으로 데이터를 변환할 수 있습니다.

시계열 데이터 생성 마법사: 데이터 선택

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

다음을 클릭하여 역할에 변수를 할당합니다.

시계열 데이터 생성 마법사: 역할에 변수 할당

- 1 데이터를 선택해야 역할에 변수를 할당할 수 있습니다.
- 2 다음과 같은 필요한 역할에 변수를 할당합니다.
 - **시간 ID** - 이 역할에 지정한 칼럼은 데이터에 대한 시간 ID입니다. 시간 ID 변수는 각 관측값의 SAS 날짜, 시간 또는 날짜/시간 값을 포함하는 입력 데이터셋의 변수입니다. 이 변수는 데이터의 빈도 및 순서를 결정하고 예측치에 대해 시간 ID 값을 추정하는 데 사용됩니다. 이 역할에 한 개의 변수만 할당할 수 있으며, 날짜 변수, 날짜/시간 변수 또는 날짜나 날짜/시간 값을 포함하는 숫자 변수여야 합니다.
 - **시계열 변수** - 이 역할에 지정한 칼럼은 모델화하고 예측할 변수입니다. 이 역할에는 한 개 이상의 숫자 변수를 할당해야 합니다. 위쪽/아래쪽 화살표를 사용하여 변수 순서를 지정합니다.
- 3 (선택 사항) 한 개의 변수를 **그룹 변수(이 변수의 값에 따라 별도의 시계열 생성됨)** 역할에 지정합니다. 이 변수는 그룹별 변수에 대한 동일한 값을 가진 관측값으로 함께 그룹화됩니다. 그룹별 변수를 지정하면 관측값 그룹에 대한 별도의 분석을 얻을 수 있습니다. 문자와 숫자 변수를 이 역할에 할당할 수 있습니다.

BY 변수의 순서는 계층 구조를 설명합니다. 위쪽/아래쪽 화살표를 사용하여 변수 순서를 지정합니다.
- 4 **다음**을 클릭하여 누적, 결측값 및 0 값에 대한 옵션을 지정합니다.

시계열 데이터 생성 마법사: 누적, 결측값 및 영(0) 값에 대한 옵션 지정

- 1 역할에 변수를 할당해야 다음과 같은 옵션을 지정할 수 있습니다.
- 2 작업에서 입력 데이터셋의 구간을 발견하면, 기본 구간이 **구간** 드롭다운 리스트에 나타납니다. 구간을 변경할 수 있지만, 드롭다운 리스트에는 **시간 ID** 역할에 지정한 변수에 적합한 구간만 포함됩니다.

큰 데이터 소스이면 작업이 구간을 탐지하는 데 오랜 시간이 걸립니다. 구간이 이 단계에 오는 데 걸린 시간까지 탐지되지 않으면, **구간** 드롭다운 리스트에서 구간을 선택할 수 있습니다. 구간이 탐지될 때까지 구간 옵션에 대한 모든 가능한 값이 나열됩니다. 구간이 결정된 후, 시간 ID 역할에 지정한 변수 유형과 구간값이 충돌을 일으키면 경고 메시지가

표시됩니다. 구간값 리스트는 변수 유형에 알맞은 값으로만 변경됩니다. 각 통계량에 대한 자세한 내용은 통계학 용어 설명을 참조하십시오.

- 3 시계열 영역에서 **시계열** 변수를 선택합니다. 이러한 변수는 **시계열 변수** 역할에 지정한 변수입니다.
- 4 **누적통계량** 드롭다운 리스트에서 계층 레벨에서 데이터 집계에 대한 통계량을 지정합니다.
- 5 **결측값 처리 값** 드롭다운 리스트에서 결측값으로 취급할 방법을 지정합니다.
- 6 각 시계열변수에 대해 3-5단계를 반복합니다.
- 7 **다음**을 클릭하여 출력 옵션을 지정합니다.

시계열 데이터 생성 마법사: 추가 출력 선택

- 1 누적, 결측값 및 0 값에 대한 옵션을 지정해야 출력 옵션을 지정할 수 있습니다.
- 2 출력에 시계열 데이터의 그래프에 포함하려면 **시계열 그래프** 체크 박스를 선택합니다.
- 3 **데이터셋에 시계열 저장** 체크 박스를 선택합니다.

주: SAS Enterprise Guide는 **옵션 창**의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

- 4 **다음**을 클릭하여 제목과 각주를 지정합니다.

시계열 데이터 생성 마법사: 제목 및 각주 지정

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

- 1 출력 옵션을 지정해야 제목과 주석을 지정할 수 있습니다.
- 2 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다. 제목과 각주에는 매크로 변수를 사용할 수 있습니다.
- 3 **마침**을 클릭하여 작업을 실행합니다.

데이터셋 속성

데이터셋 속성 마법사 정보	143
데이터셋 속성: 데이터 선택	143
데이터셋 속성: 출력 지정	144

데이터셋 속성 마법사 정보

데이터셋 속성 마법사를 사용하여 데이터셋 생성일, 위치, 관측값 수, 변수 이름, 레이블, 유형, 출력형식 등의 리포트를 생성할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	DATASETS
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	없음

데이터셋 속성: 데이터 선택

- 1 **추가**를 클릭합니다. **열기** 대화 상자가 나타납니다.
- 2 사용하려는 데이터셋을 선택합니다. 여러 데이터셋을 선택하려면 Ctrl 키를 누릅니다. **열기**를 클릭합니다. 선택한 데이터셋이 **데이터** 상자에 나타납니다.

- 3 다음을 클릭하여 출력을 지정합니다.

데이터셋 속성: 출력 지정

- 1 출력 옵션을 지정하려면 먼저 데이터를 선택해야 합니다.
- 2 출력 데이터셋의 위치를 지정합니다. 기본적으로 출력 데이터셋은 SASUSER 라이브러리에 저장됩니다. 출력 데이터셋의 위치 또는 이름을 변경하려면 **찾아보기**를 클릭합니다. 결과 데이터셋도 프로젝트에 추가됩니다.
- 3 출력 유형을 선택합니다. 다음 옵션 중에서 선택할 수 있습니다.
 - **기본 리포트**에는 DATASETS 프로시저의 출력이 포함됩니다. 이 리포트에는 데이터셋이 생성되고 마지막으로 수정된 날짜, 관측값 수, 인코딩, 엔진/호스트 종속 정보, 변수 및 해당 속성의 알파벳순 리스트 등의 데이터 속성이 포함되어 있습니다.
 - **확장 리포트**에는 REPORT 프로시저의 출력이 포함됩니다. 이 리포트에는 테이블 및 변수 속성이 표시됩니다. 이 리포트에서 테이블 유형, 테이블 생성일과 수정일, 관측값 수, 변수 레이블 및 변수 유형을 확인할 수 있습니다. 변수 테이블의 행을 다음 옵션을 기준으로 하여 오름차순 또는 내림차순으로 정렬할 수 있습니다.
 - ☐ 변수 이름
 - ☐ 테이블에서의 변수 순서
 - ☐ 변수 유형
 - ☐ 변수 출력형식
 - ☐ 변수 레이블
- 4 **마침**을 클릭하여 출력을 생성합니다.

데이터셋 및 출력형식 삭제

데이터셋 및 출력형식 삭제 작업 정보	145
데이터셋, 뷰 및 사용자 정의 출력형식 삭제	146
속성 보기	146

데이터셋 및 출력형식 삭제 작업 정보

데이터셋 및 출력형식 삭제 작업을 사용하여 데이터셋, 뷰 및 사용자 정의 출력형식을 영구적으로 삭제할 수 있습니다. 서로 다른 여러 라이브러리와 카탈로그에서 한 번에 여러 항목을 삭제할 수 있습니다.

주: 이 작업은 SAS Enterprise Guide에서만 사용할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	CATALOG, SQL
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	없음

데이터셋, 뷰 및 사용자 정의 출력형식 삭제

- 1 **서버** 드롭다운 리스트에서 데이터셋, 뷰 및 사용자 정의 출력형식이 저장된 서버를 선택합니다.
- 2 **라이브러리 및 카탈로그** 영역에서 라이브러리 또는 카탈로그를 선택합니다. 선택한 라이브러리나 카탈로그의 데이터셋, 뷰 또는 사용자 정의 출력형식이 리스트 뷰에 나타납니다. 기본적으로 이 항목은 이름을 기준으로 정렬됩니다. 칼럼 머리를 클릭하여 수정일 또는 유형을 기준으로 항목을 정렬합니다.
- 3 삭제하려는 데이터셋, 뷰 또는 사용자 정의 출력형식의 체크 박스를 하나 이상 선택합니다. 노트

주: 라이브러리 및 카탈로그 영역 아래에서 선택한 총 데이터셋, 뷰, 사용자 정의 출력형식 수와 현재 폴더에서 선택한 항목 수를 확인할 수 있습니다.

- 4 삭제하려는 모든 항목을 선택할 때까지 2, 3단계를 반복합니다.
- 5 **실행**을 클릭하여 선택한 항목을 영구적으로 삭제합니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

판별분석

<i>판별분석 작업 정보</i>	149
<i>판별분석: 분석 역할에 변수 할당</i>	150
<i>판별분석: 분석 옵션 설정</i>	151
<i>판별분석: 새로운 데이터 분류</i>	152
데이터 입력 테이블 지정	152
검정 변수 지정	153
<i>판별분석: 결과 옵션 설정</i>	154
<i>제목 및 각주 지정</i>	154
<i>속성 보기</i>	155

판별분석 작업 정보

관측값 그룹을 정의하는 양적변수 및 분류변수가 하나 이상 포함된 관측값 집합이면, 판별분석 작업을 실행하면 각 관측값을 그룹의 관측값으로 분류할 수 있는 판별 기준이 만들어집니다. 이 데이터셋에서 파생된 판별 기준은 판별함수 실행 시 두 번째 데이터셋에도 적용될 수 있습니다.

이 작업을 사용하여 자연 현상과 관련된 관측값을 분류할 수 있습니다. 예를 들어 159개 종류의 물고기에 대한 무게, 길이, 너비 등을 포함하는 SAS 데이터셋을 생성합니다. 이 데이터를 사용하여 6개의 변수를 기반으로 하고 각 물고기를 7개 중 하나로 가장 잘 분류하는 판별 함수를 찾을 수 있습니다.

주: 이 작업은 SAS Enterprise Guide에서만 사용할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	DISCRIM

요구 사항 이름	프로시저 및 제품 이름
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

판별분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분석변수	리스트의 각 변수에 대해 분석을 수행합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
분류변수	분석을 위한 그룹을 정의합니다. 분류 레벨은 이 변수의 서식화된 값에 의해 결정됩니다. 이 역할에는 변수를 하나만 할당해야 합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석변수 또는 분류변수 역할로 이미 선택한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
ID 레이블	관측값 번호 대신 분류변수의 값을 표시합니다. 이 역할에 대한 변수 할당은 모든 관측값에 대한 재대입 분류 결과 체크 박스를 선택할 때만 유효합니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.
상대 가중값 변수	분석변수의 해당 값에 가중값을 적용하는 데 이 변수의 값을 사용합니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.

판별분석: 분석 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

옵션 이름	설명
교차타당성 분류의 요약 결과	데이터셋 내의 다른 관측값에서 계산된 판별함수를 사용하여 입력 데이터셋의 각 관측값을 분류하며 분류되는 관측값은 제외합니다. 이 옵션은 기본값인 모수적 방법이 사용될 때만 적용됩니다.
각 관측값에 대한 교차타당성 분류 결과 표시	각 관측값에 대한 교차타당성 분류 결과를 표시합니다. 주: 이 옵션은 교차타당성 분류의 요약 결과 체크 박스를 선택할 때만 사용할 수 있습니다.
모든 관측값에 대한 재대입 분류 결과	각 관측값에 대한 재대입 분류 결과를 표시합니다.
사후확률 오류율 추정 값	분류 결과를 기준으로 분류 기준의 사후확률 오류율 추정값을 표시합니다.
계급 평균의 동일함에 대한 일변량검정	각 변수 모집단 내의 계급 평균이 동일하다는 가설을 검정하는 데 사용되는 일변량 통계량을 표시합니다.
계급 평균의 동일함에 대한 다변량검정	각 변수 모집단 내의 계급 평균이 동일하다는 가설을 검정하는 데 사용되는 다변량 통계량을 표시합니다.
재대입 분류 결과 표시	잘못 분류된 관측값에 대한 재대입 분류 결과만 표시합니다. 이 옵션은 입력 데이터셋이 일반 SAS 데이터셋일 때만 지정할 수 있습니다.
교차타당성 분류 결과 표시	잘못 분류된 관측값에 대한 교차타당성 분류 결과만 표시합니다.
정준판별분석 수행	정준판별분석을 수행합니다.
계산할 정준변수 개수 제한	계산할 정준변수 개수를 지정합니다. 지정하는 개수는 변수의 개수보다 적거나 동일해야 합니다. 이 옵션을 0으로 설정하면 출력에는 정준상관분석이 포함되며 정준계수, 구조 또는 평균은 포함되지 않습니다. v 를 분석변수의 개수로 하고 c 를 분류변수의 개수로 합니다. 계산할 정준변수의 개수를 제한하지 않으면 $\min(v, c - 1)$ 정준변수만 생성됩니다. 결과에 출력 데이터셋을 포함하면 v 정준변수가 생성됩니다. 이때 마지막 $v - (c - 1)$ 정준변수에 결측값이 생깁니다.

옵션 이름	설명
사전 확률	사전 확률 계산 방법을 지정합니다. 다음과 같은 방법을 사용할 수 있습니다. ■ 같은 값으로 지정 - 그룹 멤버십의 사전 확률이 동일하도록 지정합니다. ■ 표본크기에 비례해서 지정 - 그룹 멤버십의 사전 확률이 표본크기에 비례하도록 지정합니다. ■ 각 레벨에 대한 확률 입력 - 분류변수의 각 레벨에 대한 사전 확률을 지정할 수 있습니다. 확률의 합이 1이 아니면 1이 되도록 각 확률값이 적당한 비율로 조정됩니다. 주: 지정한 확률의 합은 1이 되어야 합니다. 확률값을 원래의 값으로 재설정하려면 확률값 초기화를 클릭합니다.

판별분석: 새로운 데이터 분류

데이터 입력 테이블 지정

이 옵션에 액세스하려면 선택 영역에서 **새로운 데이터 분류**를 클릭합니다.

분석하려는 테이블에서 유도된 판별 기준에 따라 분류될 행을 포함하는 테이블(여기서는 "입력 테이블")을 선택할 수 있습니다. 입력 테이블에는 **분석변수** 역할에 할당된 모든 변수가 포함되어야 합니다.

기본적으로 입력 테이블의 분류 결과는 출력 결과에 포함됩니다. 입력 테이블은 작업을 열 때 사용하는 입력 데이터입니다.

다른 입력 테이블을 지정하려면 **분류할 새로운 데이터** 체크 박스를 선택하고 **찾아보기**를 클릭합니다.

GROUP BY 변수 역할에 변수를 할당하고 원래 입력 데이터의 대응 BY 그룹의 판별 함수로 새로운 데이터의 각 BY 그룹을 분류하려면, 선택한 새로운 데이터는 작업을 여는 데 사용한 입력 데이터와 동일한 순서로 정렬되어야 합니다.

주: **분류할 새로운 데이터** 체크 박스를 활성화하려면 **분류변수** 역할에 변수를 할당해야 합니다.

주: 입력 테이블을 지정하거나 검정 변수를 지정하려면 **분류할 새로운 데이터** 체크 박스를 선택해야 합니다.

입력 테이블 옵션을 설정하려면 다음 옵션을 사용하십시오.

- 기본적으로 입력 테이블의 분류 결과는 출력 결과에 포함됩니다. 입력 테이블 분류 결과가 출력되지 않게 하려면 **새로운 데이터의 분류 결과 표시** 체크 박스의 선택을 취소합니다.
- 입력 테이블에서 얻은 모든 데이터, 사후 확률 및 각 행이 분류된 클래스를 포함하는 영구 출력 테이블을 생성하려면 **새로운 데이터의 분류 결과에 대한 출력 테이블**을 선택합니다.
- 각 행에 대한 그룹별 밀도함수 추정값 외에 입력 테이블의 모든 데이터가 포함된 영구 출력 테이블을 생성하려면 **밀도함수 추정값 그룹에 대한 출력 테이블**을 선택합니다.

주: SAS Enterprise Guide는 **옵션 창의 작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

입력 테이블을 지정한 후에는 검정 변수를 지정할 수 있습니다.

검정 변수 지정

이 옵션에 액세스하려면 선택 영역에서 **새로운 데이터 분류**를 클릭합니다.

분류 데이터가 포함된 입력 테이블을 지정한 후에는 검정 변수를 지정할 수 있습니다.

검정 변수를 지정하는 방법:

- 1 **새로운 데이터의 분류변수** 드롭다운 리스트에서 변수 이름을 선택합니다. 이 변수는 입력 테이블의 관측값이 제대로 분류되었는지 확인하는 데 사용됩니다. 여기에서 지정한 변수는 **분류변수** 역할에 할당된 변수와 같은 유형(문자 또는 숫자)이어야 합니다.

여기에서 변수를 선택하지 않으면 **분류변수** 역할에 할당한 변수가 사용됩니다.

주: **분류변수** 역할에 할당한 변수가 입력 테이블에 존재하지만 원래의 데이터셋과 이름이 다를 때, 이 옵션을 사용할 수 있습니다. 예를 들어, 분류변수 역할에 할당된 변수가 'Sex'이지만 입력 테이블의 분류변수는 'Gender'입니다.

- 2 **검정 데이터의 빈도 수** 드롭다운 리스트에서 변수 이름을 선택합니다. 여기에서 선택한 변수는 검정 변수에 대해 **빈도 수** 역할에 할당된 변수와 동일한 기능을 수행합니다.
- 3 **검정 데이터의 표시 레이블(ID) 변수** 드롭다운 리스트에서 변수 이름을 선택합니다. 여기에서 선택한 변수는 검정 변수에 대해 **ID 레이블** 역할에 지정된 변수와 동일한 기능을 수행합니다. 이러한 변수 선택은 **새로운 데이터의 분류 결과 표시** 체크 박스를 선택할 때만 유효합니다.

판별분석: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
출력 데이터 저장	출력 테이블에 포함할 결과를 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 사후 확률 및 재대입 분류 테이블에는 입력 데이터, 사후 확률 및 재대입에 의한 각 관측값의 분류 결과 등이 포함됩니다. ■ 사후 확률 및 교차 타당성 분류 테이블에는 입력 데이터, 사후 확률 및 교차 타당성에 의한 각 관측값의 분류 결과 등이 포함됩니다. ■ 요약통계량 및 판별함수 테이블에는 평균, 표준편차 및 상관계수 같은 다양한 통계량이 포함됩니다. ■ 각 관측값의 특정 밀도함수 추정값 그룹 테이블에는 사용자의 입력 데이터와 각 관측값에 대한 그룹별 밀도함수 추정값이 포함됩니다. 주: SAS Enterprise Guide는 옵션 창의 작업> 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
리포트 생략	결과를 표시하지 않습니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

분포분석

분포분석 작업 정보	157
분포분석: 분석 역할에 변수 할당	158
분포분석: 분포 선택	158
분포분석: 분포 사용자 정의	159
분포분석: 도표 생성	160
분포분석: 인셋 추가	161
분포분석: 결과 옵션 지정	161
제목 및 각주 지정	162
속성 보기	163

분포분석 작업 정보

분포분석 작업에서는 숫자 변수의 분포에 대한 정보 및 데이터 요약 도구를 제공합니다. 히스토그램, 확률 도표, Q-Q 도표, 상자 도표를 비롯한 다양한 도표를 만들 수 있습니다.

이 작업을 사용하여 제품에 대한 요약통계량을 생성할 수 있습니다. 예를 들어 SAS 데이터 셋에 5,840채의 주택 담보 대출의 가치 대비 대출 비율을 저장했다고 가정합니다. 분포분석 작업을 사용하여 다음 결과를 생성할 수 있습니다.

- 적률 추정값을 포함한 요약 측도 테이블, 극 관측값 테이블
- 대출 비율 분포를 시각화하는 데 도움을 주는 히스토그램. 히스토그램은 왜도나 정상점과 같은 분포 특징을 표시합니다.
- 데이터 분포분석. 정규성을 검정하고 확률 도표와 같은 그래프를 생성합니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
----------	--------------

사용되는 SAS 프로시저	UNIVARIATE, BOXPLOT
---------------	---------------------

요구 사항 이름	프로시저 및 제품 이름
필수 SAS 제품	Base SAS, SAS/GRAPH
추가 SAS 제품(권장)	SAS/STAT

분포분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

역할 이름	설명
분석변수	분석할 숫자 변수를 지정합니다. 분석변수는 반드시 지정해야 하며, 이 역할에는 여러 개의 변수를 지정할 수 있습니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 분석 역할에 할당한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
상대 가중값 변수	상대 가중값 변수로 사용할 변수를 지정합니다. 이 역할에 지정된 변수값은 분석변수로 지정된 변수의 대응 값에 가중값을 적용하는 데 사용됩니다. 이 역할에는 변수를 하나만 할당할 수 있습니다.
분류변수	데이터를 분류 레벨로 그룹화할 때 사용되는 변수를 지정합니다(최대 2개). 히스토그램, 확률 도표 또는 Q-Q 도표 생성을 선택하면 비교 도표가 생성됩니다. 상자 도표 생성을 선택하면 나란한 상자 도표가 생성됩니다.

분포분석: 분포 선택

이 옵션에 액세스하려면 선택 영역의 분포 머리글 아래에서 **요약**을 클릭합니다.

여러 개의 이론 분포를 적합시킬 수 있습니다. 계산할 분포를 선택한 후 해당 분포의 모수값 및 선 속성을 지정하려면 선택 영역에서 분포 이름을 선택합니다.

체크 박스를 사용하여 다음 추정된 밀도 곡선을 하나 이상 선택합니다.

- **정규** - 평균(μ)과 표준편차(σ)를 가지는 정규밀도함수를 적합시킵니다.
- **로그 정규** - 척도모수(ζ), 임계모수(θ) 및 형상모수(σ)를 가지는 로그 정규 밀도함수를 적합시킵니다.
- **지수** - 척도모수(σ) 및 임계모수(θ)를 가지는 지수밀도함수를 적합시킵니다.
- **와이블** - 척도모수(σ), 임계모수(θ) 및 형상모수(c)를 가지는 3-모수 와이블 밀도함수를 적합시킵니다.
- **베타** - 척도모수(σ), 임계모수(θ) 및 형상모수(α 및 β)를 가지는 베타 밀도함수를 적합시킵니다.
- **감마** - 척도모수(σ), 임계모수(θ) 및 형상모수(α)를 가지는 감마 밀도함수를 적합시킵니다.
- **커널** - 비모수적 커널 밀도함수 추정값을 적합시킵니다.

출력에 사용할 그래프 스타일을 선택합니다. 기존 SAS 그래프나 ODS 통계 그래프 중에 선택할 수 있습니다. ODS 그래프에 대한 자세한 내용은 [SAS Output Delivery System: User's Guide](#)를 참조하십시오.

분포분석: 분포 사용자 정의

이 옵션에 액세스하려면 선택 영역의 분포 머리글 아래에서 사용자 정의할 분포 이름을 클릭합니다.

주: 이러한 분포 옵션은 **요약** 페이지에서 분포를 선택했거나 해당 분포 페이지에서 **분포 이름** 체크 박스를 선택할 때만 사용할 수 있습니다.

정규, 로그 정규, 지수, 와이블, 베타 및 감마 분포에 대해 분포모수를 추정 또는 지정하도록 선택할 수 있습니다. 기본적으로 작업에서는 모든 분석변수의 모수 값을 추정합니다.

각 분석변수에 대한 모수를 지정하는 방법:

- 1 **모든 변수에 분포 적용** 체크 박스의 선택을 해제합니다.
- 2 **분석변수** 상자에서 변수에 해당되는 체크 박스를 선택합니다.
- 3 모수값을 지정합니다.

주: 추정 모수값을 생성하려면 **추정값 사용**을 클릭합니다.

비모수적 커널 밀도함수 추정값에 대해 대역폭(c) 및 커널 함수(k)를 지정할 수 있습니다. 기본적으로 이 작업은 근사 평균적분제곱오차(MISE)를 최소화하는 대역폭을 추정합니다. 이 대역폭을 추정하지 않으려면 **MISE 최소화** 체크 박스에 대한 선택을 취소합니다. 커널 밀도에 대해 표준화된 대역폭 모수를 지정하려면 텍스트 상자에 대역폭을 입력합니다. 각 값 사

이에 공백을 넣어 최대 다섯 개의 값을 지정할 수 있습니다. 예를 들어 텍스트 상자에 **0.5 1.0 1.5** 값을 입력합니다.

주: MISE 최소화 체크 박스를 선택 취소하지 않으면 MISE는 지정할 수 있는 다섯 개의 대역폭 값 중 하나를 선택합니다.

커널 밀도 추정값을 계산하는 데 사용되는 커널 함수(정규, 2차 또는 3차)를 지정할 수도 있습니다. 기본 커널 함수는 **정규**입니다.

최대 다섯 개의 추정값에 대해 커널 함수를 지정할 수 있습니다. 하나 이상의 커널 함수를 지정하면 함수의 순서가 항상 정규, 2차, 3차 순으로 나열됩니다. 함수를 추가하거나 순서를 변경하려면 작업 코드를 편집해야 합니다.

대역폭보다 많은 수의 커널 함수를 지정하면 마지막 대역폭이 남아 있는 추정값에 대해 반복됩니다. 예를 들어 대역폭으로 0.5 및 1.0을 지정하고 정규, 2차 및 3차를 커널 함수로 지정하면 3차 함수에 대해 1.0 대역폭이 반복됩니다.

커널 함수보다 많은 수의 대역폭을 지정하면 마지막 커널 함수가 남아 있는 대역폭에 대해 반복됩니다. 예를 들어 대역폭으로 0.5, 1.0 및 1.5를 지정하고 정규 및 2차를 커널 함수로 지정하면 2차 함수가 1.5 대역폭에 사용됩니다.

분포분석: 도표 생성

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **도표**를 클릭합니다.

주: 분석 역할에 상대 가중값 변수를 지정하면 도표를 생성할 수 없습니다.

옵션 이름	설명
히스토그램 도표	히스토그램을 생성하고 연속이론분포 및 커널 밀도함수 추정값에 대한 밀도곡선을 중첩시킵니다(선택 사항). 분포를 선택하려면 선택 영역에서 요약 을 클릭합니다. 분류변수를 지정하면 성분 히스토그램들로 구성되는 비교 히스토그램이 생성됩니다. 이 성분 히스토그램들을 사용하면 분류 레벨별 분석변수의 분포를 비교할 수 있습니다.
확률 도표	연속이론분포에 대한 확률 도표를 생성하고 이론분포에 대하여 추정된(또는 지정된) 위치모수 및 척도모수에 해당하는 참조선을 중첩시킵니다. 확률 도표는 정렬된 변수값을 지정된 이론 분포의 백분위수와 비교할 때 사용됩니다. 데이터분포가 이론분포와 일치하면, 도표의 점이 선형패턴을 형성합니다. 분류변수를 지정하면 성분 확률 도표들로 구성되는 비교 확률 도표가 생성됩니다. 분포를 선택하지 않으면 정규분포가 사용됩니다.
Q-Q 도표	연속이론분포에 대한 Q-Q 도표를 생성하고 이론 분포의 지정 또는 추정된 위치모수 및 척도모수에 해당하는 참조선을 중첩시킵니다. Q-Q 도표는 정렬된 변수값을 지정된 이론 분포의 분위수와 비교할 때 사용

옵션 이름	설명
	됩니다. 데이터분포가 이론분포와 일치하면, 도표의 점이 선형패턴을 형성합니다. 분류변수를 지정하면 성분 Q-Q 도표들로 구성되는 비교 Q-Q 도표가 생성됩니다. 분포를 선택하지 않으면 정규분포가 사용됩니다. 주: Q-Q 도표는 확률도표와 유사합니다. Q-Q 도표는 분포모수의 그래프 추정값을 구하는 데 적합하며, 확률 도표는 백분위수의 그래프 추정값을 구하는 데 적합합니다.
상자 도표	상자 도표를 생성합니다. 분류변수를 할당하면 나란한 상자 도표가 생성됩니다. 상자 도표에는 관측값 그룹의 평균, 사분위수, 최소 관측값 및 최대 관측값이 표시됩니다.
텍스트 기반 도표	다음 도표를 생성합니다. ■ 관측값 수에 따라 줄기-잎 도표 또는 막대 그래프 ■ 상자 도표 ■ 정규 확률 도표 ■ 나란한 도표(그룹 분석변수 역할에 변수를 할당할 때)
축 색상, 배경 색상, 축 너비	도표의 모양을 사용자 정의합니다. 주: 줄기-잎 도표에는 텍스트만 나타나므로 이 옵션을 사용할 수 없습니다.

분포분석: 인셋 추가

이 옵션에 액세스하려면 선택 영역의 **모양** 머리글 아래에서 **인셋**을 클릭합니다.

요약통계량 상자 또는 테이블(인셋)을 그래프에 직접 추가하여 도표를 보기 좋게 만들 수 있습니다. 비교 히스토그램을 생성하면 각각의 작은 히스토그램에 인셋이 표시됩니다.

분포분석: 결과 옵션 지정

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

분포분석 리포트에는 기본적으로 기본 신뢰구간 테이블, 기본 측도 테이블, 위치모수에 대한 검정 테이블, 적률 테이블 등이 포함됩니다.

옵션 이름	설명
기술통계량 및 공정 능력지수 테이블 생략	출력에서 모든 기술통계량 및 공정능력지수 테이블을 숨깁니다. 이 옵션을 선택해도 다음 테이블은 숨겨지지 않습니다. ■ 선택한 분포에 의해 생성된 테이블. 해당 분포에 대한 패널에서 분포테이블 생략 체크 박스를 선택하여 이러한 테이블을 숨길 수 있습니다. ■ INTERVALS 문에 의해 생성된 테이블. NOPRINT 옵션을 지정하면 이러한 테이블을 생략할 수 있습니다. INTERVALS 문에 대한 자세한 내용은 CAPABILITY 프로시저에 대한 도움말을 참조하십시오.
데이터셋에 출력 통계량 저장	출력 데이터셋을 생성합니다. 최종 결과에 포함되는 통계량은 테이블 상자에서 선택한 테이블에 따라 달라집니다. 주: SAS Enterprise Guide는 옵션 창의 작업> 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

도넛 그래프

도넛 그래프 작업 정보	165
도넛 그래프: 그래프 유형 선택	166
도넛 그래프: 분석 역할에 변수 할당	166
도넛 그래프: 모양 옵션 설정	167
도넛 그래프 옵션 설정	167
레이아웃 지정	168
고급 옵션 설정	168
추가 모양 옵션	169
제목 및 각주 지정	169
속성 보기	169

도넛 그래프 작업 정보

도넛 그래프 작업은 전체에 대한 부분의 상대적 기여도를 보여 주는 단순 그래프 또는 그룹 그래프를 생성합니다. 데이터가 부채꼴 모양의 원 "조각"으로 표시됩니다. 각 조각은 하나의 데이터 범주를 나타냅니다. 조각의 크기는 전체 그래프 통계량에 대해 해당 데이터가 차지하는 비율을 나타냅니다. 예를 들어 도넛 그래프는 체인의 전체 판매량 가운데 각 상점의 판매량이 차지하는 부분을 표시할 수 있습니다. 도넛 그래프는 가운데 구멍이 있는 것을 제외하고는 원 그래프와 비슷합니다.

자세한 내용은 [“그래픽 출력형식 선택” \(5페이지\)](#)에서 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GCHART
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

도넛 그래프: 그래프 유형 선택

이 옵션에 액세스하려면 선택 영역에서 **도넛 그래프**를 클릭합니다.

옵션 이름	설명
단순 도넛	전체에 대한 부분의 상대적인 크기를 나타내는 그래프를 생성합니다. 데이터가 부채꼴 모양의 원 "조각"으로 표시됩니다. 각 조각은 하나의 데이터 범주를 나타냅니다. 조각의 크기는 전체 그래프 통계량에 대해 해당 데이터가 차지하는 비율을 나타냅니다.
그룹형 도넛	그룹화 칼럼 역할에 할당한 변수의 각 고유향에 대해 개별 도넛 그래프를 생성합니다.

도넛 그래프: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

주: 선택한 그래프 유형에 따라 사용할 수 있는 역할이 다릅니다.

역할 이름	설명
그래프 칼럼	이 역할에 할당하는 칼럼의 값에 따라 도넛 그래프의 조각이 결정됩니다. 칼럼은 문자 또는 숫자가 될 수 있습니다.
그룹화 칼럼	이 역할에 할당하는 칼럼의 고유한 각 값에 대해 도넛 그래프가 생성됩니다. 모든 그래프는 동일한 페이지에 표시됩니다.
합계 칼럼	이 역할에 할당하는 칼럼에 따라 도넛 그래프의 조각 크기가 결정됩니다. 이 역할에 칼럼을 할당하지 않으면 그래프 칼럼 의 각 값의 빈도에 따라 조각 크기가 결정됩니다.

역할 이름	설명
그룹 분석변수	이 역할에 할당하는 칼럼의 값에 따라 생성되는 그래프의 수가 결정됩니다. 각 고유티값에 대해 별도의 그래프가 생성됩니다. 각 도넛 그래프는 새로운 페이지에 나타납니다.

도넛 그래프: 모양 옵션 설정

도넛 그래프 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **도넛**을 클릭합니다.

옵션 이름	설명
색상 지정	조각 색상을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 기본 색상 구성 - 조각에 기본 색상 구성을 사용합니다. 기본 색상 구성은 현재 스타일에 따라 결정됩니다. 스타일 관리자나 결과에 대한 속성 대화 상자를 사용하여 스타일을 지정할 수 있습니다. SAS Enterprise Guide의 옵션 대화 상자에서 각 결과 유형에 대한 스타일을 지정할 수도 있습니다. SAS Add-In for Microsoft Office에서 SAS Add-In 옵션 대화 상자를 사용하여 SAS 리포트 및 HTML 출력형식에 스타일을 적용할 수 있습니다. ■ 기존 색상(이전 버전에서 지정됨) - 이전 버전의 SAS Enterprise Guide 또는 SAS Add-In for Microsoft Office에서 지정한 색상을 사용할 수 있습니다. 이 옵션은 이전 릴리스에서 마이그레이션된 SAS 콘텐츠에 대해서만 사용 가능합니다. ■ 사용자 정의 색상 - 조각에 사용할 사용자 정의 색상을 지정할 수 있습니다. 드롭다운 리스트를 사용하여 막대 그래프에 최대 12개의 색상을 지정할 수 있습니다. 그래프에 12개가 넘는 색상을 사용해야 할 때는 현재의 기본 색상 구성에서 추가 색상이 선택됩니다. 사용자 정의 색상을 기본값으로 돌리려면 재설정을 클릭합니다. 주: 도넛 그래프 유형에 따라 일부 사용자 정의 색상을 사용하지 못할 수도 있습니다.
조각 수 지정	그래프의 조각 수를 지정합니다. 기본적으로 조각 수는 자동으로 결정됩니다. 연속 칼럼에서는 값이 여러 범위로 나뉘며 각 범위마다

옵션 이름	설명
	조각이 한 개씩 표시됩니다. 전체 4% 이하를 나타내는 조각은 "기타"라는 조각으로 합쳐집니다.
	그래프에 표시되는 조각 수를 지정하려면 조각 개수 지정 체크 박스를 선택합니다.

레이아웃 지정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **레이아웃**을 클릭합니다.

각 조각의 레이블에 이름, 백분율, 통계값 또는 이 세 가지의 조합 중 어떤 것을 사용할지를 선택합니다. 레이블의 각 요소에 대해 다음 레이블 위치 중 하나를 선택합니다.

- 조각 내부
- 원 또는 도넛 외부
- 조각 외부(선으로 연결 표시)

조각 레이블의 글꼴 색상을 조각과 동일한 색상으로 표시되도록 선택할 수 있습니다. 조각 내부에 레이블을 표시할 때는 이 옵션을 선택하지 마십시오.

"기타" 조각에 대해 백분율 및 레이블을 지정할 수 있습니다. 이 조각은 특정 백분율 값 이하인 그래프 통계량 값을 가지는 중간점을 모은 것입니다. 총합 백분율의 기본값은 4입니다. 그러므로 전체 중 4% 미만을 나타내는 조각은 "기타" 조각에 넣을 수 있습니다.

"기타" 조각의 레이블에는 최대 16자의 텍스트 문자열을 지정할 수 있습니다. 기본 레이블은 기타입니다. "기타" 조각은 조각 순서에 관계없이 원의 마지막 조각입니다.

하나의 중간점만 "기타" 범주에 속하면 해당 중간점에 대한 조각은 원의 원래 위치와 원래 레이블을 유지합니다. 예를 들어 음료의 백분율 판매를 표시하는 원에 다음 조각 및 백분율 값이 있다고 가정합니다. 합계 백분율이 5이면, Soda에 대한 조각이 마지막 조각이 되는 대신에 원에서 세 번째 조각으로 남아 있고, 이 조각은 "기타" 대신에 "Soda"라는 레이블이 붙게 됩니다.

그룹 도표는 **그룹 레이아웃** 영역을 사용하여 그래프를 정렬합니다. 출력 영역의 칼럼 수 또는 행 수를 지정합니다.

고급 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **고급**을 클릭합니다.

그래프의 조각 계산에 사용할 통계량을 드롭다운 리스트에서 선택합니다. **합계 칼럼** 역할에 변수를 할당하지 않으면 각 조각의 빈도 또는 백분율을 계산하도록 선택할 수 있습니다. 기본적으로 빈도가 계산됩니다. **합계 칼럼** 역할에 변수를 할당하면 각 조각의 합계 또는 평균을 계산하도록 선택할 수 있습니다. 기본적으로 합이 계산됩니다.

추가 모양 옵션

일부 옵션은 여러 그래프 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- [“범례 옵션 설정” \(596페이지\)](#)
- [“그래프 영역 사용자 정의” \(597페이지\)](#)

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

.....
 주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간

- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

데이터 파일을 PC로 다운로드

데이터 파일을 PC로 다운로드 작업 정보	171
데이터 파일을 PC로 다운로드: 다운로드할 SAS 데이터셋 선택	171
데이터 파일을 PC로 다운로드: 다운로드 옵션 지정	172

데이터 파일을 PC로 다운로드 작업 정보

PC에 데이터 파일 다운로드 작업을 통해 하나 이상의 SAS 서버에서 하나 이상의 SAS 데이터셋(*.sas7bdat)을 선택하고 로컬 Microsoft Windows 파일 시스템에 복사할 수 있습니다.

이 작업은 복사된 데이터 파일, 바이트 크기, 각 파일 및 그룹인 모든 파일에 대한 시간, 전송 시 발생한 오류 등이 요약된 로그를 생성합니다.

데이터 파일을 PC로 다운로드: 다운로드할 SAS 데이터셋 선택

- 1 **추가**를 클릭하여 복사할 SAS 데이터셋을 선택합니다. **열기** 대화 상자가 나타납니다.
- 2 적절한 서버 및 라이브러리를 탐색하고 복사할 파일의 이름을 선택합니다. 둘 이상의 파일을 선택하려면 Ctrl 키를 누른 상태로 추가할 파일을 선택합니다.
- 3 **열기**를 클릭하여 복사할 데이터셋 리스트에 파일을 추가합니다.

주: **추가**를 다시 클릭하고 다른 서버에서 데이터를 선택하면 여러 SAS 서버에서 데이터 파일을 선택할 수 있습니다.

.....
주: 복사되는 모든 파일에는 대상 폴더에서 소문자 이름이 지정됩니다.
.....

- 4 다음을 클릭하여 다운로드 옵션을 지정합니다.

데이터 파일을 PC로 다운로드: 다운로드 옵션 지정

- 1 다운로드 옵션을 지정하기 전에 다운로드할 데이터셋을 선택해야 합니다.
- 2 파일을 복사할 폴더 경로를 입력합니다. 기본 위치는 내 문서 폴더입니다.
- 3 로컬 컴퓨터에 복사할 파일과 이름이 같은 파일이 있을 때 이를 바꾸려면 **동일한 이름의 기존 파일 바꾸기** 체크 박스를 선택합니다. 기본적으로는 기존 파일이 바뀌지 않습니다.
- 4 **마침**을 클릭하여 데이터셋을 복사합니다.

.....
주: 소스 라이브러리가 단일 파일 경로의 Base SAS 엔진 라이브러리이면 파일이 해당 라이브러리 경로에서 대상 폴더로 직접 전송됩니다. 소스 라이브러리가 Base SAS 엔진 라이브러리가 아니거나 연결된 라이브러리이면 파일이 먼저 WORK의 staging 영역으로 전송된 다음 PROC SQL을 사용하여 대상 폴더로 복사됩니다.
.....

요인분석

요인분석 작업 정보	173
요인분석: 분석 역할에 변수 할당	174
요인분석: 요인화 방법 옵션 설정	175
요인분석: 공통성 옵션 설정	176
요인분석: 회전 및 도표 옵션 설정	177
요인분석: 결과 옵션 설정	179
제목 및 각주 지정	180

요인분석 작업 정보

요인분석은 다양한 공통 요인과 요소에 대한 분석 및 회전을 실행합니다. 다변량 데이터, 상관행렬, 공분산행렬, 요인 패턴 또는 점수화 계수 행렬 등을 입력할 수 있습니다.

이 작업을 사용하여 한 해의 직원 평가에 대한 공통 요인분석을 수행할 수 있습니다. 예를 들어 감독자가 103명의 경찰 공무원을 14개의 기준(변수)으로 평가한다고 가정합니다. 이러한 평가 이면에 있는 요인들을 확인하기 위해 공통 요인분석을 수행할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	FACTOR
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	없음

요인분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분석변수	분석할 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
편상관변수	분석할 변수의 부분 변수를 지정합니다. 이 변수는 부분상관계수 또는 공분산행렬에 기초한 분석에 사용됩니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석변수로 이미 선택한 변수로는 테이블을 그룹화할 수 없습니다.
빈도 수	관측값 내에서 다른 값의 발생 빈도를 나타내는 변수를 하나 지정합니다. 이 변수를 사용하면 작업은 각 행이 n번 발생하는 것으로 취급합니다. 유의수준에 대한 자유도 결정 시, 관측값의 합계는 변수의 합과 같은 것으로 간주됩니다. 빈도변수의 값이 없거나 1보다 작으면 관측값이 분석에 사용되지 않습니다. 값이 정수가 아니면 정수 부분만 사용됩니다.
상대 가중값 변수	입력 데이터셋의 각 관측값에 대한 상대 가중값으로 사용할 변수를 지정합니다. 각 관측값의 분산이 다르고 상대 가중값의 값이 분산의 역과 비례할 때 자주 사용됩니다. 상대 가중값이 음수이거나 결측값이면, 분석에서 제외됩니다.

요인분석: 요인화 방법 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **요인화 방법**을 클릭합니다.

옵션 이름	설명
요인화 방법	요인화 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 모든 사전값이 1이면 주성분분석을 수행하고 그렇지 않으면 주요인분석을 수행합니다. 데이터셋 유형이 FACTOR일 때를 제외하고 이 방법이 기본값입니다. ■ Harris 성분 분석 - 정준 성분 분석의 비반복 근사인 S-1RS-1의 해리스 성분 분석을 수행합니다. ■ 이미지 공분산 행렬 - 이미지 공분산 행렬에 대해 주성분분석을 수행합니다. 비특이 상관행렬이 필요합니다. ■ Alpha 요인 분석 - Alpha 요인 분석을 수행합니다. ■ 최대 우도 요인 분석 - 중요하지 않은 세부 항목을 제외하고 Fuller에 대한 알고리즘으로 최대 우도 요인 분석을 수행합니다. 비특이 상관행렬이 필요합니다. ■ 반복 주요인분석 - 반복적인 주요인분석을 수행합니다. ■ 비가중 최소제곱 요인 분석 - 비가중 최소제곱 요인 분석을 수행합니다. ■ 기존 요인 패턴 - FACTOR, CORR, UCORR, COV 또는 UCOV 데이터셋 유형에서 요인 패턴을 읽습니다. 데이터셋 유형이 FACTOR이면, 이 방법이 기본값입니다. ■ 요인 점수화 계수 - FACTOR, CORR, UCORR, COV 또는 UCOV 데이터셋 유형에서 점수화 계수를 읽습니다. 데이터셋에는 상관계수 또는 공분산행렬도 포함되어 있어야 합니다. 방법을 선택한 후 테이블의 수치적 속성 및 요인 수 옵션을 설정할 수 있습니다. 요인 수 옵션 중 두 개 이상을 변경하면 유지되는 요인의 수는 다음 기준을 만족하는 최소 수가 됩니다. ■ 요인을 유지하기 위한 최소 곱셈값 ■ 유지할 요인 수 ■ 유지 요인이 차지하는 공통분산 비율
행렬 옵션	행렬 옵션을 지정합니다. ■ 상관행렬 대신 공분산행렬을 요인화 - 요인화 방법이 주성분분석, 이미지 공분산행렬, 반복 주요인분석 또는 비가중 최소제곱 요인 분석일 때 요인화에 상관행렬 대신 공분산행렬을 사용합니다.

옵션 이름	설명
	■ 가중된 행렬 요인화 - 요인화 방법이 주성분분석, 이미지 공분산행렬, 반복 주요인분석 또는 비가중 최소제곱 분석일 때 요인화에 가중 상관행렬 또는 공분산행렬을 사용합니다. 주: 이 옵션은 요인화 방법이 주성분분석, 이미지 공분산행렬, 반복 주요인분석 또는 비가중 최소제곱 요인 분석이고 입력 데이터셋 유형이 CORR, UCORR, COV, UCOV 또는 FACTOR일 때만 사용할 수 있습니다. ■ 결과 행렬을 평균에 대해 수정 - 분석에 절편을 사용하고 공분산이나 상관계수가 평균에 대해 수정됩니다. 기본적으로 이 옵션이 선택됩니다.
분산 계산에 사용할 분모	분산 및 공분산 계산에 사용할 분모를 지정합니다. 분산 계산에 사용되는 기본 분모는 자유도입니다.

요인분석: 공통성 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **공통성**을 클릭합니다.

옵션 이름	설명
사전 공통성 추정값	다음 옵션 중 하나를 선택합니다. ■ 다중 상관계수의 제공값에 비례 - 사전 공통성 추정값을 다중 상관계수의 제공으로 설정하되, 그 합이 상관계수 절대값의 최댓값 합과 같아지도록 추정값을 조정합니다. ■ 모든 다른 칼럼과의 다중 상관계수의 제공 - 각 변수의 사전 공통성 추정값을 다른 모든 변수와의 다중 상관계수 제공으로 설정합니다. ■ 다른 칼럼과의 상관계수 절대값의 최댓값 - 각 변수의 사전 공통성 추정값을 다른 변수와의 상관계수 절대값의 최댓값으로 설정합니다. ■ 모두 1로 지정 - 모든 사전 공통성을 1.0으로 설정합니다. ■ 입력 데이터셋의 PRIORS 또는 COMMUNAL 행에서 입력 - _TYPE_ 변수에 PRIORS 또는 COMMUNAL이 포함되어 있고 데이터셋 유형이 FACTOR일 때 첫 번째 관측값에서 사전 공통성 추정값을 읽습니다. 주: 이 옵션은 선택한 요인화 방법이 주성분분석, Alpha 요인 분석, 최대 우도 요인 분석, 반복 주요인분석 또는 비가중 최소제곱 요인분석이고 입력 데이터셋 유형이 FACTOR일 때만 사용 가능합니다.

옵션 이름	설명
	■ 0과 1 사이의 난수값으로 지정 - 사전 공통성 추정값을 0과 1 사이에 분포된 난수값으로 설정합니다. ■ 각 칼럼에 값 입력 - 입력 상자에서 각 변수에 대한 사전 공통성 추정값을 선택한 후 새로운 값을 지정할 수 있습니다. 각 변수에 대한 사전 공통성 추정값의 기본값은 0입니다. 주: 사전 공통성 추정값의 기본값은 사용하는 요인화 방법에 따라 달라집니다.
Heywood 옵션	기본적으로 Alpha 요인분석, 최대우도 요인분석, 반복 주요인분석, 비가중 최소제곱 요인분석 요인화 방법은 공통성 추정값이 1을 초과하면, 반복을 멈추고 요인 수를 0으로 설정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 1보다 큰 공통성을 1로 지정 - 1보다 큰 모든 공통성을 1로 설정하고 반복을 계속합니다. ■ 1을 초과하는 공통성 허용 - 공통성이 1을 초과할 때 반복이 계속 진행되도록 허용합니다. 공통성이 지나치게 커지고 불량상태의 Hessian이 나타날 수 있으므로 수렴 문제가 발생할 가능성이 있습니다.

요인분석: 회전 및 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **회전과 도표**를 클릭합니다.

옵션 이름	설명
회전 방법	회전 방법을 지정합니다. 다음 방법 중에서 선택할 수 있습니다. 선택한 방법에 따라, 추가 옵션을 사용할 수 있습니다. ■ 회전 안 함 - 원래의 직교회전을 중지하고 회전을 수행하지 않습니다. 기본값입니다. ■ Orthogonal 직교회전 - Orthogonal 직교회전을 지정합니다. 이것은 orthomax 가중값이 $n/2$인 일반 orthomax를 사용하는 것과 동일합니다(n은 요인의 수). ■ Harris-Kaiser case II orthoblique - Harris-Kaiser case II orthoblique 회전을 지정합니다. 주 요인 방법에 의해 요인이 추출되었음을 가정하고 고유벡터를 재척도화하는 데 사용되는 고윳값 제곱근의 거듭제곱을 지정할 수 있습니다(선택 사항). 다른 추출 방법에서는 고유벡터가 비회전 요인행렬의 정규화된 칼럼으로 대체되거나 고윳값이 칼럼 정규화 상수로 대체됩니다. 기본 고윳값인 0에서는 독립 군집 해가 산출되어 각 변수가 하나의 요인에만 큰 비중을 두는 경향이 있습니다.

옵션 이름	설명
	■ Orthomax 일반회전 - 기본 Orthomax 가중값이 0인 Orthomax 회전을 지정합니다. 대부분 0과 변수 수 사이의 값을 갖지만 Orthomax 가중값의 유효값에 대한 제한은 없습니다. Orthomax 가중값은 Varimax 회전에서의 하나의 결과와 동일합니다. ■ Parsimax 직교회전 - Parsimax 직교회전을 지정합니다. 이는 일반 Orthomax가 $[(nvar \times (n - 1)) / (nvar + n - 2)]$와 같은 Orthomax 일반회전을 사용하는 것에 해당합니다. 여기서 $nvar$는 변수 개수이고 n은 요인 개수입니다. ■ Oblique Procrustes - TARGET 데이터셋에서 제공된 타겟 패턴으로 Oblique Procrustes 회전을 지정합니다. 회전 후 단위분산으로 척도화된 인수와 함께 비제약 최소제곱법이 사용됩니다. ■ Oblique Promax - Oblique Promax 회전을 지정합니다. 직각 또는 경사에 대한 원하는 선행 회전 방법을 설정할 수 있습니다. 기본값 3으로 계산된 타겟 패턴 계산에 사용되는 거듭제곱을 변경할 수도 있습니다. ■ Quartimax 직교회전 - Quartimax 직교회전을 지정합니다. ■ Varimax 직교회전 - Varimax 직교회전을 지정합니다.
요인 패턴 행의 정규화방법	행을 정규화하는 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ Kaiser 정규화 - 기본 방법인 Kaiser 정규화를 사용합니다. ■ Cureton-Mulaik 기법에 의해 가중화 - Cureton-Mulaik 기법을 사용하여 행에 가중값을 지정합니다. ■ 공분산을 나타내기 위해 재척도화 - 패턴 행렬의 행을 재척도화하여 상관계수 대신 공분산을 나타냅니다. ■ 정규화하지 않음 - 정규화가 수행되지 않습니다.
출력할 도표	결과에 포함할 도표를 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 요인 패턴 도표(회전 후) - 회전 후의 요인 패턴 도표를 표시합니다. 도표화할 요인 수 옵션을 사용하여 처음 n개 요인의 모든 쌍을 도표화하여 총 $n(n - 1)/2$개의 도표를 생성할 수 있도록 도표화할 요인의 수 n을 설정합니다. 정수값을 직접 입력하거나 위쪽/아래쪽 화살표를 사용하여 클릭할 때마다 값을 1씩 조정할 수 있습니다. 기본값은 2입니다. ■ 비회전 요인 패턴 도표 - 비회전 요인 패턴 도표를 표시합니다. ■ 고윳값의 스크리 그래프 표시 - 고윳값의 스크리 그래프를 표시합니다. 기본적으로 어떠한 도표도 표시되지 않습니다. 표시할 도표를 선택한 후 도표화할 요인 수도 지정할 수 있습니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 모든 요인 사용 - 모든 요인을 도표화합니다. 기본값입니다.

옵션 이름	설명
	■ 도표를 그릴 요인 수 - 도표화할 요인 수(n)를 지정할 수 있습니다. 지정할 수 있는 최솟값은 2입니다. 값을 지정하면 첫 번째 n 요인의 모든 쌍이 도표화되어 총 $n(n-1)/2$개의 도표가 생성됩니다.

요인분석: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
출력 데이터 저장	출력 데이터 저장 영역에서 영구적으로 저장할 출력 데이터의 요인 또는 통계량 체크 박스를 선택합니다. 통계량을 선택하면 출력 데이터셋에서 상관 행렬을 제외할 수 있습니다. 상관 행렬은 요청된 다른 출력에 대해 상관 행렬이 필요하지 않고(예: 스코어 또는 잔차 상관)이 표시되는 경우) 및 기존 요인 패턴을 인수분해 방법으로 선택한 경우에만 제외됩니다. 데이터셋 유형이 FACTOR이면, 이 방법이 기본값입니다. 주: 입력 데이터셋 유형이 CORR, UCORR, COV, UCOV 또는 FACTOR이면 요인 옵션을 사용할 수 없습니다. 이 옵션은 편상관 변수 역할에 변수를 할당할 때만 사용할 수 있습니다. 주: SAS Enterprise Guide는 옵션 창의 작업 > 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
출력 표시	출력 표시 영역의 다음 옵션 중에서 선택합니다. ■ 잔차 상관행렬 - 잔차 상관행렬 및 관련된 편상관계수행렬을 표시합니다. 잔차 상관행렬의 대각원소는 고유 분산입니다. ■ 고윳값은 대각 요소가 공통 추정값으로 대체된 감소 상관 행렬의 고유벡터를 표시합니다. maximum-likelihood factor analysis 요인화 방법 사용 시 고유벡터는 감소된 상관행렬 가중값에 대한 것입니다. ■ 요인 점수화 계수 - 요인 점수화 계수를 표시합니다. 비회전 주성분일 때를 제외하고 각 요인에 대한 다중상관계수의 제공함도 변수와 함께 표시됩니다. ■ 적재값의 절대값이 큰 순으로 행렬의 행 순서화 - 출력에서 여러 요인행렬의 행(변수) 순서를 변경합니다. 첫 번째 요인에서

옵션 이름	설명
	적재값의 절대값이 큰 변수(oblique 회전의 참조 구조 적재값)가 적재값이 큰 변수부터 작은 변수 순으로 가장 먼저 표시되고 두 번째 요인에서 적재값의 절대값이 가장 큰 변수가 그 다음에 표시됩니다.
■ 입력 칼럼의 평균과 표준편차	- 평균, 표준편차 및 관측값 개수를 표시합니다.
■ 입력 칼럼의 상관행렬	- 상관행렬 또는 편상관행렬을 표시합니다.
■ 표본 적합성에 대한 Kaiser의 측도	- 기타 모든 변수(음의 역상 상관계수) 및 표본 적합성에 대한 Kaiser의 측도를 제어하면서, 각 변수쌍 간의 편상관행렬을 생성합니다.
■ 입력 요인 패턴	- 입력 요인 패턴 및 관련 통계량을 표시합니다. oblique는 참조 및 인수 구조가 계산되어 표시됩니다. 주: 이 옵션은 기존 요인 패턴 방법을 선택할 때만 사용할 수 있습니다.
■ 점수화 계수	- 점수화 계수 및 관련 통계량을 표시합니다. oblique는 참조 및 인수 구조가 계산되어 표시됩니다. 주: 이 옵션은 요인 점수화 계수 방법을 선택할 때만 사용할 수 있습니다.
	기본적으로 요인분석 결과 또는 관련 통계량 계산이 선택되지 않습니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

Forecast Studio 프로젝트 생성

Forecast Studio 프로젝트 생성 마법사 정보	181
Forecast Studio 프로젝트 생성: 데이터 확인	182
Forecast Studio 프로젝트 생성: 예측변수 선택	182
Forecast Studio 프로젝트 생성: 분류변수 선택	183
Forecast Studio 프로젝트 생성: 프로젝트 저장	184

Forecast Studio 프로젝트 생성 마법사 정보

Forecast Studio 프로젝트 생성 마법사는 예측변수를 지정하고 데이터의 계층적 예측 여부를 선택하며 새로운 **SAS Forecast Studio** 프로젝트의 예측 시간 범위를 지정하는 데 도움을 줍니다. **SAS Forecast Studio**는 자동화를 통해 예측 과정을 신속하게 수행하도록 고안된 예측 응용 프로그램입니다. 이 소프트웨어는 타임스탬프된 데이터를 예측할 때 사용할 시계열 모델을 자동으로 선택합니다. 이 프로젝트를 저장하려면 **SAS Forecast Server** 액세스 권한이 있어야 합니다.

주: 컴퓨터에서 SAS 데이터를 사용하는 예측 프로젝트를 생성하려면 로컬 컴퓨터에서 데이터 소스를 선택하지 말고 파일 시스템을 통해 선택해야 합니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	SAS Forecast Server 프로시저
필수 SAS 제품	SAS Forecast Server
추가 SAS 제품(권장)	없음

Forecast Studio 프로젝트 생성: 데이터 확인

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

다음을 클릭하여 예측변수를 선택합니다.

Forecast Studio 프로젝트 생성: 예측변수 선택

- 1 예측변수를 선택하려면 먼저 데이터를 확인해야 합니다.
- 2 다음과 같은 역할에 변수를 할당합니다.

시간 ID

이 변수는 데이터에 시간 ID를 지정합니다. 시간 ID 변수는 각 관측값의 SAS 날짜, 날짜/시간 또는 시간값이 포함된 입력 데이터셋의 변수입니다. 이 변수는 데이터의 빈도 및 순서를 결정하고 예측치에 대해 시간 ID 값을 추정하는 데 사용됩니다. 이 역할에 한 개의 변수만 할당할 수 있으며, 날짜 변수, 날짜/시간 변수 또는 날짜나 날짜/시간 값을 포함하는 숫자 변수여야 합니다.

예측(종속변수)

이 역할에 할당한 칼럼은 모델화하고 예측할 변수를 지정합니다. 예측변수는 종속변수라고도 불립니다. 이 역할에는 한 개 이상의 숫자 변수를 할당해야 합니다. 위쪽/아래쪽 화살표를 사용하여 변수 순서를 지정합니다.

- 3 (선택 사항) **예측에 사용하는 변수:** 이 역할에 할당한 칼럼이 설명변수, 입력변수, 예측변수 또는 인과변수를 지정합니다. 이 역할에 할당한 칼럼은 설명, 입력, 예측함수 또는 원인요인변수를 지정합니다. 사용할 예측 변수는 독립변수라고도 불립니다.
- 4 다음을 클릭하여 분류변수를 선택합니다.

Forecast Studio 프로젝트 생성: 분류변수 선택

- 1 분류변수를 선택하려면 먼저 예측변수를 선택해야 합니다.
- 2 **분류변수(이 변수의 값에 따라 통계량 생성됨)** 역할에 변수를 할당합니다. 이 역할에 할당된 변수를 분류변수라고 합니다. 분류변수에 대해 같은 값을 가진 관측값은 그룹으로 지정됩니다. 분류변수를 할당하면 관측값 그룹에 대한 별도의 분석 내용을 얻을 수 있습니다. 문자와 숫자 변수를 이 역할에 할당할 수 있습니다.

분류변수의 순서는 계층 구조를 의미합니다. 위쪽/아래쪽 화살표를 사용하여 변수 순서를 지정합니다.

- 3 출력 표시 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다.
 - **리스트 뷰** - 데이터 계층을 생성하지 않습니다. 리스트에 예측 시계열이 구성됩니다.
 - **계층** - 분류변수값을 사용하여 데이터를 계층으로 구성합니다. 이 옵션이 기본적으로 선택됩니다.

주: **예측(종속변수)** 역할에 여러 변수를 할당하면 **계층** 옵션을 사용할 수 없습니다.

- 4 출력을 계층으로 표시하도록 선택하면 **설정**을 클릭하여 계층 설정을 지정할 수 있습니다. **계층 설정** 창이 나타납니다.

다음 옵션 중에서 선택할 수 있습니다.

집계 방법

계층 레벨의 집계 데이터에 사용할 통계량을 지정합니다. 기본적으로 모든 종속 및 독립변수는 테이블에 나열됩니다. **집계** 칼럼에서 각 변수에 사용할 집계 통계량을 선택합니다.

예를 들어 데이터셋에 제품 그룹에 대한 판매가 포함되어 있습니다. 범주에 대한 총 매출을 알고자 할 수 있습니다. 데이터에 각 제품의 가격이 포함되어 있을 때 제품군의 평균 가격을 알려고 한다면 **합계**를 집계 통계량으로 선택합니다. 데이터에 각 상품의 가격이 포함되어 있으면 상품 라인의 평균 가격을 알고 싶을 수 있습니다. 데이터에 각 제품의 가격이 포함되어 있을 때 제품군의 평균 가격을 알려고 한다면 **평균**을 집계 통계량으로 선택합니다.

보정 레벨

예측할 데이터가 포함된 계층에서 레벨을 선택합니다. 다음 옵션 중에서 선택할 수 있습니다.

- **상향식** - 최저 계층 레벨에서 데이터를 사용하여 예측값을 생성합니다. 이 예측값은 더 높은 계층 레벨에 맞게 조정된 예측값을 생성하는 데 사용됩니다.

상향식 방법으로 계절성과 같은 데이터의 패턴을 확인할 수 있습니다. 그러나 사용자가 최저 계층 레벨에 있기 때문에 데이터의 잡음과 임의성이 매우 높을 수도 있습니다. 또한 최저 계층 레벨의 데이터가 산발적이거나 빈약하기 때문에 예측에 실패할 수 있습니다.

- **미들아웃** - 낮은 레벨에서 데이터를 합한 다음, 이 값을 사용하여 중간 레벨의 예측값을 생성합니다. 일부 계층에는 둘 이상의 중간 레벨이 있기 때문에 사용하려는 레벨을 지정해야 합니다. 중간 레벨의 예측값은 더 높은 계층과 낮은 계층 모두의 예측값을 생성하는 데 사용됩니다.

- **하향식** - 최저 레벨에서 데이터를 합한 다음, 이 값을 사용하여 예측값을 생성합니다. 최고 레벨의 예측값은 더 낮은 계층 레벨의 예측값을 생성하는 데 사용됩니다.

하향식 방법으로 낮은 계층 레벨의 데이터에서 과도한 잡음을 제거할 수 있습니다. 그러나 계절성과 같은 예측값의 패턴이 손실될 수도 있습니다.

Disaggregation 방법

계층의 Disaggregation 방법을 선택합니다. 다음 옵션 중에서 선택할 수 있습니다.

- **예측 비율** - 계층 평균을 기반으로 하여 예측값을 조정합니다. 기본값입니다.
- **동일한 차분 분할** - 레벨 간 예측 편차를 기반으로 하여 예측값을 조정합니다.

- 5 **예측할 간격 수(수평)** 상자에서 단단계 예측에 대한 간격 수를 미래로 지정합니다. 기본값은 12입니다.
- 6 다음을 클릭하여 프로젝트를 저장합니다.

Forecast Studio 프로젝트 생성: 프로젝트 저장

- 1 프로젝트를 저장하려면 분류변수를 선택해야 합니다.
- 2 **다른 이름으로 저장** 상자에 프로젝트 이름을 입력합니다. 프로젝트 이름은 올바른 SAS 이름이어야 합니다. 이름은 32자까지 가능하고 문자(A-Z)로 시작해야 합니다. 다음 글자는 문자와 숫자(0-9) 모두 가능합니다. 대소문자가 모두 유효합니다. SAS 명명 규칙에 대한 자세한 내용은 [SAS Language Reference: Concepts](#)를 참조하십시오.
- 3 **마침**을 클릭하여 **SAS Forecast Studio** 프로젝트를 생성합니다. 프로젝트를 생성한 후 **SAS Forecast Studio**에서 열 수 있습니다.

Forecast Studio 프로젝트 열기

<i>Forecast Studio 프로젝트 열기 마법사 정보</i>	185
<i>Forecast Studio 프로젝트 열기: 예측 프로젝트 선택</i>	186
<i>Forecast Studio 프로젝트 열기: 예측 시계열 선택</i>	186
<i>Forecast Studio 프로젝트 열기: 결과 유형 선택</i>	186
<i>Forecast Studio 프로젝트 열기: 제목 및 각주 지정</i>	187

Forecast Studio 프로젝트 열기 마법사 정보

Forecast Studio 프로젝트 열기 마법사는 기존 **SAS Forecast Studio** 프로젝트에서 선택한 시계열을 열고 결과 표시 방법을 지정하는 데 도움을 줍니다. **SAS Forecast Studio**는 자동화를 통해 예측 과정을 신속하게 수행하도록 고안된 예측 응용 프로그램입니다. 이 소프트웨어는 타임스탬프된 데이터를 예측할 때 사용할 시계열 모델을 자동으로 선택합니다.

이 작업으로 **SAS Forecast Studio**에서 생성된 프로젝트를 열고 추가 분석 작업을 수행할 수 있습니다. 이 프로젝트를 열려면 **SAS Forecast Server** 액세스 권한이 있어야 합니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	SAS Forecast Server 프로시저
필수 SAS 제품	SAS Forecast Server
추가 SAS 제품(권장)	없음

Forecast Studio 프로젝트 열기: 예측 프로젝트 선택

예측 프로젝트 영역에서 사용 가능한 모든 프로젝트 리스트 및 프로젝트가 마지막으로 수정된 날짜를 확인할 수 있습니다. 열려는 예측 프로젝트를 선택합니다. 한 번에 하나의 프로젝트만 열 수 있습니다.

기본적으로 프로젝트의 상세 정보는 표시되지 않습니다. 선택한 프로젝트와 연관된 설명, 종속변수 및 분류변수(BY 변수)를 확인하려면 **자세히**를 클릭합니다. 이 정보는 예측 서버에 저장되어 있기 때문에 상세 정보를 표시하는 데 몇 초가 걸릴 수 있습니다. 상세 정보를 숨기려면 **숨기기**를 클릭합니다.

다음을 클릭하여 예측 시계열을 선택합니다.

Forecast Studio 프로젝트 열기: 예측 시계열 선택

- 1 예측 시계열을 선택하려면 먼저 예측 프로젝트를 선택해야 합니다.
- 2 **표시할 시계열 선택** 영역에서 열려는 예측 시계열의 체크 박스를 선택합니다.
프로젝트 데이터가 계층적 예측값이 아니면 시계열이 리스트에 표시됩니다. 프로젝트 데이터가 계층적 예측값이면 시계열은 계층에 표시됩니다. **모두 펼치기**를 클릭하여 계층의 모든 노드를 펼칩니다. 노드를 접으려면 **모두 접기**를 클릭합니다.
- 3 **다음**을 클릭하여 결과 유형을 선택합니다.

Forecast Studio 프로젝트 열기: 결과 유형 선택

- 1 결과 유형을 선택하려면 먼저 예측 시계열을 선택해야 합니다.
- 2 열려는 결과 유형을 선택합니다. 다음 옵션 중에서 선택할 수 있습니다.
 - **예측 도표** - 선택한 모델의 과거값, 적합값 및 조정된 예측값을 표시합니다.

- **예측 테이블** - 통계 예측값, 조정값, 조정된 예측값, 수동 Override 값 및 최종 예측값에 대한 모델을 표시합니다.
 - **모델 선택 테이블** - 현재 프로젝트에 정의된 모든 모델을 표시합니다.
 - **모델 모수 추정값 테이블** - 각 모델 모수의 추정값, 표준오차 및 유의성 검정을 표시합니다.
 - **성분 테이블** - 모델에서 사용된 각 성분의 테이블을 표시합니다.
 - **적합통계량 테이블** - 예측 모델에 대해 사용 가능한 모든 적합통계량을 표시합니다.
 - **예측 요약 테이블** - 예측 기간의 시간 범위에 대한 종속변수의 예측값을 표시합니다.
- 3 다음을 클릭하여 제목과 각주를 지정합니다.

Forecast Studio 프로젝트 열기: 제목 및 각주 지정

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

Forecast Studio 프로젝트 Override

<i>Forecast Studio 프로젝트 Override 마법사 정보</i>	189
<i>Forecast Studio 프로젝트 Override: 예측 프로젝트 선택</i>	190
<i>Forecast Studio 프로젝트 Override: 예측 계열 선택 및 Override 실행</i>	190
예측 시계열 선택	190
Override 생성 및 실행	191
선택한 시간 기간에 대한 Override 계산	191
Override 충돌 해결	192

Forecast Studio 프로젝트 Override 마법사 정보

SAS Forecast Studio는 자동화를 통해 예측 과정을 신속하게 수행하도록 고안된 예측 응용 프로그램입니다. 이 소프트웨어는 타임스탬프된 데이터를 예측할 때 사용할 시계열 모델을 자동으로 선택합니다. **SAS Enterprise Guide** 또는 **SAS Add-In to Microsoft Office**에서 이 프로젝트를 열려면 **SAS Forecast Server** 액세스 권한이 있어야 합니다.

Forecast Studio 프로젝트 Override 마법사를 사용하면 기존 **SAS Forecast Studio** 프로젝트의 예측 데이터에 대한 Override를 실행할 수 있습니다. Override는 예측된 값 대신 지정하는 대체값입니다. 미래값에 대한 Override만 실행할 수 있습니다.

Override를 실행하면 **SAS Forecast Server**에서 프로젝트를 조정합니다. Override가 적용되는 방식은 프로젝트의 데이터가 계층적으로 예측되는지 여부에 따라 달라집니다.

- 프로젝트의 데이터가 계층적으로 예측되면 Override는 조정된 통계 예측을 기준으로 수정되는 방식으로 적용됩니다.
- 프로젝트의 데이터가 계층적으로 예측되지 않으면 Override는 통계 예측을 기준으로 수정되는 방식으로 적용됩니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	SAS Forecast Server 프로시저
필수 SAS 제품	SAS Forecast Server
추가 SAS 제품(권장)	없음

Forecast Studio 프로젝트 Override: 예측 프로젝트 선택

예측 프로젝트 영역에서 사용 가능한 모든 프로젝트 리스트 및 프로젝트가 마지막으로 수정된 날짜를 확인할 수 있습니다. 열려는 예측 프로젝트를 선택합니다. 한 번에 하나의 프로젝트만 열 수 있습니다.

기본적으로 프로젝트의 상세 정보는 표시되지 않습니다. 선택한 프로젝트와 연관된 설명, 종속변수 및 분류변수(BY 변수)를 확인하려면 **자세히**를 클릭합니다. 이 정보는 예측 서버에 저장되어 있기 때문에 상세 정보를 표시하는 데 몇 초가 걸릴 수 있습니다. 상세 정보를 숨기려면 **숨기기**를 클릭합니다.

다음을 클릭하여 예측 시계열을 선택합니다.

Forecast Studio 프로젝트 Override: 예측 계열 선택 및 Override 실행

예측 시계열 선택

- 1 예측 시계열을 선택하려면 먼저 예측 프로젝트를 선택해야 합니다.
- 2 트리 뷰에서 Override를 실행할 시계열을 선택합니다.

프로젝트 데이터가 계층적으로 예측되면 시계열은 계층으로 표시됩니다. **모두 펼치기**를 클릭하여 계층의 모든 노드를 펼칩니다. 노드를 접으려면 **모두 접기**를 클릭합니다.

프로젝트 데이터가 계층적 예측값이 아니면 시계열이 리스트에 표시됩니다.

- 3 시계열을 선택한 다음 **예측 데이터 Override**를 클릭하여 Override를 생성 및 실행합니다.
- 4 Override를 실행할 모든 시계열에 대해 2 - 3 단계를 반복합니다.
- 5 Override를 모두 생성 및 실행했으면 **마침**을 클릭하여 이 마법사를 닫습니다.

Override 생성 및 실행

예측 데이터 Override 대화 상자의 맨 위에는 두 개의 그래프가 있습니다. 왼쪽의 그래프에서 파란색 선은 예측값을 표시하고 원은 실제 데이터를 나타냅니다. 오른쪽의 그래프에서 진한 파란색 선은 예측값을 표시하고, 연한 파란색 선은 예측값의 신뢰구간이며 검은색 더하기 부호는 지정하는 Override를 나타냅니다.

Override를 생성 및 실행하는 방법:

- 1 **Override 편집** 테이블에서 Override를 추가 또는 편집할 시간 구간을 선택합니다. Ctrl 키를 누른 상태에서 테이블의 행을 선택하여 여러 시간 구간을 선택할 수 있습니다.
- 2 개별 시간 구간에 대한 Override를 생성하려면 해당 시간 구간에 대한 **Override** 칼럼의 셀을 클릭합니다. Override 값을 입력하고 Enter를 누릅니다.
선택한 시간 구간에 대한 Override를 계산하려면 **Override 계산**을 클릭합니다. **Override 계산기** 대화 상자가 나타납니다.
이 세션 동안 지정한 Override를 제거하려면 **재설정**을 클릭합니다. 이전에 지정했거나 **SAS Forecast Studio**에서 지정한 Override는 제거되지 않습니다.
- 3 현재 시계열에 대한 Override를 생성한 후 **실행**을 클릭합니다. **SAS Forecast Server**가 프로젝트를 조정합니다. 프로젝트가 조정되면 **예측 데이터 Override** 창이 닫히고 마법사의 예측 시계열 선택 단계로 돌아옵니다. 트리 뷰에서 아이콘이 Override를 실행한 시계열 옆에 표시됩니다.

주: 프로젝트를 보정할 때 **SAS Forecast Server**에서 Override 충돌을 탐지하면 프로젝트를 보정할 수 없으며 **SAS Forecast Studio**를 사용하여 충돌을 보정해야 한다는 오류 메시지가 표시됩니다.

선택한 시간 구간에 대한 Override 계산

다음 옵션을 사용하여 하나 이상의 시간 구간에 대한 Override를 계산할 수 있습니다.

옵션 이름	설명
기준 예측 설정	조정된 통계 예측에 대해 지정된 증가 또는 감소값을 기준으로 Override가 지정되도록 할 수 있습니다.

옵션 이름	설명
	주: 데이터를 계층적으로 예측하지 않았거나 조정 예측을 사용할 수 없을 때에는 이 옵션을 선택하여 통계 예측에 대해 지정된 증가 또는 감소값을 기준으로 Override가 지정되도록 할 수 있습니다. 드롭다운 리스트 및 텍스트 상자를 사용하여 현재 값을 늘리거나 줄일 단위 수 또는 백분율을 지정합니다. 증가 또는 감소 옵션을 선택합니다. 선택한 옵션의 텍스트 상자에서 값을 지정하고 드롭다운 리스트에서 이 값이 단위인지 백분율인지를 지정합니다. 예를 들어 조정된 통계 예측보다 10% 더 큰 Override를 생성한다고 가정해 보겠습니다. 증가 옵션을 선택합니다. 텍스트 상자에 10을 입력하고 드롭다운 리스트에서 %(백분율)를 선택합니다. Override는 조정된 통계 예측보다 10% 더 큼니다.
선택한 주기 간의 값 배포	단일 Override를 선택한 모든 주기 사이에 배분합니다. 텍스트 상자에 이 Override를 지정합니다. Override가 적용되는 방식은 선택하는 옵션에 따라 다릅니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 각 기간의 현재 예측에 비례하여 배분 - 지정하는 값이 선택한 시간 기간 사이에 비례적으로 배분됩니다. 예를 들어 예측값이 550과 430인 두 개의 시간 기간을 선택합니다. 값 100을 지정하고 이 값이 두 시간 기간 사이에 비례적으로 배분되도록 선택합니다. 그러면 예측값은 606.12와 473.88이 됩니다. ■ 기간 내에서 동등하게 배분 - 지정하는 값이 선택한 시간 기간 사이에 균등하게 배분됩니다. 예를 들어 예측값이 550과 430인 두 개의 시간 기간을 선택합니다. 값 100을 지정하고 이 값이 두 시간 기간 사이에 균등하게 배분되도록 선택합니다. 그러면 예측값은 600과 480이 됩니다.
선택한 각 주기에 값 할당	단일 Override를 선택된 모든 기간에 적용합니다. 텍스트 상자에 이 Override를 지정합니다.

선택한 시간 기간에 대한 Override를 계산했으면 **확인**을 클릭하여 Override를 실행합니다.

Override 충돌 해결

프로젝트를 보정할 때 **SAS Forecast Server**에서 Override 충돌을 탐지하면 프로젝트를 보정할 수 없으며 **SAS Forecast Studio**를 사용하여 충돌을 보정해야 한다는 오류 메시지가 표시됩니다. **SAS Forecast Studio**에서 실행된 모든 Override를 확인하고 예측 데이터에 적용할 Override를 선택할 수 있습니다.

Override 충돌을 해결하는 방법:

- 1 **Forecast Studio Override 실행** 마법사를 종료하고 Microsoft Excel을 닫으려면 **취소**를 클릭하십시오.

- 2 **SAS Forecast Studio**에서 프로젝트를 엽니다. **SAS Forecast Studio**에서 프로젝트를 열면 "보정 필요"라는 경고가 작업 공간 맨 위에 표시됩니다.
- 3 이 경고 옆에 있는 **업데이트**를 클릭합니다. **Override 충돌** 대화 상자가 나타납니다.
- 4 Override 충돌을 해결하려면 **Override 충돌** 대화 상자를 사용하여 Override를 수정하거나 이 대화 상자를 닫고 데이터 테이블을 사용하여 Override를 업데이트합니다.

Override 충돌을 해결하는 방법에 대한 자세한 내용은 **SAS Forecast Studio**의 도움말을 참조하십시오.

일반화선형모델

일반화선형모델 작업 정보	195
일반화선형모델: 분석 역할에 변수 할당	196
일반화선형모델: 모델 생성	197
일반화선형모델: 모델 옵션 설정	197
일반화선형모델: Post Hoc 검정 선택	199
일반화선형모델: 도표 생성	199
일반화선형모델: 예측값 옵션 설정	200
제목 및 각주 지정	201
속성 보기	201

일반화선형모델 작업 정보

일반화선형모델 작업은 이산형으로 간주되는 횟수 또는 측정 비율과 같이 정규분포를 따르지 않는 데이터를 모델화할 수 있도록 일반 선형모델을 확장한 것입니다. 또한 일반화선형모델 작업을 사용하면 평균이 특정 범위의 값으로 제한된 데이터 또는 모든 관측값에 대해 분산이 일정하지 않은 데이터에도 모형을 생성할 수 있습니다.

이 작업을 사용하여 다음과 같은 작업을 수행할 수 있습니다.

- 두 제조사 제품의 수명 비교. 예를 들어 기계 부품의 고장 횟수가 있고, 그 중 일부는 제조업체 A에서, 일부는 제조업체 B에서 제조했다고 가정합니다. 이 작업을 사용하여 두 제조업체의 기계 부품 수명의 차이를 테스트할 수 있습니다.
- 이 작업으로 기계 부품의 수명에 있어서의 두 제조사의 차이점을 검정할 수 있습니다. 예를 들어 3개 브랜드의 아이스크림에 대한 가상의 미각 검정 결과를 포함하는 데이터셋이 있다고 가정합니다. 세 개의 브랜드는 매우 좋음(vg)에서 매우 나쁨(vb)까지 5점 척도로 등급이 매겨집니다. 세 브랜드에 대한 등급의 차이를 평가하기 위해 분석이 수행됩니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GENMOD
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

일반화선형모델: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
종속변수	종속(반응)변수로 사용할 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
양적변수	연속독립(설명)변수(효과라고도 함)로 사용할 변수를 지정합니다.
분류변수	이산형 독립변수 효과로 사용할 변수를 지정합니다. 이 역할에는 숫자 또는 문자 변수를 할당할 수 있으나 이러한 변수는 한정된 수의 이산 값을 가지고 있어야 합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 종속변수나 독립변수로 이미 선택한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
상대 가중값 변수	반응변수로 선택한 변수의 대응값에 대해 상대 가중값으로 사용할 변수를 하나 지정합니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.

역할 이름	설명
-------	----

상대 오프셋	오프셋 변수로 사용할 변수를 지정합니다. 주: 이 역할에 할당할 수 있는 변수는 최대 한 개입니다. 종속변수, 독립변수 또는 분류변수로 이미 선택한 변수로는 오프셋을 수행할 수 없습니다.
---------------	---

일반화선형모델: 모델 생성

여러 SAS 작업에 모델을 생성할 수 있습니다. 자세한 내용은 [부록 3, “모델 생성” \(609페이지\)](#)에서 참조하십시오.

일반화선형모델: 모델 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **모델 옵션**을 클릭합니다.

옵션 이름	설명
-------	----

분포	모델의 반응변수 확률 분포를 지정합니다. 주: 이항 분포는 종속변수에 두 개의 레벨이 있을 때만 사용할 수 있습니다. 주: 종속변수 역할에 문자 변수가 할당되면 다항 및 이항 분포만 사용할 수 있습니다.
-----------	--

연결 함수	선형예측변수에 대한 기대값의 관계를 설명합니다. 생성하는 모델에 따라 여러 가지 연결 함수를 사용할 수 있습니다.
--------------	---

연결 함수 드롭다운 리스트의 다음 옵션 중에서 선택할 수 있습니다.

- 없음
- 항등
- 로짓
- 프로빗
- 로그
- 보 로그-로그
- 수의 거듭제곱

옵션 이름	설명
	주: 누적 로짓, 누적 보 로그-로그 및 누적 프로빗 연결 함수는 다항 분포를 선택할 때만 사용할 수 있습니다. 연결 함수 드롭다운 리스트에서 연결 함수를 선택하지 않고 사용자 정의 연결 함수가 있으면 사용자 정의 연결 함수가 사용됩니다. 사용자 정의 연결 함수를 지정하지 않으면 모델 분포에 대한 기본 연결 함수가 사용됩니다. ■ 이항 분포에서는 기본 연결 함수가 로짓입니다. ■ 감마 분포에서는 기본 연결 함수가 역(역(-1))입니다. ■ 역가우시안 분포에서는 기본 연결 함수가 역 제곱(역(-2))입니다. ■ 다항 분포에서는 기본 연결 함수가 누적 로짓입니다. ■ 음이항 분포에서는 기본 연결 함수가 로그입니다. ■ 정규 분포에서는 기본 연결 함수가 항등입니다. ■ 포아송 분포에서는 기본 연결 함수가 로그입니다. 주: 분포 및 연결 함수에 대해 없음을 선택하면 연결 함수가 항등인 정규분포가 분석에 사용됩니다.
효과 검정	분석에 포함할 통계량을 지정합니다. ■ 연속하는 모델 쌍 간의 우도비 통계량을 표시하려면 Type 1 또는 순차 분석 계산 체크 박스를 선택합니다. 순차적 적합 모형은 영(절편항만 있음) 모형에서 시작하여 정의한 모형까지 계속됩니다. ■ 정의한 모델의 각 효과에 대한 대비 통계량을 표시하려면 Type 3 대비에 대한 통계량 계산 체크 박스를 선택합니다. 기본 분석은 대비에 대한 우도비 통계량을 계산하는 것입니다. wald 통계량 계산 체크 박스를 선택하면 wald 통계량도 계산됩니다.
신뢰한계	예측값에 대한 신뢰수준을 지정합니다. 기본값은 95% 신뢰수준입니다.
추정값의 상세 정보	모수 추정값에 대한 상관행렬을 표시할지, 모수 추정값에 대한 공분산 행렬을 표시할지 또는 두 가지 모두를 표시할지를 지정합니다.
신뢰구간	양측 신뢰구간을 사용할지를 지정합니다. ■ 프로파일 우도함수(부분최대우도함수라고도 함)를 기준으로 모든 모델 모수에 대한 양측 신뢰구간을 계산하려면 양측 체크 박스를 선택합니다. ■ 모수 추정값의 근사 정규성을 기준으로 모든 모델 모수에 대한 양측 wald 신뢰구간을 계산하려면 양측 wald 체크 박스를 선택합니다.

일반화선형모델: Post Hoc 검정 선택

이 옵션에 액세스하려면 선택 영역에서 **Post Hoc 검정**을 클릭합니다.

추정할 효과의 각 효과에 대해 **평균 검정을 위한 옵션** 상자에서 옵션을 지정합니다.

일반화선형모델: 도표 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

기본적으로 현재 선택된 데이터에 적용 가능한 모든 도표가 출력에 포함됩니다. 그러나 **사용자 정의 도표 리스트** 옵션을 선택하여 출력에 포함할 도표를 선택할 수 있습니다. 다음 옵션 중에서 선택할 수 있습니다.

옵션 이름	설명
Cook의 D 도표	Cook의 D 통계량 대 관측값 번호 도표를 생성합니다.
DFBETA 도표	모수 추정값에서 관측값을 제외한 차이에 대한 도표를 생성합니다.
DFBETAS 도표	모수 추정값에서 관측값을 제외한 표준화된 차이를 도표로 생성합니다. 개별 관측값이 적합 모델의 각 추정 모수에 미치는 영향을 평가하는 데 유용합니다.
지레값 도표	H 행렬에서 대응되는 대각원소의 값에 대한 도표를 생성합니다. 행렬의 대각원소는 더 큰 값을 보이는 경향이 있는 계획 공간에서 극값을 탐지하는 데 유용합니다.
예측값 도표	예측값 도표를 생성합니다.
피어슨 잔차 도표	모델로 제대로 설명되지 못한 관측값을 식별하는 피어슨(카이) 잔차 대 예측값에 대한 도표를 생성합니다.
이탈도 잔차 도표	적합이 잘되지 않은 값을 식별하는 이탈도 잔차 대 예측값에 대한 도표를 생성합니다.
우도 잔차 도표	적합이 잘되지 않은 값을 식별하는 우도 잔차 대 예측값에 대한 도표를 생성합니다.
원잔차 도표	원잔차 대 예측값 도표를 생성합니다.

옵션 이름	설명
표준화 피어슨 잔차	모델로 제대로 설명되지 못한 관측값을 식별하는 표준화 피어슨(카이) 잔차 대 예측값에 대한 도표를 생성합니다.
표준화 이탈도 잔차	적합이 잘되지 않은 값을 식별하는 표준화 이탈도 잔차 대 예측값에 대한 도표를 생성합니다.

일반화선형모델: 예측값 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **예측값**을 클릭합니다.

옵션 이름	설명
예측할 데이터	출력 데이터를 예측하는 데 사용할 데이터 소스를 식별합니다. 추가 데이터 체크 박스를 선택하면 이 데이터가 포함됩니다. 추가 데이터의 위치를 지정하려면 찾아보기 를 클릭합니다.
출력 데이터 저장	예측통계량 및 진단통계량을 출력 데이터셋에 저장할지를 지정합니다. 주: SAS Enterprise Guide는 옵션 창의 작업> 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기 를 클릭합니다.
추가적인 통계량	추가적인 통계량을 저장할지를 지정합니다. ■ 잔차 - 부적합한 관측값을 식별하는 우도잔차, 모델로 잘 설명되지 않는 관측값을 식별하는 피어슨(카이) 잔차 및 부적합한 관측값을 식별하는 이탈도 잔차를 저장합니다. ■ 예측한계 - 예측값에 대한 신뢰상한 및 하한을 저장합니다.
예측값 표시	결과에 예측값을 표시합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

High-Performance 선형회귀

<i>High-Performance 선형회귀 작업 정보</i>	203
<i>High-Performance 선형회귀: 분석 역할에 변수 할당</i>	204
<i>High-Performance 선형회귀: 모델 생성</i>	205
모델 생성	205
모형 선택 방법 지정	205
<i>High-Performance 선형회귀: 모델 옵션 설정</i>	208
<i>High-Performance 선형회귀: 파티션 옵션 설정</i>	208
<i>High-Performance 회귀: 성능 옵션 설정</i>	209
<i>High-Performance 회귀: 예측값 옵션 설정</i>	210
<i>제목 및 각주 지정</i>	210

High-Performance 선형회귀 작업 정보

High-Performance 선형회귀 작업은 일반적인 최소제곱 모델에 대한 모델 선택을 적합하고 수행합니다. 선형 회귀작업과 달리 분류 변수를 포함하는 일반적인 선형 모델을 지정할 수 있습니다. High-Performance 선형회귀 작업은 LAR(Least Angle Regression) 및 Lasso 선택 방법도 지원합니다.

High-Performance 선형회귀 작업은 분류변수에 대해 두 가지 모수화(GLM 및 참조)를 허용합니다. 다중 모수화를 허용하는 다른 작업과 달리 이 작업의 분류변수는 동일한 모수화를 가집니다. GLM 모수화가 기본값입니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	HPREG
필수 SAS 제품	Base SAS, SAS High-Performance Analytics Server, SAS/STAT

요구 사항 이름	프로시저 및 제품 이름
추가 SAS 제품(권장)	SAS/GRAPH

High-Performance 선형회귀: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
종속변수	종속(반응)변수로 사용할 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
설명 효과	회귀(설명)변수로 사용할 변수를 지정합니다. 회귀변수를 지정하지 않으면, 각 종속변수는 평균만 가진 모델에 적합됩니다. 즉, 각 종속변수는 상수항만 가진 모형에 적합됩니다.
분류변수	분석에서 사용할 분류변수를 지정합니다. 문자 또는 숫자 변수가 될 수 있습니다. 분류변수가 숫자이면 값은 연속형이 아닌 이산형이어야 합니다. 이러한 변수가 고유티값이 매핑될 레벨을 통해 분석을 입력합니다. 값이 레벨화되면 같은 값을 공유하는 관측값은 같은 레벨로 할당됩니다. 모든 분류 변수 또는 개별 분류 변수에 대한 코딩 스타일을 지정할 수 있습니다. 모든 분류 변수에 대해 동일한 코딩 스타일을 지정하려면, 글로벌 분류 변수 옵션 영역에서 옵션을 사용합니다. ■ GLM 옵션을 전체 순위보다 낮은 참조 셀 코딩에 사용합니다. 이 모수화는 예를 들어 SAS/STAT의 GLM, MIXED 및 GLIMMIX 프로시저에서 사용됩니다. ■ 참조 셀 인코딩에 대해 우선 설정 옵션을 사용합니다. 첫 번째 또는 마지막 순서의 값을 사용할지를 지정합니다. 참조를 선택하면, 개별 분류변수에 대해 참조 셀 인코딩을 덮어쓸 수 있습니다. 분할 체크 박스는 분할 분류변수를 포함하는 효과에 대응할 디자인 행렬 칼럼을 지정합니다. 해당 효과의 기타 디자인 칼럼에 독립적으로 모델에 들어가거나 나오도록 각 칼럼을 선택할 수 있습니다. 글로벌 옵션 또는 개별 분류변수의 옵션으로 지정할 수 있습니다. Temp 변수에 hot, Warm, cold 값이 있는 세 개의 레벨이 있고 성별 변수에 M 및 F 값을 가진 두 개의 레벨이 있다고 가정합니다.

역할 이름	설명
	■ 분할이 글로벌 옵션으로 선택되면, 두 효과가 8개의 독립효과로 분할됩니다. 효과 성별은 두 개의 효과인 gender_M 및 gender_F로 분할됩니다. 효과 gender*temp는 gender_M*temp_hot, gender_F*temp_hot, gender_M*temp_warm, gender_F*temp_warm, gender_M*temp_cold, gender_F*temp_cold 등 6개 효과로 분할됩니다. ■ 임시 변수에 대한 옵션으로 분할을 선택하면 성별*임시 효과는 세 가지 효과인 *gender*temp_hot, gender*temp_Warm, *gender*temp_cold로 분할됩니다. 이러한 세 개의 분할 효과는 각각 두 성별 레벨에 대응하는 두 모수를 현재 가지고 있습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
상대 가중값 변수	가중 최소제곱적합에 대한 상대 가중변수의 값을 포함하는 변수를 지정합니다. 상대 가중변수의 값은 음이 아니어야 합니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.

High-Performance 선형회귀: 모델 생성

모델 생성

여러 SAS 작업에 모델을 생성할 수 있습니다. 자세한 내용은 [부록 3, “모델 생성” \(609페이지\)](#)에서 참조하십시오.

모형 선택 방법 지정

이 옵션에 액세스하려면 선택 영역의 모델 머리글 아래에서 **선택**을 클릭합니다.

옵션 이름	설명
모형 선택 방법	모형 선택 방법 드롭다운 리스트에서 방법을 선택합니다. 방법을 선택하지 않으면 전체 모형이 사용됩니다. 모형 선택 방법에 따라 해당 방법에 대한 추가 옵션을 사용할 수 있습니다. 사용할 수 있는 방법은 다음과 같습니다.

옵션 이름	설명
	■ 전체 모형 적합(선택 없음) - 기본값이며 효과 선택을 제공하지 않습니다. 여기에서 모형은 모형 적합에 사용되는 종속변수와 양적변수를 할당할 때 생성됩니다. ■ 전진선택 - 효과가 없는 모델에서 시작합니다. 각 설명변수에 대해 이 방법은 변수가 포함되었을 때 모델에 대한 변수의 기여도를 반영하는 F 통계량을 계산합니다. F 통계량에 대한 p-값은 모델에서 변수를 포함하도록 지정된 유의수준과 비교됩니다. 기본적으로 이 값은 0.05입니다. 유의수준을 변경하려면 모형에 추가되는 기준 텍스트 상자에 값을 입력합니다. 이 값보다 큰 유의수준을 가지는 F 통계량이 없으면 전진선택은 중단됩니다. 그렇지 않으면 전진선택법은 가장 큰 F 통계량을 가진 변수를 모델에 추가합니다. 그런 다음, 모델 외부에 남아 있는 변수에 대해 F 통계량을 다시 계산하여 평가 절차를 반복합니다. 결국, 유의한 F 통계량을 생성하는 변수가 없을 때까지 변수는 모델에 하나씩 추가됩니다. 변수가 모형에 추가되면 변수는 모델에서 계속 유지됩니다. ■ 후진제거 - 모든 설명변수를 포함하여 모델에 대한 F 통계량을 계산함으로써 시작됩니다. 모형에 남아 있는 모든 변수가 유의한 F 통계량을 계산할 때까지 변수가 모델에서 하나씩 삭제됩니다. 유의수준은 모형에 남는 기준 텍스트 상자에 지정됩니다. 기본적으로 이 값은 0.05입니다. 각 단계에서 모델에 대한 최소 분포를 보여 주는 변수가 삭제됩니다. ■ 단계별 선택 - 전진선택법의 변형입니다. 단계별 방법에서 모델에 이미 있는 변수가 반드시 유지될 필요가 없습니다. 전진선택법은 F 통계량이 유의하면 모델에 변수를 하나씩 추가합니다. 유의수준은 모델에 추가되는 기준 텍스트 상자에 지정됩니다. 그러나, 단계별 선택법은 변수가 추가된 후 이미 모델에 포함된 변수를 모두 확인하여 유의한 F 통계량을 생성하지 않은 변수를 모두 삭제합니다. 유의수준은 모형에 남는 기준 텍스트 상자에 지정됩니다. 이러한 확인이 끝나고 해당 변수가 삭제된 후에만 다른 변수가 모델에 추가될 수 있습니다. 단계별 프로세스는 다음 조건 중 하나에 충족되면 종료됩니다. □ 모델 외부의 변수가 유의한 F 통계량을 가지지 않고 모델의 모든 변수가 모델에 남도록 지정된 유의수준에서 유의할 때 □ 모델에 추가될 변수가 방금 삭제된 변수일 때 ■ 전진스왑 선택 - 확장된 전진선택법입니다. 이 방법은 선택 기준을 향상시키는 현재 모델의 안팎에서 으로 전환하거나 종료하는 쌍별 효과 스왑을 사용합니다. ■ LARS(Least angle regression) - 효과가 없는 모델에서 시작해서 효과를 추가합니다. 어떤 단계에서든 모수 추정값은 대응하는 최소제곱값과 비교할 때 더 작습니다. 모델에 분류변수가 포함되면 이러한 분류변수는 분할됩니다. ■ Lasso 방법 - 절대 회귀계수 합계가 제한되는 보통 최소제곱 버전에 근거하여 모수를 추가하거나 삭제합니다. 모델에 분류변수가 포함되면 이러한 분류변수는 분할됩니다.
유의수준	전진선택 및 LARS 방법에서는 모델에 양적(설명)변수를 추가하기 위한 유의수준을 지정합니다. 후진제거법에서는 양적변수를 모델에 남기기 위한

옵션 이름	설명
	유의수준을 지정합니다. 단계별 선택 및 Lasso 방법에서는 두 가지 유의수준을 모두 지정합니다.
최적값을 가진 모델 선택	기준을 사용하여 최적값을 산출하는 모델을 선택 프로세스의 각 단계에서 선택합니다. 둘 이상의 단계 모델에서 지정된 기준의 최적값이 발생하면 최소 모수 개수를 가진 모델이 선택됩니다. 기준을 지정하지 않았으면, 선택한 모델은 선택 프로세스의 마지막 단계에 있는 모델이 됩니다. 사용 가능한 기준: ■ AIC - Akaike의 정보기준 ■ AICC - Akaike의 정보기준의 수정 버전인 작은 표본의 특성 ■ BIC - Schwarz Bayesian 기준 ■ ADJRSQ - 조정 R-제곱 기준 ■ CP - Mallows' C_p 통계량 ■ PRESS - 제곱 통계량의 예측 잔차 합계 ■ VALIDATE - 평가용 데이터에 대한 평균제곱오차
선택 프로세스 종단을 위한 기준 선택	기준을 사용하여 선택 프로세스를 중단합니다. 안 함을 선택하면 선택 프로세스는 알맞은 추가 또는 제거 후보를 찾을 수 없거나 크기 한계에 도달할 때 중단됩니다. 사용 가능한 기준: ■ SL - 스코어 검증 유의 ■ AIC - Akaike의 정보기준 ■ AICC - Akaike의 정보기준의 수정 버전인 작은 표본의 특성 ■ BIC - Schwarz Bayesian 기준 ■ ADJRSQ - 조정 R-제곱 기준 ■ CP - Mallows' C_p 통계량 ■ PRESS - 제곱 통계량의 예측 잔차 합계 ■ VALIDATE - 평가용 데이터에 대한 평균제곱오차
모형에 포함시킬 효과	모형에 추가할 효과를 선택합니다. 변수가 모형에 포함시킬 효과 상자에 추가됩니다. 변수를 선택하고 화살표 버튼을 클릭하여 이 변수의 순서를 재지정할 수 있습니다. 주: 전체 모형 적합(선택 없음), LARS(Least angle regression) 또는 Lasso 방법 모형 선택 방법을 선택하면 이 옵션을 사용할 수 없습니다.

High-Performance 선형회귀: 모델 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

옵션 이름	설명
추정값의 상세 정보	추정값의 상세 정보 영역에서 다음 통계량을 선택할 수 있습니다. ■ 표준화회귀계수 - 표준화회귀계수를 생성합니다. 표준화된 회귀계수는 회귀변수의 표본 표준편차에 대한 종속변수의 표본 표준편차 비율로 모수 추정값을 나눈 값입니다. ■ 모수 추정값에 대한 신뢰한계 - 모수 추정값에 대한 신뢰상한 및 하한을 요청합니다. 기본적으로 95% 한계가 계산됩니다.
공선성 진단	다음과 같은 진단을 선택할 수 있습니다. ■ 추정 허용값 - 추정치에 대한 허용값을 생성합니다. 모수 허용은 $1 - R^2$로 정의되며, 여기서 R^2는 모델의 기타 모든 모수의 모수 회귀분석에서 가져옵니다. ■ 분산팽창값 - 모수 추정값으로 분산팽창 요인을 생성합니다. 분산팽창값은 허용값의 상대적 개념입니다. 주: LARS(Least angle regression) 또는 Lasso 방법 모형 선택 방법을 선택하면 이 옵션을 사용할 수 없습니다.

High-Performance 선형회귀: 파티션 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **파티션**을 클릭합니다.

옵션 이름	설명
모델 분석, 평가 및 검정을 위해 데이터를 나누지 않음	관측값을 나누지 않도록 지정

옵션 이름	설명
관측값 임의 할당에 대해 사용할 비율 지정	논리적으로 입력 데이터셋의 관측값을 모델 분석, 평가 및 검증에 대한 배반 서브셋으로 나누도록 지정합니다. 관측값 비율이 분석 및 평가 역할에 임의로 할당되도록 지정할 수 있습니다. ■ 분석 비율 - 분석 역할로 임의 할당된 입력 데이터셋의 관측값 비율을 표시합니다. 평가 또는 검증 역할에 할당되지 않은 관측값은 분석 역할로 할당됩니다. ■ 평가 비율 - 평가 역할로 임의 할당된 입력 데이터셋의 관측값 비율을 지정합니다. ■ 검증 비율 - 검증 역할로 임의 할당된 입력 데이터셋의 관측값 비율을 지정합니다.

High-Performance 회귀: 성능 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **성능**을 클릭합니다.

옵션 이름	설명
이 작업에 대해 기본 설정 사용	기본 설정을 사용합니다.
이 작업에 대한 사용자 정의 설정 사용	작업에 대한 노드, 그리드 호스트, 설치 위치 및 데이터 서버를 지정할 수 있습니다. ■ 노드 - 데이터가 데이터베이스에 따라 처리되지 않음을 고려할 때 분산 컴퓨팅 환경의 노드 개수를 지정합니다. ■ 그리드 호스트 - 어플라이언스 호스트의 이름을 지정합니다. ■ 설치 위치 - 어플라이언스에서 High-Performance Analytics 공유 라이브러리가 설치되는 디렉터리를 지정합니다. ■ 데이터 서버 - 호스트 파일을 통해 정의되고 Teradata에 대한 LIBNAME 문에서 사용되는 Teradata 시스템의 서버 이름을 지정합니다.
상세 프로시저 단계 표시	프로시저 단계의 통계량 계산을 표시하는 테이블을 생성합니다.

High-Performance 회귀: 예측값 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **예측값**을 클릭합니다.

도표 이름	설명
출력 데이터 저장	예측통계량 및 진단통계량을 출력 데이터셋에 영구히 저장할지를 지정합니다. ■ 예측값 - 각 반응레벨의 예측 확률을 저장합니다. 세 개의 레벨 1, 2, 3이 있는 반응변수 Y에 대한 개별 확률은 $\Pr(Y=1)$, $\Pr(Y=2)$, $\Pr(Y=3)$입니다. ■ 진단 통계량 - DFBETAS 진단 통계량, 신뢰구간 변위 진단(C 및 CBAR), H 행렬, 표준오차 추정값을 가진 선형예측변수 추정값을 저장합니다. 주: SAS Enterprise Guide는 옵션 창의 작업> 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
추가적인 통계량	■ 잔차를 선택하여 다음과 같은 통계량을 저장합니다. □ 이탈도의 변경 사항 □ 이탈도 잔차 □ 피어슨(카이제곱) 잔차 □ 개별 관측값 삭제로 인한 이탈도의 변경 사항 □ 개별 관측값 삭제로 인한 카이제곱 적합도 통계량의 변경 사항 ■ 예측한계를 선택하면 각 예측값에 대한 신뢰구간의 상한 및 하한값과 종속변수의 예측값 평균이 출력에 포함됩니다.
출력 변수 선택	출력 데이터셋에 포함할 칼럼을 지정합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

.....
주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.
.....

High-Performance 로지스틱 회귀

<i>High-Performance 로지스틱 회귀 작업 정보</i>	213
<i>High-Performance 로지스틱 회귀: 분석 역할에 변수 할당</i>	214
<i>High-Performance 로지스틱 회귀: 모델 생성</i>	215
반응 지정	215
모델 생성	216
모형 선택 방법 지정	216
<i>High-Performance 로지스틱 회귀: 모형 옵션 설정</i>	218
<i>High-Performance 회귀: 성능 옵션 설정</i>	219
<i>High-Performance 회귀: 예측값 옵션 설정</i>	219
<i>제목 및 각주 지정</i>	220

High-Performance 로지스틱 회귀 작업 정보

High Performance 로지스틱 회귀 작업은 SAS High-Performance Analytics 환경에서 이진, 이항 및 다항 데이터에 대한 로지스틱 회귀 모델을 적합시킵니다. High-Performance 로지스틱 회귀 작업은 분류변수에 대해 두 가지 모수화(GLM 및 참조)를 허용합니다. 다중 모수화를 허용하는 다른 작업과 달리 이 작업의 분류변수는 동일한 모수화를 가집니다. GLM 모수화가 기본값입니다.

High-Performance 로지스틱 회귀 작업은 High-Performance 분산 환경에서 작동하도록 특수하게 고안되었습니다. 기본적으로 여러 스레드에서 계산이 수행됩니다. 로지스틱 회귀 작업은 단일 스레드를 사용합니다.

주: High-Performance 로지스틱 회귀 작업을 열면 **SAS Enterprise Guide**가 Binary 값을 가지는 변수를 지정할 때 약간의 지연이 있음을 알 수 있습니다. 이 분석이 완료되면, 작업에 **종속변수** 역할에 지정된 변수에 관한 정보가 나타납니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	HPLOGISTIC
필수 SAS 제품	Base SAS, SAS High-Performance Analytics Server, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

High-Performance 로지스틱 회귀: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
종속변수	종속(반응)변수로 사용할 변수를 지정합니다. 이 역할에는 하나의 변수만 지정해야 합니다.
양적변수	독립(설명)변수로 사용할 변수를 지정합니다.
분류변수	분석에서 사용할 분류변수를 지정합니다. 문자 또는 숫자 변수가 될 수 있습니다. 분류변수가 숫자이면 값은 연속형이 아닌 이산형이어야 합니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
상대 가중값 변수	반응변수로 선택한 변수의 대응값에 대해 상대 가중값으로 사용할 변수를 하나 지정합니다. 상대 가중값이 음수이거나 결측값이면 분석에서 제외됩니다.

High-Performance 로지스틱 회귀: 모델 생성

반응 지정

이 옵션에 액세스하려면 선택 영역의 모델 머리글 아래에서 **반응**을 클릭합니다.

역할 이름	설명
반응 유형	사용할 수 있는 반응 유형은 반응변수의 레벨 수에 따라 결정됩니다. ■ 반응변수에 레벨이 두 개만 있으면 반응 유형은 이진(binary)이 됩니다. ■ 반응변수에 레벨이 두 개 이상 있으면 반응 유형을 순서가 있음 또는 순서가 없음 중에서 선택할 수 있습니다.
모델 유형	다음과 같은 모델 유형 중에서 선택할 수 있습니다. ■ 로짓 - 로그 오즈 함수의 결과를 표시합니다. 두 가지 반응 범주가 있으면 이진 로짓 모델이 사용됩니다. 셋 이상의 반응 범주가 있으면 누적 로짓 모형이 사용됩니다. ■ 프로빗 - 역표준정규확률적분함수의 결과를 표시합니다. 두 가지 반응 범주가 있으면 이진 프로빗 모델이 사용됩니다. 셋 이상의 반응 범주가 있으면 누적 프로빗 모형이 사용됩니다. ■ 보 로그-로그 - 보 로그-로그 함수의 결과를 표시합니다. 두 가지 반응 범주가 있으면 이진 보 로그-로그 모델이 사용됩니다. 셋 이상의 반응 범주가 있으면 누적 보 로그-로그 모형이 사용됩니다. ■ 로그-로그 - 로그-로그 함수의 결과를 표시합니다. 두 가지 반응 범주가 있으면 이진 로그-로그 모델이 사용됩니다. 셋 이상의 반응 범주가 있으면 누적 로그-로그 모형이 사용됩니다. ■ Glogit - 일반화된 로짓 함수의 출력을 표시합니다. 일반화된 로짓 모델에서 각 비참조 범주는 참조 범주와 대조를 이룹니다. 주: 이 옵션은 반응 유형으로 순서가 없음을 지정했을 때에만 사용할 수 있습니다.
반응 레벨	반응 레벨 상자에는 반응변수에 대한 출력형식 레벨이 오름차순으로 나열되어 있습니다.
다음 레벨에 모형 적합	레벨에 모형을 적합시키는 방법을 지정할 수 있습니다. ■ 이진을 반응 유형으로 선택했으면 다음 레벨에 모형 적합 드롭다운 리스트에서 이진 반응 모형의 사건 범주를 선택합니다.

역할 이름	설명
	■ 순서가 없음을 반응 유형으로 선택했으면 다음 레벨에 모형 적합 드롭다운 리스트에서 일반화된 로짓 모델의 참조 범주를 선택합니다. 각 로짓에서 비참조 범주와 참조 범주의 차이점이 표시됩니다. ■ 순서가 있음을 반응 유형으로 선택했으면 다음 레벨에 모형 적합 드롭다운 리스트의 값을 사용하여 반응 범주의 순서를 지정할 수 있습니다. 정렬된 값 중 첫 번째 값을 선택하면 반응 범주가 반응 레벨 상자에 표시되는 순서로 정렬됩니다. 정렬된 값 중 마지막 값을 선택하면 반응 범주가 역순으로 정렬됩니다.

모델 생성

여러 SAS 작업에 모델을 생성할 수 있습니다. 자세한 내용은 [부록 3, “모델 생성” \(609페이지\)](#)에서 참조하십시오.

모형 선택 방법 지정

이 옵션에 액세스하려면 선택 영역의 모델 머리글 아래에서 **선택**을 클릭합니다.

옵션 이름	설명
모형 선택 방법	모형 선택 방법 드롭다운 리스트에서 방법을 선택합니다. 방법을 선택하지 않으면 전체 모형이 사용됩니다. 모형 선택 방법에 따라 해당 방법에 대한 추가 옵션을 사용할 수 있습니다. 사용할 수 있는 방법은 다음과 같습니다. ■ 전체 모형 적합(선택 없음) - 기본값이며 효과 선택을 제공하지 않습니다. 여기에서 모형은 모형 적합에 사용되는 종속변수와 양적변수를 할당할 때 생성됩니다. ■ 전진선택 - 효과가 없는 모델에서 시작합니다. 각 설명변수에 대해 이 방법은 변수가 포함되었을 때 모델에 대한 변수의 기여도를 반영하는 F 통계량을 계산합니다. F 통계량에 대한 p-값은 모델에서 변수를 포함하도록 지정된 유의수준과 비교됩니다. 기본적으로 이 값은 0.05입니다. 유의수준을 변경하려면 모형에 추가되는 기준 텍스트 상자에 값을 입력합니다. 이 값보다 큰 유의수준을 가지는 F 통계량이 없으면 전진선택은 중단됩니다. 그렇지 않으면 전진선택법은 가장 큰 F 통계량을 가진 변수를 모델에 추가합니다. 그런 다음, 모델 외부에 남아 있는 변수에 대해 F 통계량을 다시 계산하여 평가 절차를 반복합니다. 결국, 유의한

옵션 이름	설명
	F 통계량을 생성하는 변수가 없을 때까지 변수는 모델에 하나씩 추가됩니다. 변수가 모형에 추가되면 변수는 모델에서 계속 유지됩니다. ■ 후진제거 - 모든 설명변수를 포함하여 모델에 대한 F 통계량을 계산함으로써 시작됩니다. 모형에 남아 있는 모든 변수가 유의한 F 통계량을 계산할 때까지 변수가 모델에서 하나씩 삭제됩니다. 유의수준은 모형에 남는 기준 텍스트 상자에 지정됩니다. 기본적으로 이 값은 0.05입니다. 각 단계에서 모델에 대한 최소 분포를 보여 주는 변수가 삭제됩니다. ■ 단계별 선택 - 전진선택법의 변형입니다. 단계별 방법에서 모델에 이미 있는 변수가 반드시 유지될 필요가 없습니다. 전진선택법은 F 통계량이 유의하면 모델에 변수를 하나씩 추가합니다. 유의수준은 모델에 추가되는 기준 텍스트 상자에 지정됩니다. 그러나, 단계별 선택법은 변수가 추가된 후 이미 모델에 포함된 변수를 모두 확인하여 유의한 F 통계량을 생성하지 않은 변수를 모두 삭제합니다. 유의수준은 모형에 남는 기준 텍스트 상자에 지정됩니다. 이러한 확인이 끝나고 해당 변수가 삭제된 후에만 다른 변수가 모델에 추가될 수 있습니다. 단계별 프로세스는 다음 조건 중 하나에 충족되면 종료됩니다. □ 모델 외부의 변수가 유의한 F 통계량을 가지지 않고 모델의 모든 변수가 모델에 남도록 지정된 유의수준에서 유의할 때 □ 모델에 추가될 변수가 방금 삭제된 변수일 때 ■ 후진제거(모델 재적합없이 빠르게) - 빠른 후진제거를 수행합니다. 이 방법은 모델의 모든 효과로 시작하고 모델을 재적합하지 않고 효과를 삭제합니다.
유의수준	■ 전진선택법에서는 모델에 양적(설명)변수를 추가하기 위한 유의수준을 지정합니다. ■ 후진제거 및 후진제거(모델 재적합 없이 빠르게)법에서는 유의수준을 지정하여 양적변수가 모델에 남아 있도록 합니다. ■ 단계별 선택법에서는 두 가지 유의수준을 모두 지정합니다.
최적값을 가진 모델 선택	기준을 사용하여 최적값을 산출하는 모델을 선택 프로세스의 각 단계에서 선택합니다. 둘 이상의 단계 모델에서 지정된 기준의 최적값이 발생하면 최소 모수 개수를 가진 모델이 선택됩니다. 기준을 지정하지 않았으면, 선택한 모델은 선택 프로세스의 마지막 단계에 있는 모델이 됩니다. 사용 가능한 기준: ■ AIC - Akaike의 정보기준 ■ AICC - Akaike의 정보기준의 수정 버전인 작은 표본의 특성 ■ BIC - Schwarz Bayesian 기준
선택 프로세스 중단 을 위한 기준 선택	기준을 사용하여 선택 프로세스를 중단합니다. 안 함을 선택하면 선택 프로세스는 알맞은 추가 또는 제거 후보를 찾을 수 없거나 크기 한계에 도달할 때 중단됩니다. 사용 가능한 기준: ■ SL - 스코어 검증 유의

옵션 이름	설명
	■ AIC - Akaike의 정보기준 ■ AICC - Akaike의 정보기준의 수정 버전인 작은 표본의 특성 ■ BIC - Schwarz Bayesian 기준
모형에 포함시킬 효과	모델에 추가할 효과를 선택합니다. 변수가 모형에 포함시킬 효과 상자에 추가됩니다. 변수를 선택하고 화살표 버튼을 클릭하여 이 변수의 순서를 재지정할 수 있습니다. 주: 이 옵션은 모형 선택 방법으로 전체 모형(선택 없음)을 선택하면 사용할 수 없습니다.

High-Performance 로지스틱 회귀: 모형 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

옵션 이름	설명
모델 적합 평가	다음과 같은 모델 적합통계량을 지정합니다. ■ Hosmer와 Lemeshow 적합도 검정 - 이진 반응 모델에서는 Hosmer와 Lemeshow 적합도 검정을 표시합니다. 주: 이 옵션은 종속변수 역할에 이진 변수를 할당할 때만 사용할 수 있습니다. ■ 일반화 R-제곱 - 적합 모델에 대한 일반화 결정계수를 표시합니다.
신뢰수준	개별 왈드 검중에 기반하여 신뢰구간을 계산할지를 지정합니다. 신뢰수준 드롭다운 리스트에서 값을 선택하여 회귀모수 또는 오즈비의 신뢰구간에 대한 유의수준을 설정할 수도 있습니다.
클래스 레벨 정보 표시 억제	클래스 레벨 정보 테이블 표시를 억제합니다. 로지스틱 회귀 계수 계산 억제를 선택하여 공분산 행렬과 로지스틱 회귀 계수의 표준오차 계산을 억제합니다. 모델에 수천 개의 변수가 포함되어 있으면 회귀계수의 공분산 행렬과 표준오차를 파생하기 위한 Hessian 역행렬의 역방향 작업에 시간이 많이 걸릴 수 있습니다.
분류 변수 코딩 스타일	분류 변수에 대한 모수화 방법을 지정합니다. ■ GLM 옵션을 전체 순위보다 낮은 참조 셀 코딩에 사용합니다. ■ 참조 - 참조 셀 코딩에 사용됩니다.

옵션 이름	설명
CLASS 변수를 대문자로 변환	레벨화하기 전에 문자값의 분류변수 값을 대문자로 변환할지를 지정합니다. 예를 들어 이 옵션을 선택하고 분류변수의 값이 'a', 'A' 및 'b'인 경우 'a' 및 'A'는 동일한 레벨을 나타내며 분류변수는 두 개의 값, 'A' 및 'B'만 있습니다.

High-Performance 회귀: 성능 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **성능**을 클릭합니다.

옵션 이름	설명
이 작업에 대해 기본 설정 사용	기본 설정을 사용합니다.
이 작업에 대한 사용자 정의 설정 사용	작업에 대한 노드, 그리드 호스트, 설치 위치 및 데이터 서버를 지정할 수 있습니다. ■ 노드 - 데이터가 데이터베이스에 따라 처리되지 않음을 고려할 때 분산 컴퓨팅 환경의 노드 개수를 지정합니다. ■ 그리드 호스트 - 어플라이언스 호스트의 이름을 지정합니다. ■ 설치 위치 - 어플라이언스에서 High-Performance Analytics 공유 라이브러리가 설치되는 디렉터리를 지정합니다. ■ 데이터 서버 - 호스트 파일을 통해 정의되고 Teradata에 대한 LIBNAME 문에서 사용되는 Teradata 시스템의 서버 이름을 지정합니다.
상세 프로시저 단계 표시	프로시저 단계의 통계량 계산을 표시하는 테이블을 생성합니다.

High-Performance 회귀: 예측값 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **예측값**을 클릭합니다.

도표 이름	설명
출력 데이터 저장	예측통계량 및 진단통계량을 출력 데이터셋에 영구히 저장할지를 지정합니다. ■ 예측값 - 각 반응레벨의 예측 확률을 저장합니다. 세 개의 레벨 1, 2, 3이 있는 반응변수 Y에 대한 개별 확률은 $\Pr(Y=1)$, $\Pr(Y=2)$, $\Pr(Y=3)$입니다. ■ 진단 통계량 - DFBETAS 진단 통계량, 신뢰구간 변위 진단(C 및 CBAR), H 행렬, 표준오차 추정값을 가진 선형예측변수 추정값을 저장합니다. 주: SAS Enterprise Guide는 옵션 창의 작업 > 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
추가적인 통계량	■ 잔차를 선택하여 다음과 같은 통계량을 저장합니다. □ 이탈도의 변경 사항 □ 이탈도 잔차 □ 피어슨(카이제곱) 잔차 □ 개별 관측값 삭제로 인한 이탈도의 변경 사항 □ 개별 관측값 삭제로 인한 카이제곱 적합도 통계량의 변경 사항 ■ 예측한계를 선택하면 각 예측값에 대한 신뢰구간의 상한 및 하한값과 종속변수의 예측값 평균이 출력에 포함됩니다.
출력 변수 선택	출력 데이터셋에 포함할 칼럼을 지정합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

히스토그램

히스토그램 작업 정보	221
공정능력분석: 분석 역할에 변수 할당	222
히스토그램: 분포 선택	223
히스토그램: 분포 사용자 정의	224
분포 옵션에 액세스하는 방법	224
분포 모수 지정	224
커널 밀도함수 추정값 지정	224
공정능력분석: 도표 모양 옵션 설정	225
축 옵션 설정	225
참조선 추가	225
인셋 추가	225
기타 도표 옵션 설정	226
규격 한계 옵션 설정	228
공정능력분석: 분석 유형 지정	228
제목 및 각주 지정	229
속성 보기	229

히스토그램 작업 정보

공정능력분석은 통계적 관리 상태 공정의 출력분포를 해당 규격한계와 비교하여 규격과 일치하는지를 파악합니다. 공정능력분석에서는 히스토그램을 사용하여 통계 관리 상태 공정의 측정값 분포를 해당 규격한계와 비교합니다.

히스토그램 작업을 사용하여 제조 과정을 분석할 수 있습니다. 예를 들어 반도체 회사에서 인쇄 회로기판을 만드는 데 동판의 두께가 규격하한 3.45mm와 3.55mm 사이에 속하는지 확인하기 위해 표본을 추출했다고 가정합니다. 히스토그램의 각 막대는 관측값 백분율과 측정 두께가 표시됩니다. 규격상한 및 하한이 히스토그램에 포함되어 있어 해당 규격한계에 속하는 관측값의 백분율을 확인할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	CAPABILITY
필수 SAS 제품	Base SAS, SAS/QC, SAS/GRAPH
추가 SAS 제품(권장)	SAS/STAT

공정능력분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분석변수	분석할 숫자 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다. P-P 도표를 제외한 모든 도표에서는 변수 이름의 규격한계 영역에서 각 분석변수에 대한 상한, 목표 값 및 하한을 지정할 수 있습니다. 이 한계를 지정하면 정규성 검정을 보여 주는 테이블, 규격한계를 보여 주는 테이블 및 공정능력지수를 보여 주는 테이블이 자동으로 출력에 포함됩니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n은 해당 행의 빈도변수값입니다.
상대 가중값 변수	상대 가중값으로 사용할 변수를 지정합니다. 이 역할에 변수를 할당하면 각 통계량의 변수값이 가중 통계량 계산에 사용됩니다. 이 역할에는 변수를 하나만 할당할 수 있습니다. 주: 히스토그램, 확률 도표, P-P 도표, Q-Q 도표 또는 CDF 도표를 생성할 때는 상대 가중값 변수 역할에 변수를 할당하지 않는 것이 좋습니다. 이 역할에 변수를 할당했으면 출력을 해석할 때 주의해야 합니다.

역할 이름	설명
분류변수 (히스토그램 작업에 서만 사용 가능)	분류변수로 사용할 하나 이상의 변수를 지정합니다. 이 역할에 변수를 할당하면 전체 분류 레벨에서 분석변수의 분포를 비교할 수 있는 비교 히스토그램이 생성됩니다. 주: 비교 확률 도표 또는 비교 Q-Q 도표를 생성하려면 분포분석 작업을 사용하십시오.

히스토그램: 분포 선택

이 옵션에 액세스하려면 선택 영역의 분포 머리글 아래에서 **요약**을 클릭합니다.

하나의 히스토그램에 여러 개의 이론분포를 적합시킬 수 있습니다. 계산할 분포를 선택한 후 해당 분포의 모수값 및 선 속성을 지정하려면 선택 영역에서 분포 이름을 선택합니다.

옵션 이름	설명
분포	체크 박스를 사용하여 다음 추정된 밀도 곡선을 하나 이상 선택합니다. ■ 정규 - 평균(μ)과 표준편차(σ)를 가지는 정규밀도함수를 적합시킵니다. ■ 로그 정규 - 척도모수(ζ), 임계모수(θ) 및 형상모수(σ)를 가지는 로그 정규 밀도함수를 적합시킵니다. ■ 지수 - 척도모수(σ) 및 임계모수(θ)를 가지는 지수밀도함수를 적합시킵니다. ■ 와이블 - 척도모수(σ), 임계모수(θ) 및 형상모수(c)를 가지는 3-모수 와이블 밀도함수를 적합시킵니다. ■ 베타 - 척도모수(σ), 임계모수(θ) 및 형상모수(α 및 β)를 가지는 베타 밀도함수를 적합시킵니다. ■ 감마 - 척도모수(σ), 임계모수(θ) 및 형상모수(α)를 가지는 감마 밀도함수를 적합시킵니다. ■ 커널 - 비모수적 커널 밀도함수 추정값을 적합시킵니다. 주: 비교 히스토그램에서는 정규 분포와 커널 밀도함수 추정값만 사용할 수 있습니다.
그래프 스타일	출력에 사용할 그래프 스타일을 선택합니다. 기존 SAS 그래프나 ODS 통계 그래프 중에 선택할 수 있습니다. ODS 그래프에 대한 자세한 내용은 SAS Output Delivery System: User's Guide를 참조하십시오.

히스토그램: 분포 사용자 정의

분포 옵션에 액세스하는 방법

이 옵션에 액세스하려면 선택 영역의 분포 머리글 아래에서 사용자 정의할 분포 이름을 클릭합니다.

주: 이러한 분포 옵션은 **요약** 페이지에서 분포를 선택했거나 해당 분포 페이지에서 **분포 이름** 체크 박스를 선택할 때만 사용할 수 있습니다.

분포 모수 지정

정규, 로그 정규, 지수, 와이블, 베타 및 감마 분포에서는 각 분석변수 또는 모든 분석변수에 대해 모수를 지정할 수 있습니다. 기본적으로 작업에서는 모든 분석변수의 모수 값을 추정합니다.

각 분석변수에 대한 모수를 지정하는 방법:

- 1 **모든 변수에 분포 적용** 체크 박스의 선택을 해제합니다.
- 2 **분석변수** 상자에서 변수에 해당되는 체크 박스를 선택합니다.
- 3 모수값을 지정합니다.

커널 밀도함수 추정값 지정

비모수적 커널 밀도함수 추정값에 대해 대역폭(c) 및 커널 함수(k)를 지정할 수 있습니다. 기본적으로 이 작업은 근사 평균적분제곱오차(MISE)를 최소화하는 대역폭을 추정합니다. 이 대역폭을 추정하지 않으려면 **MISE 최소화** 체크 박스에 대한 선택을 취소합니다. 커널 밀도에 대해 표준화된 대역폭 모수를 지정하려면 텍스트 상자에 대역폭을 입력합니다. 각 값 사이에 공백을 넣어 최대 다섯 개의 값을 지정할 수 있습니다. 예를 들어 텍스트 상자에 **0.5 1.0 1.5** 값을 입력합니다.

주: **MISE 최소화** 체크 박스에 대한 선택을 취소하지 않으면 MISE는 지정할 수 있는 다섯 개의 대역폭 값 중 하나를 선택합니다.

커널 밀도 추정값을 계산하는 데 사용되는 커널 함수(정규, 2차 또는 3차)를 지정할 수도 있습니다. 기본 커널 함수는 **정규**입니다.

최대 다섯 개의 추정값에 대해 커널 함수를 지정할 수 있습니다. 하나 이상의 커널 함수를 지정하면 함수의 순서가 항상 정규, 2차, 3차 순으로 나열됩니다. 함수를 추가하거나 순서를 변경하려면 작업 코드를 편집해야 합니다.

대역폭보다 많은 수의 커널 함수를 지정하면 마지막 대역폭이 남아 있는 추정값에 대해 반복됩니다. 예를 들어 대역폭으로 0.5 및 1.0을 지정하고 정규, 2차 및 3차를 커널 함수로 지정하면 3차 함수에 대해 1.0 대역폭이 반복됩니다. 커널 함수보다 많은 수의 대역폭을 지정하면 마지막 커널 함수가 남아 있는 대역폭에 대해 반복됩니다.

예를 들어 대역폭으로 0.5, 1.0 및 1.5를 지정하고 정규 및 2차를 커널 함수로 지정하면 2차 함수가 1.5 대역폭에 사용됩니다.

공정능력분석: 도표 모양 옵션 설정

축 옵션 설정

이 옵션에 액세스하려면 선택 영역의 **모양** ⇨ **축** 머리글 아래에서 **축**을 클릭합니다.

다음 옵션을 설정할 수 있습니다.

- 축 색상
- 눈금 표시 색상
- 가로축 및 세로축 레이블

참조선 추가

이 옵션에 액세스하려면 선택 영역의 **모양** ⇨ **축** 머리글 아래에서 **가로** 및 **세로**를 클릭합니다.

참조선을 사용하려면 **참조선 사용**을 선택합니다. 참조선의 출력형식을 지정하고 각 참조선의 위치를 지정할 수 있습니다.

인셋 추가

이 옵션에 액세스하려면 선택 영역의 **모양** 머리글 아래에서 **인셋**을 클릭합니다.

요약통계량 상자 또는 테이블(인셋)을 그래프에 직접 추가하여 도표를 보기 좋게 만들 수 있습니다. 비교 히스토그램을 생성하면 각각의 작은 히스토그램에 인셋이 표시됩니다.

인셋 포함을 선택하여 도표에 요약통계량을 추가합니다. 인셋에 포함할 통계량을 선택합니다.

주: 공정능력지수를 계산하기 위한 통계량은 **분석변수** 역할에 할당한 변수에 대한 규격한계를 생성할 때만 사용할 수 있습니다. 히스토그램, CDF 도표, 확률 도표 및 Q-Q 도표에 대한 규격한계를 생성할 수 있습니다. 공정능력지수 계산에 대한 자세한 내용은 CAPABILITY 프로시저에 대한 도움말을 참조하십시오.

기타 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **옵션**을 클릭합니다.

사용할 수 있는 도표 모양 옵션은 작업에서 생성되는 도표의 유형에 따라 다릅니다.

역할 이름	설명
배경	도표의 배경 색상을 지정합니다.
막대 외곽선 및 막대 채우기	히스토그램의 막대 색상을 지정합니다. 주: 이 옵션은 비교 히스토그램에서만 사용할 수 있습니다.
축 바꾸기 : 백분위수를 세로축으로 사용	도표에서 축을 반대로 바꿉니다. 이 옵션을 선택하면 도표의 방향이 바뀌므로 선택한 축 옵션 및 참조선에서 지정한 내용은 적용되지 않습니다. 주: 이 옵션은 확률 도표 및 Q-Q 도표에서만 사용할 수 있습니다.
기본 범례 생략	도표에 기본 범례를 표시하지 않습니다.
중간점 사용	히스토그램에 대한 중간점을 지정합니다. 드롭다운 리스트에서 사용할 수 있는 옵션은 히스토그램을 생성하는지 또는 비교 히스토그램을 생성하는지에 따라 달라집니다. 분류변수 역할에 변수를 할당할 때에는 비교 히스토그램을 생성합니다. 비교 히스토그램에서는 다음 옵션을 사용할 수 있습니다. ■ 중간점 지정 - 시작 중간점과 마지막 중간점을 지정할 수 있습니다. 다른 중간점을 계산하는 데 사용할 단계 간격도 지정할 수 있습니다. 중간점은 각 막대로 표시되는 값 범위의 중간 값입니다. ■ 키 - 키 셀의 데이터에 대해 중간점을 결정합니다. 기본적으로 분류변수의 레벨은 첫 번째 분류변수의 내부(서식화되지 않은) 값의 증가 순으로 위에서 아래로(왼쪽에서 오른쪽으로) 표시됩니다. 분류변수를 하나만 지정하면 키 셀은 이러한 순서로 첫 번째에 나타나는 레벨이 됩니다. 분류변수를 두 개 지정하면 키 셀은 이러한 순서로 첫 번째에 나타나는 변수 1과 변수 2의 레벨 결합이 됩니다. 따라서 키 셀의 선택에 따라 모든 셀에 사용되는 균등 수평 축이 결정됩니다. 그런 다음 키 셀에 대한 중간점 리스트에 나머지

역할 이름	설명
	셀의 데이터가 포함될 때까지 리스트가 어느 방향이든지 필요한 만큼 확장됩니다. ■ 균등 - 총 표본 크기를 기반으로 중간점 개수를 결정합니다. 분류변수 역할에 변수를 할당하지 않을 때에는 히스토그램을 생성합니다. 히스토그램에서는 다음 옵션을 사용할 수 있습니다. ■ 중간점 지정 - 시작 중간점과 마지막 중간점을 지정할 수 있습니다. 다른 중간점을 계산하는 데 사용할 단계 간격도 지정할 수 있습니다. 중간점은 각 막대로 표시되는 값 범위의 중간 값입니다. ■ 그룹 수 지정 - 데이터에 대한 막대 수(히스토그램 구간이라고도 함)를 지정할 수 있습니다. 기본적으로 표시되는 구간 수는 데이터 범위에 따라 달라집니다. 그러나 표시할 최대 구간 수는 지정할 수 있습니다. 주: 히스토그램 작업을 사용하는 경우에만 중간점을 지정할 수 있습니다.

사용자 정의 그래프 크기 지정(픽셀)	이미지 크기를 지정합니다. 참조할 수 있도록 전체 화면 크기가 너비 및 높이 상자에 제공됩니다.
-----------------------------	---

주: 비교 옵션은 비교 히스토그램에서만 사용할 수 있습니다. 히스토그램 작업에서 **분류변수** 역할에 변수를 할당하여 비교 히스토그램을 생성할 수 있습니다.

역할 이름	설명
행 및 칼럼	분석 역할에 하나 이상의 분류변수를 할당하면, 페이지당 행 수 및 페이지당 칼럼 수 상자를 사용하여 비교 히스토그램에서의 칼럼과 행 배열 방식을 변경할 수 있습니다. 한 개의 분류변수를 할당하면 페이지당 한 개의 칼럼과 두 개의 행이 기본 정렬 방식이 됩니다. 두 개의 분류변수를 할당하면 페이지당 두 개의 행과 두 개의 칼럼이 기본 정렬 방식이 됩니다. 옆 프레임 및 위 프레임 드롭다운 리스트를 사용하여 행 레이블과 칼럼 레이블의 프레임 영역을 채울 색상을 지정할 수도 있습니다.
막대 수	기본적으로 표시되는 구간 수는 데이터 범위에 따라 달라집니다. 그러나 표시할 최대 구간 수는 지정할 수 있습니다. 또한 표시되는 구간 수가 데이터 평균 이상 및 이하의 표준편차 값 범위를 기반으로 하도록 표준편차를 지정할 수 있습니다.

규격 한계 옵션 설정

주: 히스토그램, 확률 도표, Q-Q 도표 및 CDF 도표에서 이러한 옵션을 사용할 수 있습니다.

선택 영역의 **모양** 머리글 아래에서 **규격한계**를 클릭하여 규격하한, 목표 값 및 규격상한을 지정합니다.

주: **규격한계** 옵션은 적어도 하나의 분석변수에 대해 하나 이상의 한계를 지정할 때만 사용할 수 있습니다. 분석변수에 대한 규격한계를 지정하려면 선택 영역에서 **데이터**를 클릭하고 분석변수를 선택합니다.

공정능력분석: 분석 유형 지정

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

공정능력분석 리포트에는 기본적으로 기본 신뢰구간 테이블, 기본 측도 테이블, 위치모수에 대한 검정 테이블, 적률 테이블 등이 포함됩니다.

리포트에 테이블을 추가하려면 **테이블** 상자에서 테이블을 선택합니다.

분석 옵션을 지정하려면 **테이블** 상자에서 테이블을 선택합니다. 옵션을 사용할 수 있으면, 옵션을 지정하는 데 사용한 오른쪽 영역에 상자가 나타납니다.

리포트에서 모든 기술통계량 테이블을 숨기려면 **기술통계량 및 공정능력지수 테이블 생략** 체크 박스를 선택합니다. 이 옵션을 선택해도 다음 테이블은 숨겨지지 않습니다.

- 선택한 분포에 의해 생성된 테이블. 해당 분포에 대한 패널에서 **분포테이블 생략** 체크 박스를 선택하여 이러한 테이블을 숨길 수 있습니다.
- INTERVALS 문에 의해 생성된 테이블. NOPRINT 옵션을 지정하면 이러한 테이블을 생략할 수 있습니다. INTERVALS 문에 대한 자세한 내용은 CAPABILITY 프로시저에 대한 도움말을 참조하십시오.

결과를 출력 데이터셋에 저장하려면 **데이터셋에 출력 통계량 저장** 체크 박스를 선택합니다. 데이터셋에는 일변량 통계량 및 공정능력지수가 포함됩니다.

주: SAS Enterprise Guide는 **옵션 창**의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

개별 측정값관리도

개별 측정값관리도 정보	231
개별 측정값관리도: 분석 역할에 변수 할당	232
관리도: 관리한계 지정	233
관리도: 수행할 검정 선택	234
관리도: 도표 모양 옵션 설정	235
관리도: 분석 결과 저장	235
제목 및 각주 지정	236
속성 보기	236

개별 측정값관리도 정보

Shewhart 관리도는 공정이 통계적 관리 상태에 있는지 여부를 결정하기 위한 그래픽 및 분석 도구입니다. 개별 측정값관리도 작업은 개별 측정값 및 이동 범위에 대한 관리도를 생성합니다. 이 관리도는 각 하위 그룹 표본에 대해 하나의 측도만 사용할 수 있고 측도가 독립적이고 정규 분포를 따르는 경우에 적합합니다.

이 작업을 사용하여 제조 공정을 분석할 수 있습니다. 예를 들어 제트 엔진을 제조하는 항공 회사가 각 엔진의 앞부분의 내부 지름(단위: 센티미터)을 측정한다고 가정합니다. 20개 엔진의 직경 측정값이 SAS 데이터셋으로 저장되었습니다. 개별 측정값관리도의 각 포인트는 특정 엔진의 안쪽 직경을 가리킵니다. 이동 범위 그래프의 각 포인트는 2개의 가장 최근의 측정값을 나타냅니다. 모든 개별 측정값 및 이동 범위가 관리한계 내에 위치하면 공정이 통계적 관리 상태에 있다고 결론을 내릴 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	SHEWHART
필수 SAS 제품	Base SAS, SAS/QC, SAS/GRAPH

요구 사항 이름	프로시저 및 제품 이름
추가 SAS 제품(권장)	SAS/STAT

개별 측정값관리도: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
공정 측정 변수	분석할 공정 변수를 지정합니다. 이 변수에는 개별 측정값이 포함됩니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
측정값 표시 변수	데이터의 하위그룹을 식별하는 변수를 지정합니다. 변수값은 입력 테이블의 관측값이 논리적 하위그룹으로 배열되는 방법을 나타냅니다. 일반적으로 값은 다음과 같습니다. ■ 하위그룹 표본의 수집 순서를 지정하는 지수 ■ 하위그룹 표본이 수집된 날짜 또는 시간 ■ 하위그룹 표본을 고유하게 식별하는 레이블 측정값 표시 변수는 필수 변수이며 이 역할에는 단 하나의 변수만 할당할 수 있습니다. 식별된 하위그룹으로 데이터를 정렬하려면 하위그룹별 정렬 체크 박스를 선택합니다.
그룹 분석 변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
블록 변수	데이터를 연속된 하위그룹 블록으로 그룹화하는 변수를 최대 두 개까지 지정합니다. 이 블록은 범례에서 레이블로 표시되며 각 블록 변수는 범례에서 한 종류의 레이블이 됩니다.

관리도: 관리한계 지정

이 옵션에 액세스하려면 선택 영역에서 **관리한계**를 클릭합니다.

옵션 이름	설명
Sigma 한계	관리도에 그려지는 요약 하위그룹 통계량의 표준오차(sigma)의 배수로 관리한계의 너비를 지정합니다. 너비는 양수여야 합니다. 기본 배수는 3입니다.
계산 방법 선택	관리한계의 계산 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 활성 데이터에서 관리한계 계산 ■ 선택한 데이터에서 관리한계 계산 관리한계 또는 관리한계를 계산할 수 있는 모수를 포함하는 데이터셋을 지정하려면 찾아보기를 클릭합니다. 사용할 데이터셋을 선택하고 열기를 클릭합니다. 주: 선택한 데이터셋은 LIMITS= 데이터셋이어야 합니다. LIMITS= 데이터셋에 필요한 변수에 대한 자세한 내용은 생성 중인 그래프의 프로시저 문서를 참조하십시오. 모든 하위 그룹에 대한 요약통계량을 도표화하려면 하위그룹 표본 크기가 명목 관리한계 표본 크기와 동일한지 관계없이 모든 하위그룹에 대한 요약통계량 도표를 생성합니다. 체크 박스를 선택합니다. 거의 모든 하위 그룹의 표본 크기가 공통이고 Nominal 표본 크기 n에 대응하는(가변이 아닌) 고정된 관리한계를 표시하려고 할 때 이 통계량을 도표화할 수 있습니다. 이 옵션을 선택하고 하위 그룹에 대한 표본 크기가 유의하게 다르면 관리한계 해석은 표본 크기가 n인 하위 그룹에 대해서만 의미가 있습니다. 표본 크기가 n과 같지 않은 데이터 포인트에 대해 특수 표식을 추가하려면 표본 크기에 해당하는 포인트에 대해 특수 표식을 추가하며, 표본 크기는 고정 관리한계의 명목표본 크기와 동일하지 않습니다. 체크 박스를 선택합니다. ■ 관리한계 지정 상한, 중심 및 하한 필드에 관리상한, 중심선 및 관리하한을 각각 지정합니다. 하위그룹의 표본 크기가 다양하면, 명목표본 크기를 지정해야 합니다. 주: 그룹 분석변수 역할에 변수를 할당하면 관리한계 값을 입력할 수 없습니다.

관리도: 수행할 검정 선택

이 옵션에 액세스하려면 선택 영역에서 **검정**을 클릭합니다.

주: 이 검정은 3 Sigma 한계가 아니거나 하위그룹 표본 크기에 따라 달라지는 관리한계에는 적용되지 않습니다.

옵션 이름	설명
검정 선택	체크 박스를 사용하여 런 검정, 패턴 검정 및 Western Electric 규칙이라고도 하는 특수 원인에 대한 검정을 하나 이상 요청합니다. 이러한 검정은 주관리도에 그려진 포인트에서 특정 비임의 패턴을 찾아냅니다. 이러한 비임의 패턴의 발생을 신호라고 합니다. 이러한 검정을 통해 특수한 변동의 원인이 존재하는지를 알 수 있습니다. 각 검정을 선택하면 해당 검정에 대한 설명이 설명 상자에 나타납니다.
레이블	텍스트 상자를 사용하여 특수 원인에 대한 검정이 양성으로 나타난 포인트에 대한 레이블을 입력합니다. 레이블의 문자 수는 16자를 초과할 수 없습니다. 각 레이블은 검정에서 신호가 발생한 모든 점에 표시됩니다.
레이블 색상	■ 드롭다운 리스트를 사용하여 지정한 레이블의 색상을 선택합니다. ■ 영역 선 표시 체크 박스를 사용하여 주관리도에 A, B, C 영역을 구분하는 선을 표시합니다. 이 영역은 관리한계 사이의 간격을 여섯 개의 동일한 간격을 지닌 하위 간격으로 분할함으로써 구성됩니다. ■ 3 sigma 한계 무시 체크 박스를 사용하여 기본 배수 3 이외의 sigma 관리한계를 지정했을 때 특수한 원인을 찾는 검정을 수행할 수 있습니다. sigma 한계를 변경하려면 선택 영역에서 관리한계를 클릭합니다. ■ 모든 점을 검정의 시작점으로 사용 체크 박스를 사용하여 중복되는 점 패턴에 특수 원인 검정을 적용합니다.

관리도: 도표 모양 옵션 설정

일부 모양 옵션은 여러 관리도 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- “축 옵션 설정” (605페이지)
- “참조선 추가” (605페이지)
- “기타 도표 옵션 설정” (606페이지)
- “블록 변수 옵션 설정” (608페이지)

관리도: 분석 결과 저장

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

분석 결과를 저장하기 위해 다양한 통계량을 포함하는 출력 데이터셋을 생성할 수 있습니다. 영구적으로 저장하려는 출력 데이터의 체크 박스를 선택합니다.

역할 이름	설명
하위그룹 통계량과 관리한계 데이터셋	관리도에 그려지는 정보를 포함하는 출력 데이터셋을 생성합니다. 여기에는 하위그룹 식별자값 및 해당 요약통계량과 관리한계가 포함됩니다. 주: 이 옵션은 OUTTABLE= 데이터셋을 생성합니다. SHEWHART 프로시저에서 TABLE= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.
하위그룹 통계량 출력 데이터셋	하위그룹 요약통계량을 포함하는 출력 데이터셋을 생성합니다. 주: 이 옵션은 OUTHISTORY= 데이터셋을 생성합니다. SHEWHART 프로시저에서 HISTORY= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.
관리한계 출력 데이터셋	관리한계를 저장하는 출력 데이터셋을 생성합니다. 추후에 이 작업 사용 시 관리한계의 계산 방법을 선택할 때 이 데이터셋을 사용할 수 있습니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식

■ 결과의 그래픽 출력에 대한 출력형식

.....
주: 기존 작업의 레이블을 변경하는 경우 속성 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.
.....

생명표 분석

생명표 분석 작업 정보	239
생명표 분석: 분석 역할에 변수 할당	240
생명표 분석: 방법 옵션 설정	241
생명표 분석: 도표 생성	242
생명표 분석: 결과 옵션 설정	243
제목 및 각주 지정	244
속성 보기	244

생명표 분석 작업 정보

생명표 분석 작업은 연구의 중단 또는 종료로 인해 오른쪽이 절단될 수 있는 데이터의 생존 분포에 대한 비모수 추정값을 계산합니다. 이러한 데이터에서는 알 수 있는 사실은 수명이 입력값을 초과한다는 것뿐입니다. 생존곡선을 비교하여 두 표본이 동일한 생존분포의 표본 인지를 확인할 수 있습니다.

이 작업은 계층에 걸친 생존함수의 동질성을 검정하기 위한 순위검정 및 우도비 검정을 계산합니다. 생존곡선을 검사할 수 있는 도표도 생성됩니다.

이 작업을 사용하여 특정 인구에 대해 생존분포를 연구할 수 있습니다. 예를 들어 발암 물질에 노출된 40마리의 쥐를 무작위로 추출하여 두 처리 집단 중 하나로 할당해서 생존시간을 연구한다고 가정합니다. 생존시간은 임의 추출부터 사망까지의 시간입니다. 관심 사건은 발암 물질에 의해 유발된 암으로 인한 사망입니다. 생명표 분석 작업을 사용하여 두 처리 집단 사이에 생존분포가 다른지를 결정할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	LIFETEST
필수 SAS 제품	Base SAS, SAS/STAT

요구 사항 이름	프로시저 및 제품 이름
추가 SAS 제품(권장)	SAS/GRAPH

생명표 분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
생존시간	생존시간 변수로 사용할 변수를 지정합니다. 생존시간은 고장시간이라고도 하며, 이벤트 시간은 비절단 생존시간을 말합니다. 생존 시간은 필수이며, 이 역할에는 변수를 하나만 할당할 수 있습니다.
절단 표시 변수	절단 표시 변수로 사용할 변수를 지정합니다. 오른쪽 절단 표시값 영역에서 생존시간의 절단 여부를 나타내는 변수값을 선택합니다. 절단 표시 변수값은 숫자여야 하며, 결측값이 아니어야 합니다. 절단 표시 변수는 필수 변수가 아닙니다. 그러나 절단 표시 변수가 사용되면 하나의 변수만 지정할 수 있으며, 반드시 절단 표시값을 지정해야 합니다.
계층변수	값이 계층 레벨을 결정하는 변수를 지정합니다. 숫자 변수에 대해 끝점을 지정하지 않으면 계층변수 개별값의 레벨 조합에 의해 계층이 결정됩니다. 체크 박스를 클릭하여 결측값이 올바른 계층 레벨을 형성하도록 할 수 있습니다. 숫자 계층변수에서는 계층 레벨 영역의 구간 지정을 선택하여 계층 구간을 지정할 수 있습니다. 그런 다음 끝점을 추가하여 구간을 정의합니다. 기본적으로 고유타값이 계층 레벨을 형성합니다. 계층을 결정하는 끝점을 지정합니다. 끝점 아래의 구간에는 끝점이 포함되지 않고, 끝점 위의 구간에는 끝점이 포함됩니다. 변수에 대한 끝점을 지정하려면 끝점 값을 입력하고 추가를 클릭합니다. 끝점은 정수 값으로 제한되지 않습니다. 쉼표나 공백으로 구분된 단일 값을 입력하거나 범위를 입력할 수 있습니다. 예를 들어 $\{-\infty, 5), [5, 10), [10, 15), [15, \infty)\}$ 구간에 속하는 값에 대해 고유 층화를 생성하려면 5 10 15를 입력합니다. 범위 5 ... 10은 다음과 같은 구간으로 값의 고유 계층을 만듭니다: $\{-\infty, 5), [5, 6), [6, 7), [7, 8), [8, 9), [9, 10), [10, \infty)\}$.

역할 이름	설명
	이 역할에 여러 변수를 지정하면 각 변수별로 구간을 정의하십시오.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수 값입니다.
공변량 검정	실패 시간 연결에 대한 검증에 사용되는 숫자(연속) 공변량의 리스트를 지정합니다. 두 개의 순위 통계량 집합이 계산됩니다. 통계량과 분산은 모든 층에 대해 풀링됩니다. 각 공변량에 대해 일변량(주변) 검정 통계량이 표시되고, 공분산의 결합 효과에 대한 검정 통계량도 표시됩니다.
ID 변수	ID 변수로 사용할 변수를 하나 이상 지정합니다. ID 변수값은 누적-한계 생존함수 추정값의 관측값 레이블을 생성하는 데 사용됩니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.

생명표 분석: 방법 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **방법**을 클릭합니다.

옵션 이름	설명
추정 방법	추정 방법을 지정하려면 누적-한계(Kaplan-Meier) 또는 생명표 분석 버튼을 클릭합니다. 두 방법 모두 비슷한 결과를 생성합니다. 그러나 Kaplan-Meier 방법이라고도 하는 제품 한계 방법은 실제 생존 횟수를 기반으로 하며, 생명표 분석 방법은 생존 횟수를 구간으로 그룹화합니다. 표본의 크기가 크면 생존시간을 구간으로 그룹화하는 생명표 분석이 유리합니다. 기본 방법은 누적-한계입니다.

옵션 이름	설명
구간	추정 방법으로 누적·한계를 선택하면, 각 구간에 대해 끝점을 정의하여 표시되는 생존함수 추정값의 개수를 생성할 수 있습니다. 작업은 각 지정된 구간 내의 최소 시간에 대한 추정값만 표시합니다. 추정 방법으로 수명 테이블을 선택하면 생존 함수 계산에 사용되는 구간 구성에 대한 정보를 지정할 수 있습니다. 기본적으로는 10개의 구간이 사용되지만, 개수를 선택하고 값을 입력하면 다른 값을 지정할 수 있습니다. 구간의 너비를 지정할 수도 있습니다. 너비를 선택하고 값을 입력합니다. 구간을 구성하는 데 사용할 정확 끝점을 정의하려면 끝점을 선택합니다.
생존함수에 대한 신뢰 수준	신뢰수준 값은 직접 입력하거나 드롭다운 리스트에서 선택할 수 있습니다. 기본값은 95% 신뢰수준입니다.

생명표 분석: 도표 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

이 옵션을 사용하면 시간에 대한 추정생존함수, 위험함수, 확률밀도함수의 도표를 그릴 수 있습니다.

옵션 이름	설명
생존함수 도표 표시	이 옵션은 해당 모집단의 수명을 설명하는 데 사용되는 생존분포함수의 도표를 생성합니다. ■ 신뢰구간을 포함하려면 각 점에 대한 신뢰구간 포함 체크 박스를 선택합니다. 도표의 신뢰한계는 신뢰구간의 방법 설정에 따라 결정됩니다. 기본값은 95% 신뢰구간입니다. ■ 범례가 있는 단일 도표에 각기 다른 계층의 도표를 모두 표시하려면 한 개의 도표에 각 계층 중첩 체크 박스를 선택합니다. 계층변수 역할에 변수를 할당하고 생존함수 도표 표시 체크 박스를 선택하면 기본적으로 각기 다른 계층의 도표가 모두 하나의 도표에 표시됩니다. 주: 한 개의 도표에 각 계층 중첩 체크 박스는 계층변수 역할에 할당된 변수가 하나일 때만 활성화됩니다. 두 개 이상의 변수를 지정하면 이 옵션을 선택할 수 없습니다.
Probit(1 — 생존함수)	이 옵션은 1에서 시간 로그값에 대한 추정생존함수를 뺀 프로비트 도표를 생성합니다. 이 도표는 로그 정규분포 평가에 유용합니다. 로그 정규 모델이 적합하면 직선에 가까운 곡선이 나타납니다.

옵션 이름	설명
-Log(생존함수)	이 옵션은 시간에 대한 추정생존함수의 음의 로그값 도표를 생성합니다. 이 도표는 지수분포 평가에 유용합니다. 지수 모델이 적합하면 곡선이 원점을 지나는 직선의 형태로 나타납니다.
Log(-log(생존함수))	이 옵션은 시간 로그값에 대한 추정생존함수의 음의 로그값 도표를 생성합니다. 이 유형의 도표는 와이블 분포 평가에 유용합니다. 와이블 모델이 적합하면 직선에 가까운 곡선이 나타납니다.
생명표 추정값	시간에 대한 위험함수 및 확률밀도함수의 도표를 생성하려면 먼저 생명표 분석 방법을 선택하여 생존함수 추정값을 계산해야 합니다. 그런 다음 위험함수 또는 확률밀도함수 체크 박스를 선택할 수 있습니다.

생명표 분석: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
출력 데이터 저장	다양한 통계량과 생존함수 추정값이 포함된 출력 데이터셋을 생성합니다. ■ 생존함수 추정값 체크 박스를 선택하면 생존함수의 추정값과 대응하는 모든 계층의 신뢰한계를 포함하는 출력 데이터셋이 생성됩니다. ■ 통계량 체크 박스를 선택하면 선택한 공변량 검정 역할 변수의 고장시간과의 관계에 대한 전체 카이제곱 검정통계량을 포함하는 출력 데이터셋이 생성됩니다. 이 데이터셋에는 일변량 순위통계량과 관련 공분산행렬 값도 포함됩니다. 주: SAS Enterprise Guide는 옵션 창의 작업> 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
모든 출력 숨기기	결과를 표시하지 않습니다. 이 체크 박스를 선택하면 도표 패널의 옵션을 사용할 수 없습니다. 결과는 기본적으로 표시됩니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

선 도표

선 도표 작업 정보	245
선 도표: 도표 유형 선택	246
선 도표: 분석 역할에 변수 할당	247
선 도표: 모양 옵션 설정	248
기타 도표 옵션 설정	248
추가 모양 옵션	249
제목 및 각주 지정	249
속성 보기	249

선 도표 작업 정보

선 도표 작업은 데이터 포인트의 추세 또는 패턴을 표현함으로써 변수 간의 수학적 관계를 보여 주는 선, 산점도, 스플라인, 바늘, 스텝, 회귀, 평활, STD, 라그랑지 보간 또는 중첩 도표를 생성합니다.

예를 들어 선 도표는 특정 제품에 대하여 광고 지출량과 판매량 사이에 관계가 있는지 나타낼 수 있습니다.

자세한 내용은 [“그래픽 출력형식 선택” \(5페이지\)](#)에서 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GPLOT
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

선 도표: 도표 유형 선택

이 옵션에 액세스하려면 선택 영역에서 **선 도표**를 클릭합니다.

역할 이름	설명
선 도표	데이터 포인트가 직선으로 연결된 도표를 생성합니다. 포인트는 입력 데이터셋에 나타난 순서대로 연결됩니다.
스플라인 도표	선을 보간하는 데 사용되는 스플라인 루틴입니다.
바늘 도표	각 데이터 포인트에서 세로축이 0 값을 갖는 수평선까지, 수평선이 0보다 크면 세로축의 최솟값까지 수직선을 그립니다. 수평선은 자동으로 표시됩니다.
스텝 도표	데이터 영역을 스텝함수로 표시합니다. 도표에서 스텝은 수직선으로 연결됩니다. 도표를 그리기 전에 데이터는 독립변수로 정렬됩니다.
회귀선을 포함한 산점도	산점도를 생성하고 데이터에 최적의 선형 회귀선을 적합시킵니다. 기본적으로, 회귀선은 도표의 원점을 지나지 않으며 신뢰한계는 표시되지 않습니다.
평활 도표	스플라인 루틴을 사용하여 매끄러운 선을 데이터에 적합시킵니다. 이 보간법은 잡음 데이터를 평활화하기 위해 사용됩니다. 도표에서 점은 반드시 선에 위치하지 않아도 됩니다. 도표를 그리기 전에 데이터는 독립변수로 정렬됩니다.
표준편차 도표	Y 변수의 평균 값과 각 X 변수값에 대한 두 표준편차를 실선으로 연결합니다.
라그랑지 보간 도표	1차항의 라그랑지 다항식을 사용하여 도표 선을 평활하게 합니다. 도표를 그리기 전에 데이터는 독립변수로 정렬됩니다.
중첩을 사용한 다중 수직 칼럼 선 도표	각각의 Y 변수 대 동일한 X 변수로 이루어진 개별 도표를 생성합니다. 모든 도표가 동일한 그래프에 나타납니다.
그룹 칼럼별 다중 선 도표	그룹 변수의 각 개별 값에 대한 별도의 도표를 생성합니다. 모든 도표가 동일한 그래프에 나타납니다.
산점도	데이터 포인트의 패턴이나 집중도를 나타냄으로써 2개 또는 3개 변수 사이의 관계를 보여 주는 산점도를 생성합니다.

선 도표: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
수평	이 역할에 할당하는 칼럼은 그래프의 가로축 또는 X축 변수입니다.
수직	이 역할에 할당하는 칼럼은 그래프의 세로축 또는 Y축 변수입니다. 가로축 변수의 고윳값에 대한 요약통계를 생성할 수 있습니다. 세로 역할에서 변수를 선택합니다. 가로축의 개별값 요약을 선택하고 요약 함수를 선택합니다. 예를 들어 입력 데이터는 각 국가의 판매 및 비용을 지역별로 나열합니다. 이때 가로 역할에 국가 변수를, 세로 역할에 판매 변수를 할당합니다. 각 국가에 대한 판매 합계를 도표로 나타내려면 체크 박스를 선택한 다음 드롭다운 리스트에서 합을 선택합니다. 이 작업은 국가의 지역별 판매를 추가하고 최종 출력에 각 국가의 단일 판매 값을 표시합니다.
세로(오른쪽)	이 역할에 할당하는 칼럼은 그래프의 오른쪽 면에 대한 세로축 또는 Y축 변수입니다. 이 변수는 X축 변수에 대응하여 그려지므로 그래프에 별도의 도표가 생성됩니다. 가로축 변수의 고윳값에 대한 요약통계를 생성할 수 있습니다. 세로 역할에서 변수를 선택합니다. 가로축의 개별값 요약을 선택하고 요약 함수를 선택합니다. 예를 들어 입력 데이터는 각 국가의 판매 및 비용을 지역별로 나열합니다. 이때 가로 역할에 국가 변수를, 세로(오른쪽) 역할에는 비용 변수를 할당합니다. 각 국가에 대한 비용 합계를 도표로 나타내려면 체크 박스를 선택한 다음 드롭다운 리스트에서 합을 선택합니다. 이 작업은 국가의 지역별 비용을 추가하고 최종 출력에 각 국가에 대해 단일 비용 값을 표시합니다. 주: 선 도표 유형으로 그룹 칼럼별 다중 선 도표를 선택하면 세로(오른쪽) 역할을 사용할 수 없습니다.
그룹	이 역할에 할당하는 칼럼 값에 따라 동일한 그래프에 표시되는 도표 선 개수가 결정됩니다. 이 칼럼의 각 고윳값에 대해 별도의 도표가 생성됩니다. 모든 선은 동일한 축을 사용합니다.

역할 이름	설명
	주: 선 도표 유형으로 그룹 칼럼별 다중 선 도표 를 선택하면 세로(오른쪽) 역할을 사용할 수 없습니다.
그룹 분석 칼럼	이 역할에 할당하는 칼럼의 값에 따라 생성되는 그래프의 수가 결정됩니다. 각 고유탄에 대해 별도의 그래프가 생성됩니다. 생성된 모든 그래프에 대해 각 축에 대한 배율을 동일하게 하려면 모든 그래프에 대해 축 길이 같음 체크 박스를 선택합니다. 기본적으로, 각 축에 대한 값의 범위는 도표화된 데이터에서 최소 및 최대값에 기반합니다. 이 옵션을 선택하면 각 축의 범위는 모든 그래프에 대해 동일합니다.

선 도표: 모양 옵션 설정

기타 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **도표**를 클릭합니다.

주: 이 옵션은 선 도표 또는 2D 산점도를 선택할 때만 사용할 수 있습니다.

특정 변수에 대한 도표 모양을 제어하는 방법

- 1 리스트에서 변수를 선택합니다. 리스트의 변수는 **세로** 및 **세로(오른쪽)** 역할에 할당된 변수입니다. **그룹 칼럼별 다중 선 도표**를 선 도표 유형으로 선택했으면 이 리스트에는 **그룹** 역할에 할당된 고유한 변수값이 포함됩니다.
- 2 **선** 영역에서 선 유형, 선 두께 및 선 색상을 지정합니다.
- 3 외곽선 색상을 선택합니다. 이 옵션은 다음 부분의 색상을 지정합니다.
 - 둘러싸인 영역의 외곽선
 - 회귀분석의 신뢰한계
 - 표준 편차 도표의 선 또는 외곽선
- 4 **데이터 포인트 표식** 영역에서 표식 유형, 사용할 기호, 각 표식의 높이 및 데이터 포인트 표식의 색상을 지정합니다.

주: 일부 기호는 모든 그래픽 출력형식에서 지원되지 않습니다. 결과 출력형식에서 지원되지 않는 기호를 선택하면 지원되는 다른 기호가 대신 표시됩니다.

- 5 결측값에서 도표 선을 끊을지를 선택합니다.

주: **결측값에서 두 번째 도표 선 분리** 체크 박스는 **세로(오른쪽)** 역할에 변수를 할당할 때만 활성화됩니다.

추가 모양 옵션

일부 옵션은 여러 그래프 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- [“보간법 선택” \(601페이지\)](#)
- [“축 옵션 설정” \(593페이지\)](#)
- [“참조선 추가” \(595페이지\)](#)
- [“범례 옵션 설정” \(596페이지\)](#)
- [“그래프 영역 사용자 정의” \(597페이지\)](#)

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

선 도표 마법사

선 도표 마법사 정보	251
선 도표 마법사: 데이터 선택	251
선 도표 마법사: 역할에 변수 할당	252
선 도표 마법사: 도표 옵션 설정	252
선 도표 마법사: 모양 옵션 설정	253
선 도표 마법사: 제목 및 각주 지정	253

선 도표 마법사 정보

선 도표 마법사를 사용하여 표식 및 선보간에 대한 옵션을 추가하여 2차원 선 도표를 생성할 수 있습니다.

선 도표 마법사: 데이터 선택

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

다음을 클릭하여 역할에 변수를 할당합니다.

선 도표 마법사: 역할에 변수 할당

- 1 데이터를 선택해야 역할에 변수를 할당할 수 있습니다.
- 2 다음과 같은 역할에 변수를 할당합니다.
 - **가로** - 이 역할에 할당하는 칼럼은 도표의 가로축 또는 X축 변수입니다.
 - **세로** - 이 역할에 할당하는 칼럼은 도표의 세로축 또는 Y축 변수입니다.
- 3 (선택 사항) **그래프 분류** 역할에 변수를 할당합니다. 이 역할에 할당하는 칼럼의 값에 따라 생성할 선 도표 개수가 결정됩니다. 각 고유티값에 대해 별도의 도표가 생성됩니다. 이 역할에 여러 변수를 할당할 수 있습니다.
- 4 **가로축의 개별값에 대하여 요약** 체크 박스를 선택하여 가로축의 고유티값에 대한 요약통계를 생성합니다(선택 사항). **함수** 드롭다운 리스트에서 요약 함수를 선택합니다.
 예를 들어 입력 데이터는 각 국가의 판매 및 비용을 지역별로 나열합니다. 이때 **가로** 역할에 국가 변수를, **세로** 역할에 판매 변수를 할당합니다. 각 국가에 대한 판매 합계를 도표로 나타내려면 체크 박스를 선택한 다음 드롭다운 리스트에서 **합**을 선택합니다. 이 작업은 국가의 지역별 판매를 추가하고 최종 출력에 각 국가의 단일 판매 값을 표시합니다.
- 5 **다음**을 클릭하여 도표 옵션을 지정합니다.

선 도표 마법사: 도표 옵션 설정

- 1 역할에 변수를 할당해야 도표 옵션을 지정할 수 있습니다.
- 2 **유형** 드롭다운 리스트에서 선 보간을 선택합니다. **자세히**를 클릭하여 선택한 보간에 대해 옵션을 지정합니다. 다음과 같은 보간을 사용할 수 있습니다.
 - **라그랑지** - 데이터 포인트를 지나는 매끄러운 곡선을 그리는 방법입니다. 이 방법은 데이터가 테이블식의 정확한 값으로 구성되었을 때 주로 사용됩니다. 지정된 차수 (1,3,5)의 다항식이 가장 근사한 2, 4, 6개의 포인트에 적용됩니다. 일반적으로 1차 미분계수는 연속적이지 않습니다. X 변수의 값이 반드시 증가하지만 않을 때 해당 모수적 방법이 사용됩니다.
 - **선** - 데이터 포인트를 직선으로 연결합니다. 포인트는 입력 데이터에 나타난 순서대로 연결됩니다.
 - **바늘** - 각 데이터 포인트에서 세로축이 0 값을 갖는 수평선까지, 수평선이 0보다 크면 세로축의 최솟값까지 수직선을 그립니다. 수평선은 자동으로 표시됩니다.
 - **회귀** - 회귀분석 도표를 생성합니다. 이 보간법은 숫자 변수를 가로 및 세로 역할에 할당할 때만 사용할 수 있습니다.
 - **산점도** - 보간을 생략합니다. 데이터 포인트 도표가 생성됩니다.

- **평활** - 스플라인 루틴을 사용하여 데이터에 평활선을 적합시킵니다. 잡음 데이터를 평활화하기 위한 방법입니다. 도표에서 점은 반드시 선에 위치하지 않아도 됩니다.
 - **스플라인** - 스플라인 루틴을 사용하여 선 도표를 그립니다.
 - **표준편차** - 평균 Y 값과 각 X에 대한 $\pm 1, 2$ 또는 3 표준편차를 실선으로 연결합니다.
 - **스텝** - 스텝함수를 사용하여 데이터를 도표로 그립니다.
- 3 선 유형, 선 두께 및 선 색상을 지정합니다.
 - 4 사용할 기호, 각 표식의 높이 및 데이터 포인트 표식의 색상을 지정합니다.
-
- 주: 일부 기호는 모든 그래픽 출력형식에서 지원되지 않습니다. 결과 출력형식에서 지원되지 않는 기호를 선택하면 지원되는 다른 기호가 대신 표시됩니다.
-
- 5 **다음**을 클릭하여 모양 옵션을 지정합니다.

선 도표 마법사: 모양 옵션 설정

- 1 도표 옵션을 지정해야 도표 모양을 설정할 수 있습니다.
- 2 도표에 참조선 및 눈금 표식을 포함할지를 지정합니다.
- 3 (선택 사항) 세로축 및 가로축에 대한 레이블을 편집합니다.
- 4 **다음**을 클릭하여 제목과 각주를 지정합니다.

선 도표 마법사: 제목 및 각주 지정

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

- 1 모양 옵션을 설정해야 제목과 주석을 지정할 수 있습니다.
- 2 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다. 제목과 각주에는 매크로 변수를 사용할 수 있습니다.
- 3 **마침**을 클릭하여 작업을 실행합니다.

선형모델

선형모델 작업 정보	255
선형모델: 분석 역할에 변수 할당	256
선형모델: 모델 생성	257
선형모델: 모델 옵션 설정	257
선형모델: 고급 옵션 설정	258
선형모델: Post Hoc 검정 선택	259
Post Hoc 검정 정보	259
최소제곱평균에 대한 평균 검정 옵션 지정	259
산술평균에 대한 평균 검정 옵션 지정	260
선형모델: 도표 생성	262
선형모델: 예측값 옵션 설정	262
제목 및 각주 지정	263
속성 보기	264

선형모델 작업 정보

선형모델 작업에서는 일반선형모델에 맞는 최소제곱법을 사용합니다. 이 작업은 일반선형 모형의 구조 내에서 데이터를 분석합니다. 이 작업은 하나 이상의 연속 종속변수를 하나 이상의 독립변수에 연결하는 모델을 처리합니다. 독립변수는 관측값을 이산 그룹으로 나누는 분류변수이거나 연속변수일 수 있습니다.

이 작업을 사용하여 다음과 같은 작업을 수행할 수 있습니다.

- 회귀분석을 수행합니다. 예를 들어 자동차가 가장 높은 주행 거리를 달성하는 속도를 결정하기 위해 다양한 속도에서의 주행 거리에 대해 테스트할 수 있습니다. 2차 모델은 실험 데이터에 적합합니다.
- 다변량 반복 측도 설계를 분석합니다. 예를 들어 두 반응 Y1과 Y2가 각 대상에 대해 세 번(처리 전, 처리 후, 후속 조치에서) 측정되었다고 가정합니다. 각 대상은 A, B 또는 대조군의 세 가지 처리 중 하나를 받습니다. 이 작업에서 다른 반응과 반복되는 요인을 식별

하여 여러 측정 시간을 확인할 수 있습니다. 반복 측정 분석에는 시간 및 처리 주효과에 대한 다변량검정과 이들 간의 상호작용이 포함됩니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GLM
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

선형모델: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
종속변수	종속(반응)변수로 사용할 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
양적변수	연속독립(설명)변수(효과라고도 함)로 사용할 변수를 지정합니다.
분류변수	이산형 독립변수 효과로 사용할 변수를 지정합니다. 이 역할에는 숫자 또는 문자 변수를 할당할 수 있으나 이러한 변수는 한정된 수의 이산 값을 가지고 있어야 합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 종속변수나 독립변수로 이미 선택한 변수로는 분석을 그룹화할 수 없습니다.

역할 이름	설명
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수 값입니다.
상대 가중값 변수	반응변수로 선택한 변수의 대응값에 대해 상대 가중값으로 사용할 변수를 하나 지정합니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.

선형모델: 모델 생성

여러 SAS 작업에 모델을 생성할 수 있습니다. 자세한 내용은 [부록 3, “모델 생성” \(609페이지\)](#)에서 참조하십시오.

선형모델: 모델 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **모델 옵션**을 클릭합니다.

옵션 이름	설명
절편에 대한 검정 표시	절편에 대해 수행한 가설 검정의 결과를 표시합니다.
표시할 제곱합	각 효과에 대한 추정 가능 함수와 관련된 제곱합을 표시합니다. 추정 가능 함수는 모델 모수의 선형 조합입니다. 제곱합은 가설 검정을 위해 계산됩니다. Type I과 Type III 제곱합이 기본값입니다. 해당 제곱합 유형 체크 박스를 사용하여 유형을 추가 또는 제거할 수 있습니다. 주: 적어도 하나의 표시할 제곱합 체크 박스를 선택해야 합니다.
모수 추정값 표시	모수 추정값을 표시합니다. 모수는 일반적으로 고정되어 있으나 알려지지 않아 모델을 사용하여 표본 데이터로부터 예측되는 모집단 수량입니다. 모수 추정값은 기본적으로 표시됩니다. 분류변수 역할에 변수가 할당될 때만 모수 추정값 표시 체크 박스의 선택을 취소할 수 있습니다. 모수 추정값의 신뢰한계 체크 박스를 사용하여 모수 추정값에 대한 신뢰 구간을 표시합니다. 신뢰구간은 미리 지정된 신뢰수준으로 추정되는 합리적인 모수값의 범위를 정하는 상한 및 하한 경계입니다.

옵션 이름	설명
	신뢰수준 값은 직접 입력하거나 드롭다운 리스트에서 선택할 수 있습니다. 기본값은 95% 신뢰수준입니다.

선형모델: 고급 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **고급 옵션**을 클릭합니다.

옵션 이름	설명
각 모델 효과에 대해 추정 가능한 함수	■ 각 모델 효과에 대한 Type I - Type IV 추정가능 함수의 행렬 계수를 표시하려면 모든 추정 가능 함수에 대한 일반 형식 표시 체크 박스를 선택합니다. ■ 특정 추정가능 함수에 대한 행렬 계수를 표시하려면 제공합 유형 옆의 체크 박스를 선택합니다. ■ 추정가능 함수를 별칭 구조로 표시하려면 별칭 구조로 추정 가능한 함수 표시 체크 박스를 선택합니다. 별칭 구조에서 각 행은 각각의 추정 가능 함수에 의해 어떤 모수 선형 결합이 추정되는지를 표시하며 동일한 정보를 가진 칼럼이 모수 추정값 테이블에 추가됩니다. 주: 절편에 대한 검정 표시 체크 박스를 선택하면 절편에 대한 추정 가능 함수를 표시할 수 있습니다.
행렬 표시	■ 확대 X'X 교차곱 행렬을 표시하려면 X'X 체크 박스를 선택합니다. ■ 확대 X'X 역행렬(또는 일반 역행렬)을 표시하려면 X'X의 역행렬을 선택합니다.
조율모수	주: 이러한 조율모수는 문제가 불량조건일 때만 유용합니다. ■ Type III 또는 IV 함수의 민감도를 변경하려면 Type III 및 IV 함수의 민감도 상자에 새로운 값을 입력합니다. 지정 값보다 작은 절대값을 가진 추정 가능 함수의 모든 요소는 0으로 설정됩니다. 기본값은 1E-08입니다. ■ 행렬을 지울 때 사용되는 민감도를 변경하려면 비정칙성 확인값 상자에 원하는 값을 입력합니다. 이 옵션은 회귀 루틴의 민감도를 계획의 선형 의존성에 맞춰 튜닝합니다. 대각피벗원소가 변수의 수정제공합에 지정한 값을 곱한 값보다 작으면, 해당 칼럼은 이전 칼럼에 대해 1차 종속인 것으로 선언되고 관련 모수는 0으로 설정됩니다. 기본값은 1E-07입니다.

옵션 이름	설명
	회귀 루틴의 민감도를 비정칙성 확인값 상자에서 지정한 값을 기준으로 튜닝할 때 사용되는 허용도를 표시하려면 허용도 표시 체크 박스를 선택합니다.

선형모델: Post Hoc 검정 선택

Post Hoc 검정 정보

이 옵션에 액세스하려면 선택 영역의 Post Hoc 검정 머리글 아래에서 **최소제곱** 또는 **산술**을 클릭합니다.

- 1 **추가**를 클릭하여 효과를 **추정할 효과** 상자에 추가합니다.
- 2 리스트의 각 효과에 대해 **평균 검정을 위한 옵션** 상자에 옵션을 지정합니다.

주: **평균 검정을 위한 옵션** 상자의 옵션은 최소제곱평균 및 산술평균의 옵션과는 다릅니다.

최소제곱평균에 대한 평균 검정 옵션 지정

다음 옵션은 **평균 검정을 위한 옵션** 상자에서 사용할 수 있습니다.

옵션 이름	설명
사용할 분류변수 효과	분석에서 사용할 분류변수 효과를 선택합니다. 기본적으로 분류변수 효과는 사용되지 않습니다.
평균 비교	분석에 사용할 비교 방법을 지정합니다. ■ 차이에 대한 P-값 표시 - 최소제곱평균의 차이에 대한 p-값을 계산합니다. 다음과 같은 옵션을 사용할 수 있습니다. □ None - p-값을 계산하지 않습니다. □ 기본값 - 수정하지 않고 쌍별 차이를 계산합니다. □ 모든 쌍별 차 - 모든 쌍별 차이를 계산합니다.

옵션 이름	설명
	□ 첫 번째 레벨을 사용하는 대조군 - 기본적으로 지정된 각 최소제곱 평균 효과의 첫 번째 레벨을 사용하여 차이를 계산합니다. □ 대조군보다 작다는 가설 검정 - 비관리 레벨이 관리 레벨보다 유의하게 작은지를 검정합니다. 비관리 레벨에서 관리를 뺀 값에 대한 관련 신뢰상한이 0보다 작거나, 마이너스 무한대로 설정된 신뢰하한을 무시하면 비관리 레벨이 관리보다 유의하게 더 작음을 선언합니다. □ 대조군보다 크다는 가설 검정 - 비관리 레벨이 관리 레벨보다 유의하게 큰지를 검정합니다. 비관리 레벨에서 관리를 뺀 것에 대한 관련 신뢰하한이 0보다 크면 비관리 레벨을 관리보다 유의하게 더 큼니다. 또한, 무한대로 설정되는 관련 신뢰상한을 무시합니다. ■ 비교를 위한 수정 방법 - 최소제곱평균의 차이에 대한 p-값 및 신뢰한계에 대해 다중비교 수정을 요청합니다.
신뢰한계	각 최소제곱평균에 대해 신뢰한계를 표시하려면 신뢰한계 표시 옵션에서 예 를 선택합니다. 차이에 대한 P-값 표시 옵션에 값을 지정하면, 평균 간 차이에 대한 신뢰한계도 함께 표시됩니다.

산술평균에 대한 평균 검정 옵션 지정

다음 옵션은 **평균 검정을 위한 옵션** 상자에서 사용할 수 있습니다.

옵션 이름	설명
사용할 분류변수 효과	분석에서 사용할 분류변수 효과를 선택합니다. 기본적으로 분류변수 효과는 사용되지 않습니다.
평균 비교	분석에 사용할 비교 방법을 지정합니다. ■ 기본값 - 평균 비교 방법이 수행되지 않습니다. ■ Bonferroni t-검정 - 모든 주효과 평균에 대한 평균 차이의 Bonferroni t 검정을 수행합니다. ■ Duncan의 다중범위검정 - 모든 주효과 평균에 대해 Duncan의 다중범위검정을 수행합니다. ■ Gabriel의 다중비교검정 - 모든 주효과 평균에 대해 Gabriel의 다중비교절차를 수행합니다. ■ 쌍별 t 검정 - 모든 주효과 평균에 대해 쌍별 t 검정을 수행합니다. 이 검정은 셀 크기가 같을 때의 Fisher의 최소 유의차 검정에 해당합니다. ■ Ryan-Einot-Gabriel-Welsh 다중범위검정 - 모든 주효과 평균에 대해 Ryan-Einot-Gabriel-Welsh 다중범위검정을 수행합니다.

옵션 이름	설명
	■ Scheffe의 다중비교절차 - 모든 주효과 평균에 대해 Scheffé의 다중비교절차를 수행합니다. ■ Sidak의 수정된 쌍별 t 검정 - 모든 주효과 평균에 대해 Sidak의 부등식에 따라 조정된 레벨에서 평균 간 차이에 대한 쌍별 t 검정을 수행합니다. ■ Student-Newman-Keuls 다중범위검정 - 모든 주효과 평균에 대해 Student-Newman-Keuls 다중범위검정을 수행합니다. ■ Tukey의 스튜던트화 범위검정 - 모든 주효과 평균에 대해 Tukey의 스튜던트화 범위검정을 수행합니다. ■ Walter-Duncan K 비율 - 모든 주효과 평균에 대해 <i>Waller-Duncan K</i> 비율 검정을 수행합니다.
오차평균제곱	다중비교에 사용할 오류 효과를 선택합니다. 오류 효과에 대한 평균 제곱 유형을 지정할 수도 있습니다.
평균 옵션	평균에 대해 다음 옵션을 지정합니다. ■ 평균 표시 - 모든 모델 변수에 대한 평균을 표시할지 종속변수에 대해서만 표시할지를 지정합니다. 기본적으로 작업에서는 모든 연속변수에 대한 평균이 표시됩니다. ■ 유의하지 않은 서브셋 조인 - 비교 방법을 통해 계산된 평균이 내림차순으로 나열되고 유의하지 않은 서브셋이 해당 평균 옆에 선으로 표시됩니다. ■ 내림차순으로 평균값 정렬하기 - 평균을 내림차순으로 정렬합니다. 내림차순으로 평균값 정렬하기 옵션에 대해 예를 선택하거나 신뢰구간 영역의 옵션에서 예를 선택하면 평균이 정렬되지 않습니다.
신뢰구간	분석에 신뢰구간을 포함할지를 지정합니다. 다음과 같은 옵션을 사용할 수 있습니다. ■ 평균에 대해 표시 - 지정한 변수의 각 레벨 평균에 대한 간격으로 비교 방법의 결과를 표시합니다. ■ 모든 쌍별 차에 대해 표시 - 평균 간 모든 쌍별 차에 대한 신뢰구간으로 비교 방법의 결과를 표시합니다.
분산의 동질성	평균 효과에 의해 정의된 그룹에 대한 분산의 동질성 검정 요청 여부를 지정합니다. 다음과 같은 옵션을 사용할 수 있습니다. ■ None - 검정이 수행되지 않습니다. 기본값입니다. ■ Bartlett - Bartlett 검정을 지정합니다. 이 검정은 정규이론 우도비 검정 수정한 것입니다. ■ Brown and Forsythe - Levene 검정의 Brown과 Forsythe 변동을 지정합니다. ■ Levene(제곱 잔차) - Levene 검정에 제곱 잔차를 사용합니다. ■ Levene(절대 잔차) - Levene 검정에 절대 잔차를 사용합니다.

옵션 이름	설명
	■ O'Brien - O'Brien의 검정을 지정합니다. 이 검정은 기본적으로 제곱잔차를 사용하여 Levene의 검정을 수정한 것입니다.

선형모델: 도표 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

기본적으로 현재 선택된 데이터에 적용 가능한 모든 도표가 출력에 포함됩니다. 그러나 **사용자 정의 도표 리스트** 옵션을 선택하여 출력에 포함할 도표를 선택할 수 있습니다. 다음 옵션 중에서 선택할 수 있습니다.

옵션 이름	설명
진단 도표	표시할 적합에 대한 회귀진단 패널을 생성합니다. 이 패널에는 잔차 대 예측값, 절대 잔차 대 예측값, 스튜던트화 잔차 대 예측값, 관측된 반응변수 대 예측값, 스튜던트화 잔차 대 지레값, Cook의 D 대 관측값, 잔차에 대한 Q-Q 도표, 잔차 히스토그램, 잔차 적합 도표에 대한 산점도가 표시됩니다.
잔차 도표	표시될 각 연속공변량에 대한 잔차의 산점도를 생성합니다.
상자 도표	일원분산분석 모델의 모델 효과에 대한 도표를 생성합니다. 이 도표를 생성하려면 모형에 분류변수가 한 개만 포함되어야 합니다. 한 개의 분류변수를 할당하지 않으면 상자 도표가 생성되지 않습니다. 상대 가중값 변수를 할당하면 상자 도표를 생성할 때 이 가중값은 무시합니다.
교호 작용 도표	이원분산분석 모델에 대한 도표를 생성합니다. 이 도표를 생성하려면 모형에 분류변수가 두 개만 포함되어야 합니다. 두 개의 분류변수를 할당하지 않으면 교호 작용 도표가 생성되지 않습니다.

선형모델: 예측값 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **예측값**을 클릭합니다.

옵션 이름	설명
예측할 데이터	출력 데이터를 예측하는 데 사용할 데이터 소스를 식별합니다. 추가 데이터 체크 박스를 선택하면 이 데이터가 포함됩니다. 추가 데이터의 위치를 지정하려면 찾아보기 를 클릭합니다.
출력 데이터 저장	예측통계량 및 진단통계량을 출력 데이터셋에 저장할지를 지정합니다. 주: SAS Enterprise Guide는 옵션 창의 작업 > 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기 를 클릭합니다.
추가적인 통계량	추가적인 통계량을 저장할지를 지정합니다. ■ 잔차 - 부적합한 관측값을 식별하는 우도잔차, 모델로 잘 설명되지 않는 관측값을 식별하는 피어슨(카이) 잔차 및 부적합한 관측값을 식별하는 이탈도 잔차를 저장합니다. ■ 예측한계 - 예측값에 대한 신뢰상한 및 하한을 저장합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

선형회귀

선형회귀 작업 정보	265
선형회귀: 분석 역할에 변수 할당	266
선형회귀: 모델 옵션 설정	266
선형회귀: 통계량 옵션 설정	268
선형회귀: 도표 생성	269
선형회귀: 예측값 옵션 설정	271
제목 및 각주 지정	272
속성 보기	272

선형회귀 작업 정보

선형회귀 분석은 최소제곱법을 사용하여 데이터에 선형 함수를 할당하려고 시도합니다. 선형회귀 작업을 사용하여 여러 종속변수 및 독립변수에 대해 선형회귀 분석을 수행할 수 있습니다.

예를 들어 회귀 분석을 사용하여 키를 알고 있는 어린이의 몸무게를 얼마나 잘 예측하는지 알아낼 수 있습니다. SAS 데이터셋에 19명의 어린이의 키와 몸무게가 포함되어 있다고 가정합니다. 몸무게를 독립변수로, 키를 종속변수로 사용하여 이 데이터에 대한 선형회귀 분석을 수행할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	REG
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

선형회귀: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
종속변수	종속(반응)변수로 사용할 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
설명변수	회귀(설명)변수로 사용할 변수를 지정합니다. 회귀변수를 지정하지 않으면, 각 종속변수는 평균만 가진 모델에 적합됩니다. 즉, 각 종속변수는 상수항만 가진 모형에 적합됩니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 반응 변수로 선택한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
상대 가중값 변수	가중 최소제곱적합에 대한 상대 가중변수의 값을 포함하는 변수를 지정합니다. 상대 가중변수의 값은 음이 아니어야 합니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.

선형회귀: 모델 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **모델**을 클릭합니다.

옵션 이름	설명
모형 선택 방법	모형 선택 방법 드롭다운 리스트에서 방법을 선택합니다. 방법을 선택하지 않으면 전체 모형이 사용됩니다. 모형 선택 방법에 따라 해당 방법에 대한 추가 옵션을 사용할 수 있습니다.

옵션 이름

설명

전체 모형 적합(선택 없음)

이 방법은 기본값이며 모델 선택 기능을 제공하지 않습니다. 여기에서 모형은 종속변수와 양적(설명)변수를 할당할 때 생성됩니다.

전진선택

전진선택법은 변수가 없는 모델에서 시작합니다. 각 설명변수에 대해 이 방법은 변수가 포함되었을 때 모델에 대한 변수의 기여도를 반영하는 F 통계량을 계산합니다. F 통계량에 대한 p -값은 모델에서 변수를 포함하도록 지정된 유의수준과 비교됩니다. 기본적으로 이 값은 **0.5**입니다. 유의수준을 변경하려면 **모형에 추가되는 기준** 텍스트 상자에 값을 입력합니다.

이 값보다 큰 유의수준을 가지는 F 통계량이 없으면 전진선택은 중단됩니다. 그렇지 않으면 전진선택법은 가장 큰 F 통계량을 가진 변수를 모델에 추가합니다. 그런 다음, 모델 외부에 남아 있는 변수에 대해 F 통계량을 다시 계산하여 평가 절차를 반복합니다. 결국, 유의한 F 통계량을 생성하는 변수가 없을 때까지 변수는 모델에 하나씩 추가됩니다. 변수가 모형에 추가되면 변수는 모델에서 계속 유지됩니다.

후진제거

후진제거법은 모든 설명변수를 포함하여 모형에 대한 F 통계량을 계산함으로써 시작됩니다. 모형에 남아 있는 모든 변수가 유의한 F 통계량을 계산할 때까지 변수가 모형에서 하나씩 삭제됩니다. 유의수준은 **모형에 남는 기준** 텍스트 상자에 지정됩니다. 기본적으로 이 값은 **0.10**입니다. 각 단계에서 모델에 대한 최소 분포를 보여 주는 변수가 삭제됩니다.

단계별 선택

단계별 선택법은 전진선택법의 변형입니다. 단계별 방법에서 모델에 이미 있는 변수가 반드시 유지될 필요가 없습니다. 전진선택법에서 변수는 모형에 하나씩 추가되고, 추가될 변수에 대한 F 통계량은 유의해야 합니다. 유의수준은 **모형에 추가되는 기준** 텍스트 상자에 지정됩니다.

그러나, 단계별 선택법은 변수가 추가된 후 이미 모델에 포함된 변수를 모두 확인하여 유의한 F 통계량을 생성하지 않은 변수를 모두 삭제합니다. 유의수준은 **모형에 남는 기준** 텍스트 상자에 지정됩니다. 이러한 확인이 끝나고 해당 변수가 삭제된 후에만 다른 변수가 모델에 추가될 수 있습니다.

단계별 프로세스는 다음 조건 중 하나에 충족되면 종료됩니다.

- 모델 외부의 변수가 유의한 F 통계량을 가지지 않고 모델의 모든 변수가 모델에 남도록 지정된 유의수준에서 유의할 때
- 모델에 추가될 변수가 방금 삭제된 변수일 때

모형 선택 방법

(계속)

최대 R-제곱 향상

최대 R-제곱 향상법은 단일 모델만을 대상으로 하지 않습니다. 대신 이 방법은 "최적의" 일변수 모형, "최적의" 이변수 모형 등을 계속해서 검색합니다. 그러나 이 방법으로 각 크기에 대해 최대 R^2 을 가진 모형을 반드시 찾을 수 있다는 보장은 없습니다. 이 방법은 최고 R^2 을 산출하는 일변수 모형을 검색하는 것으로 시작됩니다. 그런 다음, R^2 의 최대 증가량을 가져오는 또 다른 변수가 추가됩니다. 이변수 모형이 나타난 후에는 모형의 각 변수를 모형에 없는 각 변수와 비교합니다. 이 방법은 각 비교 시 변수 하나를 제거하고 다른 변수로 대체함으로써 R^2 이 증가될 수 있는지를 결정합니다. 가능한 모든 전환을 비교한 후, R^2 에서 최대 증가량을 가져오는 전환을 수행합니다. 다시 비교가 시작되고, 전환을 통해 더 이상 R^2 을 증가시킬 수 없을 때까지 과정이 계속됩니다. 그러므로 생성된 이변수 모형은 이 방법으로 찾을 수 있는 최적의 이변수 모형으로 간주됩니다. 이제 다른 변수가 모형에 추가되고 "최적의" 삼변수 모형을 찾기 위해 비교 및 전환 과정이 반복됩니다. 단계별 선택법과 최대 R^2 선택법의 차이는 최대 R^2 방법에서는 전환이 이루어지기 전에 모든 전환 대상을 평가한다는 점입니다. 단계별 선택법에서는 최적의 남은 변수가 무엇을 수행하는지를 고려하지 않은 채 최악의 변수가 제거될 수 있습니다.

옵션 이름	설명
	최소 R-제곱 향상 최소 R-제곱 향상법은 최대 R-제곱 향상법과 아주 유사하지만, 이 방법에서는 R^2의 최소 증가를 낚는 전환이 선택됩니다. 모델의 특정 개수의 변수에 대해, 최대 R-제곱 및 최소 R-제곱법은 일반적으로 동일한 최적 모델을 생성하지만, 최소 R-제곱법은 각 크기에 대해 더 많은 모델을 검토합니다. R 제곱 선택 방법은 주어진 표본에서 선형회귀를 통해 종속변수를 가장 잘 예측하는 설명변수의 서브셋을 찾습니다. 서브셋에 나타나는 설명변수의 최대 개수와 최소 개수 및 선택할 각 크기의 서브셋 개수를 지정할 수 있습니다. R-제곱법은 가능한 모든 서브셋 회귀분석을 효과적으로 수행하여, 각 서브셋 크기 내에서 R^2 크기의 내림차순으로 모형을 표시합니다. 크기가 다른 서브셋을 비교하기 위해 다른 통계량을 사용할 수 있습니다. 회귀계수 추정값을 비롯한 이러한 통계량은 표시하거나 SAS 데이터셋으로 출력을 전송할 수 있습니다. R-제곱 선택법은 적용되는 각각의 변수 개수에 대해 최대 R^2을 가지는 모형을 식별한다는 점에서 다른 선택법과 차이가 있습니다. 수정된 R-제곱 선택 수정된 R-제곱 선택법은 R-제곱 선택법과 비슷하지만, 모델 선택 기준으로 수정된 R^2 통계량이 사용되고 해당 범위의 크기 내에서 수정된 R^2 최댓값을 갖는 모델을 검색한다는 점이 다릅니다. Mallows의 C_p 선택 Mallows의 C_p 선택법은 모델 선택 기준으로 Mallows의 C_p 통계량이 사용된다는 점을 제외하면 수정된 R-제곱 선택법과 비슷합니다. 모형이 C_p의 오름차순으로 나열됩니다.
유의수준	전진선택법에서는 모델에 설명변수를 추가하기 위한 유의수준을 지정합니다. 후진제거법에서는 설명변수를 모형에 남기기 위한 유의수준을 지정합니다. 단계별 선택법에서는 두 가지 유의수준을 모두 지정합니다.
모형 적합통계량	R-제곱, 수정된 R-제곱 선택 또는 Mallows의 C_p 선택에서 체크 박스를 사용하여 결과에 포함할 모형 적합통계량을 선택합니다.
모형에 포함시킬 효과	전체 적합 모형 방법을 제외한 모든 방법에서 모형에 포함시킬 효과의 체크 박스를 선택합니다. 화살표를 사용하여 효과 순서를 변경할 수 있습니다. 효과를 제거하려면 체크 박스의 선택을 취소합니다.
절편 포함	모든 방법에 대해, 분석에서 절편 효과를 제거하려면 이 체크 박스의 선택을 취소합니다. 절편은 기본적으로 포함됩니다.

선형회귀: 통계량 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **통계량**을 클릭합니다.

옵션 이름	설명
추정값의 상세 정보	추정 옵션을 설정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 표준화된 회귀계수: 표준화된 회귀계수는 회귀변수의 표본 표준편차에 대한 종속변수의 표본 표준편차 비율로 모수 추정값을 나눈 값입니다. ■ 모델의 각 항에 대한 모수 추정값이 포함된 순차 제곱합(Type 1) 및 부분 제곱합(Type 2) ■ 추정값에 대한 상관행렬 ■ 추정값에 대한 공분산행렬 ■ 모수 추정값의 신뢰한계: 기본값은 95% 신뢰수준입니다.
상관분석	부분상관계수 및 준-부분상관계수의 제곱을 포함할지를 지정합니다. ■ 편상관계수의 제곱은 $SS/(SS+SSE)$로 계산됩니다(SSE: 오차제곱합). ■ 준-편상관계수의 제곱은 SS/SST로 계산됩니다(SST: 수정 오차제곱합). 모델에 절편을 포함하지 않도록 선택하면 분모에 비수정 총제곱합이 사용됩니다. 각 상관 유형은 Type 1 및 Type 2 제곱합을 모두 사용하여 상관계수를 계산합니다.
진단	수행할 진단 분석을 지정합니다. ■ 공선성 분석. 이 옵션은 회귀변수 간 공선성에 대한 상세 분석을 요청합니다. 여기에는 고윳값, 조건 인덱스 및 각 고윳값에 대한 추정값 분산의 분해가 포함됩니다. ■ 절편없는 공선성분석. 이 옵션은 절편이 분석에 포함되지 않는다는 점을 제외하고는 공선성분석 옵션과 동일한 분석을 요청합니다. ■ 추정값에 대한 허용값. Durbin-Watson 통계량은 오차에 1차 자기상관계수가 있는지를 표시합니다. 분산팽창값은 허용값의 상대적 개념입니다. ■ 이분산성 검정. 이 옵션은 모델의 1차 적률과 2차 적률이 올바르게 지정되었는지 여부를 검정합니다. ■ 점근 공분산 행렬. 이 옵션은 이분산성 가설하에서 추정값의 추정 점근 공분산 행렬을 표시합니다. ■ Durbin-Watson 통계량. Durbin-Watson 통계량은 오차에 1차 자기상관이 있는지 여부를 보여줍니다. (이 검정은 시계열 데이터에만 적합합니다.) 잔차의 표본 자기상관도 생성됩니다.

선형회귀: 도표 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

기본적으로 현재 선택된 데이터에 적용 가능한 모든 도표가 출력에 포함됩니다. 그러나 **사용자 정의 도표 리스트** 옵션을 선택하여 출력에 포함할 도표를 선택할 수 있습니다. 다음 옵션 중에서 선택할 수 있습니다.

도표 이름	설명
잔차의 히스토그램 도표	적합 잔차 히스토그램을 생성합니다.
잔차 대 예측값 도표	잔차 대 예측값 도표를 생성합니다. 잔차는 관측값과 예측값의 차이입니다.
스튜던트화 잔차 대 예측값 도표	스튜던트화 잔차 대 예측값 도표를 생성합니다.
관측값 대 예측값 도표	관측값 대 예측값 도표를 생성합니다.
Cook의 D 통계량 도표	Cook의 D 통계량 대 관측값 번호 도표를 생성합니다.
스튜던트화 잔차 대 지레값 도표	스튜던트화 잔차 대 지레값 도표를 생성합니다. 스튜던트화 잔차는 현재 관측값을 제외한 잔차를 표준오차로 나눈 값입니다. 관측값에 대한 지레값은 H 행렬에서 대응되는 대각원소의 값입니다.
잔차의 정규 분위수 도표	Q-Q 도표를 생성합니다.
잔차 적합 도표	중심 적합 분위수 및 잔차의 나란한 도표를 생성합니다.
잔차의 상자 도표	잔차의 상자 도표를 생성합니다.
진단 도표	적합 진단 도표를 생성합니다.
DFFITS 도표	DFFITS(표준화된 적합 차이) 도표를 생성합니다. DFFITS 통계량은 예측값에 대한 관측값의 표준 영향력입니다. DFFITS 통계량은 예측값에 대한 관측값의 표준 영향도입니다. 값이 크면 관측값이 매우 영향력이 있음을 나타냅니다.
DFBETAS 도표	모수 추정값에서 관측값을 제외한 표준화된 차이를 도표로 생성합니다. 개별 관측값이 적합 모델의 각 추정 모수에 미치는 영향을 평가하는 데 유용합니다.
잔차 도표	잔차 대 회귀변수 도표를 생성합니다.
회귀선을 포함한 산점도	데이터의 산점도에 중첩된 회귀선 도표를 생성합니다.

선형회귀: 예측값 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **예측값**을 클릭합니다.

옵션 이름	설명
예측할 데이터	출력 데이터를 예측하는 데 사용할 데이터 소스를 식별합니다. 추가 데이터 체크 박스를 선택하면 이 데이터가 포함됩니다.
추가적인 통계량	추가로 포함할 통계량을 지정합니다. ■ 잔차를 선택하면 일반잔차(ordinary residuals), 표준화잔차 및 스튜던트화 잔차가 출력에 포함됩니다. 일반잔차(ordinary residuals)는 관측값과 예측값의 차이입니다. 표준화잔차는 잔차를 표준오차로 나눈 값입니다. 스튜던트화 잔차는 현재 관측값을 제외한 잔차를 표준오차로 나눈 값입니다. ■ 예측한계를 선택하면 각 예측값에 대한 $100 \times (1 - \alpha)\%$ 신뢰구간의 상한 및 하한값과 종속변수의 예측값 평균이 출력에 포함됩니다.
출력 데이터 저장	예측통계량 및 진단통계량을 출력 데이터셋에 영구히 저장할지를 지정합니다. 진단 통계량은 다음과 같습니다. ■ Cook의 D 영향력 통계량 ■ 예측값에 대한 관측값의 표준 영향력(DFFITS 통계량) ■ 지레값 ■ 개별 예측값의 표준오차 ■ 평균 예측값의 표준오차 ■ 잔차의 표준오차 주: SAS Enterprise Guide는 옵션 창의 작업 > 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
출력 및 도표 표시	결과를 표시합니다. 결과를 출력하지 않으려면 체크 박스의 선택을 취소합니다. 이 체크 박스를 선택 취소하면 도표 패널의 옵션을 사용할 수 없습니다. 결과는 기본적으로 표시됩니다. 예측값을 결과에 포함하려면 예측값 표시 체크 박스를 선택합니다. 이 옵션은 예측할 데이터 영역의 데이터 소스 및 출력 및 도표 표시 체크 박스를 선택할 때만 사용할 수 있습니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

데이터 리스트

데이터 리스트 작업 정보	273
데이터 리스트: 분석 역할에 변수 할당	274
데이터 리스트: 리스트 옵션 설정	274
제목 및 각주 지정	276
속성 보기	277

데이터 리스트 작업 정보

데이터 리스트 작업은 전체 또는 일부 변수를 사용하여 SAS 데이터셋의 관측값을 출력합니다. 단순한 리스트에서부터, 데이터를 그룹화하고 숫자 변수의 합 및 부분합을 계산하는 고급 사용자 정의 리포트에 이르기까지 다양한 리포트를 생성할 수 있습니다.

예를 들어 데이터 리스트 작업을 사용하여 각 판매 지역에 대한 비용과 수익을 합하는 리포트를 생성하고 회사에 대한 총 비용 및 수익과 이러한 값을 비교하며, 각 하위그룹과 전체 리포트에 관측값 수를 표시하고 지역 이름이 포함된 사용자 정의 제목을 지정할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	PRINT
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	없음

데이터 리스트: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
변수 리스트	지정된 변수를 지정된 순서로 출력합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
그룹 분석변수	테이블 정렬 방법을 지정합니다. 변수 또는 변수 조합에서 각 개별값(또는 BY 그룹)에 대한 리스트가 생성됩니다.
페이지 분리 변수	지정한 변수값이 바뀌거나 다음 BY 그룹이 시작되면 새로운 페이지로 인쇄합니다. 이 역할에 지정되는 변수는 그룹 분석변수 역할의 변수여야 합니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.
합 계산 변수	선택한 변수의 합계를 리스트 리포트 하단에 출력합니다.
부분합 계산 변수	지정한 변수값이 변하거나 다음 BY 그룹이 시작되면 부분합을 출력합니다. 이 역할에 지정되는 변수는 그룹 분석변수 역할의 변수여야 합니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.
ID 레이블	관측값 번호를 사용하는 대신 행을 식별하는 방법을 지정합니다.

데이터 리스트: 리스트 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

옵션 이름	설명
나열할 행	나열할 행을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 모든 행 - 모든 행을 출력합니다.

옵션 이름	설명
	■ 1행부터 n번째 행 - n 지정 텍스트 상자에서 지정한 개수의 행을 출력합니다. 예를 들어 n 값으로 50을 지정하면 입력 데이터셋의 1행부터 50행까지가 출력 데이터에 포함됩니다. ■ 1행부터 전체의 n퍼센트 행 - n 지정 텍스트 상자에서 지정한 퍼센트의 행을 출력합니다. 예를 들어 입력 데이터셋에 100행이 있고 n을 25%로 설정하면 입력 데이터셋의 1행부터 25행까지가 출력 데이터에 포함됩니다. ■ n의 배수 행 - 데이터셋에서 n번째의 모든 행을 출력합니다. n 지정 텍스트 상자에 n 값을 지정합니다. 예를 들어 n 값으로 10을 지정하면 입력 데이터셋에서 모든 10번째 행이 출력 데이터에 포함됩니다.
행 번호 인쇄	각 관측값에 대한 행 번호가 표시되는 칼럼을 출력에 포함합니다. 칼럼 머리글 텍스트 상자에서 이 칼럼에 대한 레이블을 지정할 수 있습니다. 이 칼럼의 기본 이름은 행 번호입니다.
변수 레이블을 칼럼 머리글로 사용	변수 이름이 아닌 변수 레이블을 칼럼 머리글로 사용합니다.
행 개수 인쇄	결과의 마지막에 테이블 행 수를 표시하거나, 각 BY 그룹 결과의 마지막에 각 BY 그룹의 행 수를 표시합니다.
합을 계산하기 전에 반올림	각 숫자 값을 해당 출력형식의 자릿수로 반올림하고 지정된 출력 형식이 없으면 소수점 이하 두 자리로 반올림합니다. 이 옵션을 지정하면 데이터 리스트 작업은 변수의 합을 계산하기 전에 반올림을 수행하게 됩니다.
섹션으로 페이지 나누기	페이지에 표시할 수 있는 것보다 칼럼이 많으면 각 페이지를 섹션으로 분리합니다. 페이지를 섹션으로 나누기 옵션을 선택 취소하면 데이터 리스트 작업에서 각 페이지에 가능한 한 많은 관측값을 넣습니다. 페이지에서 적합할 것보다 더 많은 칼럼이 있는 경우, 적합하지 않은 칼럼은 다음 페이지에 배치됩니다. 페이지에 표시할 수 있는 것보다 칼럼이 많으면 해당 페이지에 표시되지 못하는 칼럼은 다음 페이지에 출력됩니다. 주: 테이블의 행 및 칼럼 수에 따라 출력 페이지 수는 줄어들 수 있습니다. 이 옵션은 SAS 리포트, HTML, RTF 또는 PDF 출력형식의 결과에 영향을 주지 않습니다.
머리글 방향	칼럼 머리글을 수평 또는 수직으로 출력합니다. 기본을 선택하여 SAS가 각 칼럼에 대해 최적의 배열을 결정하도록 합니다.
칼럼 너비	데이터 리스트 작업의 칼럼 너비 결정 방식을 지정합니다. ■ 기본값: 데이터 리스트 작업은 페이지당 기준으로 칼럼 너비를 결정합니다. ■ 전체: 데이터 리스트 작업은 모든 페이지에 대해 출력형식 너비(또는 출력형식이 지정되지 않으면 기본 너비)를 사용합니다.

옵션 이름	설명
	■ 최소: 데이터 리스트 작업은 페이지당 기준으로 가능한 가장 작은 칼럼 너비를 사용합니다. ■ 균등하게: 전체 테이블을 읽어서 출력을 생성하기 전에 적절한 칼럼 너비를 결정합니다. 이 옵션을 선택하지 않으면 동일한 칼럼의 너비가 페이지마다 다르게 출력될 수 있습니다. ■ 그룹별로 균등하게: 데이터 리스트 작업은 각 BY 그룹 내에서 모든 페이지를 균등하게 인쇄합니다. 주: 테이블의 행 및 칼럼 수에 따라 출력 페이지 수는 줄어들 수 있습니다. 이 옵션은 SAS 리포트, HTML, RTF 또는 PDF 출력형식의 결과에 영향을 주지 않습니다.
레이블 분리 문자	변수 레이블에 분리 문자(*, !, @, #, \$, %, ^, &, +) 중 하나가 포함되면 레이블이 분리되도록 지정합니다. 예를 들어 "이것은 레이블입니다."라는 변수 레이블이고 * 문자가 분할 문자로 선택되어 있으면 칼럼 머리글은 다음과 같습니다. This is a label 변수 레이블을 칼럼 머리글로 사용 및 레이블 분리 문자를 모두 선택할 필요는 없습니다. 레이블 분리 문자 옵션에는 레이블 사용이 함축되어 있습니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

리스트 리포트 마법사

리스트 리포트 마법사 정보	279
리스트 리포트 마법사: 데이터 선택	280
리스트 리포트 마법사: 리스트 정의	280
리스트 리포트 마법사: 합계 표시	280
리스트 리포트 마법사: 제목 및 각주 지정	281

리스트 리포트 마법사 정보

리스트 리포트 마법사를 사용하면 상세 또는 요약 리포트를 생성할 수 있습니다. 상세 리포트에는 리포트에 대해 선택한 각 관측값당 한 행씩 포함됩니다. 요약 리포트에서는 데이터를 통합하므로 각 행이 여러 관측값을 나타냅니다. 상세 및 요약 리포트 모두에는 행 집합 또는 모든 행에 대한 숫자 데이터를 요약하는 총합이 포함될 수 있습니다. 또한 변수의 고윳값을 기준으로 별도의 테이블이나 페이지 나누기를 생성할 수 있습니다.

예를 들어 **리스트 리포트** 마법사를 사용하여 각 지역의 식료품점에서의 소다 및 커피 판매를 나열하는 리포트를 생성할 수 있습니다. 각 영역의 관리자가 별도의 테이블로 그룹화됩니다. 각 지역의 관리자를 별도의 테이블로 그룹화하고 각 테이블 내에서 각 관리자에 대해 개별 제품 판매량과 총 판매량을 나열합니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	REPORT
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	없음

리스트 리포트 마법사: 데이터 선택

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

다음을 클릭하여 리스트를 정의합니다.

리스트 리포트 마법사: 리스트 정의

데이터 소스에 20개 미만의 칼럼이 포함되어 있으면 **리스트 리포트** 마법사에 데이터 소스의 칼럼이 모두 포함됩니다. 이러한 칼럼은 **미리 보기** 영역에 표시됩니다. 데이터에 20개가 넘는 칼럼이 포함되어 있으면 리스트 리포트에 칼럼이 포함되지 않으며 **미리 보기** 영역이 비어 있습니다.

이때 리포트에 추가할 칼럼을 지정해야 합니다. 데이터 소스의 모든 칼럼, 특정 칼럼 또는 빈도를 추가할 수 있습니다.

다음을 클릭하여 총합을 표시합니다.

리스트 리포트 마법사: 합계 표시

분류변수가 아닌 모든 숫자 변수나 전체 변수에 대한 부분합 또는 총합을 표시할 수 있습니다.

총합을 지정하는 방법:

- 1 리스트를 정의해야 총합을 지정할 수 있습니다.
- 2 **총합 선택** 영역에서 총합을 계산할 변수에 대한 체크 박스를 선택합니다. 모든 변수에 대한 총합을 표시하려면 **모두 선택** 체크 박스를 선택합니다.
- 3 총합 유형, 총합 값의 위치 및 총합이 포함된 행의 레이블을 지정하려면 **편집**을 클릭합니다. **총합 편집** 대화 상자가 나타납니다.

다음 옵션 중에서 선택할 수 있습니다.

- **총합 유형** - 테이블에 표시할 총합 유형을 지정합니다. 총합 및 부분합을 표시할 수 있습니다. 표시할 총합 유형을 하나 이상 선택해야 합니다. 테이블에서 모든 총합을 표시하려면 **모두 선택** 체크 박스를 선택합니다.

- **위치** - 테이블에서 총합 행을 표시할 위치를 지정합니다. 총합을 계산하는 데 사용된 행 또는 관측값의 앞이나 뒤에 총합을 표시할 수 있습니다.
 - **총합 레이블** - 총합 행에 사용할 레이블을 지정합니다. 기본적으로 이 레이블은 총합입니다.
 - **부분합 레이블** - 부분합 행에 사용할 레이블을 지정합니다. 기본적으로 이 레이블은 부분합입니다. 또한 레이블을 이 행의 텍스트 앞에 표시할지 뒤에 표시할지를 지정할 수 있습니다. 기본적으로 레이블과 그 뒤나 앞에 오는 텍스트 사이에 공백이 삽입되지 않습니다. 공백을 포함하려면 **부분합 레이블** 상자에서 레이블의 앞이나 뒤에 공백을 추가해야 합니다.
- 4 **다음**을 클릭하여 제목 및 각주를 지정합니다.

리스트 리포트 마법사: 제목 및 각주 지정

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

- 1 제목 및 각주를 지정하려면 먼저 모든 숫자 변수에 대한 총합을 표시해야 합니다.
- 2 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다. 제목과 각주에는 매크로 변수를 사용할 수 있습니다.
- 3 **마침**을 클릭하여 작업을 실행합니다.

로지스틱 회귀

로지스틱 회귀 작업 정보	283
로지스틱 회귀: 분석 역할에 변수 할당	284
로지스틱 회귀: 모델 생성	285
반응 지정	285
모델 생성	286
모형 선택 방법 지정	286
로지스틱 회귀: 모델 옵션 설정	288
로지스틱 회귀: 도표 생성	289
로지스틱 회귀: 예측값 옵션 설정	290
제목 및 각주 지정	291
속성 보기	292

로지스틱 회귀 작업 정보

다양한 연구 분야에서 성공과 실패와 같은 이진 반응 및 보통, 미약, 심각 등과 같은 순차적 반응이 발생합니다. 로지스틱 회귀분석은 이러한 이산 반응과 설명변수 집합 간의 관계를 연구합니다.

로지스틱 회귀를 사용하여 임상 연구의 효과를 결정할 수 있습니다. 예를 들어 로지스틱 회귀 작업을 사용하여 신경통이 있는 노인 환자에 대한 진통 치료 효과에 대한 연구의 결과를 분석한다고 가정합니다. 두 가지 테스트 치료와 하나의 위약을 비교합니다. 응답 변수는 환자가 고통을 호소할 것인지에 대한 여부입니다. 연구자는 환자의 나이와 성, 치료가 시작되기 전의 통증 기간을 기록합니다. 로지스틱 회귀 작업에 대한 출력에서 두 가지 테스트 치료의 통증 감소 효과가 다른지를 확인할 수 있습니다.

주: 로지스틱 회귀 작업을 열면 **SAS Enterprise Guide**가 Binary 값을 가지는 변수를 지정할 때 약간의 지연이 있음을 알 수 있습니다. 이 분석이 완료되면, 작업에 **종속변수** 역할에 지정된 변수에 관한 정보가 나타납니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	LOGISTIC
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

로지스틱 회귀: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
종속변수	종속(반응)변수로 사용할 변수를 지정합니다. 이 역할에는 하나의 변수만 지정해야 합니다.
양적변수	독립(설명)변수로 사용할 변수를 지정합니다. 양적변수의 변경 단위를 지정할 수도 있습니다. 예를 들어 결과에서 추정 사용자 정의 오즈비를 생성하려면 단위 상자에 -2 를 입력합니다.
분류변수	분석에서 사용할 분류변수를 지정합니다. 문자 또는 숫자 변수가 될 수 있습니다. 분류변수가 숫자이면 값은 연속형이 아닌 이산형이어야 합니다. 각 변수의 모수화 방법을 선택하려면 분류변수를 선택합니다. 변수 이름의 코딩 유형 영역에서 변수의 모수화 방법을 효과 코딩(효과)으로 할지 참조 셀 코딩(참조)으로 할지를 선택합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 반응 변수로 선택한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
상대 가중값 변수	반응변수로 선택한 변수의 대응값에 대해 상대 가중값으로 사용할 변수를 하나 지정합니다. 상대 가중값이 음수이거나 결측값이면 분석에서 제외됩니다.

역할 이름	설명
	상대 가중값 변수를 정규화하려면 정규화 체크 박스를 선택합니다.

로지스틱 회귀: 모델 생성

반응 지정

이 옵션에 액세스하려면 선택 영역의 모델 머리글 아래에서 **반응**을 클릭합니다.

역할 이름	설명
반응 유형	사용할 수 있는 반응 유형은 반응변수의 레벨 수에 따라 결정됩니다. ■ 반응변수에 레벨이 두 개만 있으면 반응 유형은 이진(binary)이 됩니다. ■ 반응변수에 레벨이 두 개 이상 있으면 반응 유형을 순서가 있음 또는 순서가 없음 중에서 선택할 수 있습니다.
모델 유형	다음과 같은 모델 유형 중에서 선택할 수 있습니다. ■ 로짓 - 로그 오즈 함수의 결과를 표시합니다. 두 가지 반응 범주가 있으면 이진 로짓 모델이 사용됩니다. 셋 이상의 반응 범주가 있으면 누적 로짓 모형이 사용됩니다. ■ 프로빗 - 역표준정규확률적분함수의 결과를 표시합니다. 두 가지 반응 범주가 있으면 이진 프로빗 모델이 사용됩니다. 셋 이상의 반응 범주가 있으면 누적 프로빗 모형이 사용됩니다. ■ 보 로그-로그 - 보 로그-로그 함수의 결과를 표시합니다. 두 가지 반응 범주가 있으면 이진 보 로그-로그 모델이 사용됩니다. 셋 이상의 반응 범주가 있으면 누적 보 로그-로그 모형이 사용됩니다. ■ 로그-로그 - 로그-로그 함수의 결과를 표시합니다. 두 가지 반응 범주가 있으면 이진 로그-로그 모델이 사용됩니다. 셋 이상의 반응 범주가 있으면 누적 로그-로그 모형이 사용됩니다. ■ Glogit - 일반화된 로짓 함수의 출력을 표시합니다. 일반화된 로짓 모델에서 각 비참조 범주는 참조 범주와 대조를 이룹니다. 주: 이 옵션은 반응 유형으로 순서가 없음을 지정했을 때에만 사용할 수 있습니다.
반응 레벨	반응 레벨 상자에는 반응변수에 대한 출력형식 레벨이 오름차순으로 나열되어 있습니다.

역할 이름	설명
다음 레벨에 모형 적합	레벨에 모델을 적합시키는 방법을 지정할 수 있습니다. ■ 이진을 반응 유형으로 선택했으면 다음 레벨에 모형 적합 드롭다운 리스트에서 이진 반응 모형의 사건 범주를 선택합니다. ■ 순서가 있음을 반응 유형으로 선택했으면 다음 레벨에 모형 적합 드롭다운 리스트의 값을 사용하여 반응 범주의 순서를 지정할 수 있습니다. 정렬된 값 중 첫 번째 값을 선택하면 반응 범주가 반응 레벨 상자에 표시되는 순서로 정렬됩니다. 정렬된 값 중 마지막 값을 선택하면 반응 범주가 역순으로 정렬됩니다.
참조 레벨	이 옵션은 순서가 없음을 반응 유형으로 선택할 때만 사용할 수 있습니다. 일반화된 로짓 모델에서는 참조 레벨 드롭다운 리스트에서 참조 범주를 지정할 수 있습니다. 일반화된 로짓 모형에서는 각 로짓에서 비참조 범주와 참조 범주의 차이점이 표시됩니다.

모델 생성

여러 SAS 작업에 모델을 생성할 수 있습니다. 자세한 내용은 [부록 3, “모델 생성” \(609페이지\)](#)에서 참조하십시오.

모형 선택 방법 지정

이 옵션에 액세스하려면 선택 영역의 모델 머리글 아래에서 **선택**을 클릭합니다.

옵션 이름	설명
모형 선택 방법	모형 선택 방법 드롭다운 리스트에서 방법을 선택합니다. 방법을 선택하지 않으면 전체 모형이 사용됩니다. 모형 선택 방법에 따라 해당 방법에 대한 추가 옵션을 사용할 수 있습니다. 사용할 수 있는 방법은 다음과 같습니다. ■ 전체 모형 적합(선택 없음) - 기본값이며 효과 선택을 제공하지 않습니다. 여기에서 모형은 모형 적합에 사용되는 종속변수와 양적변수를 할당할 때 생성됩니다. ■ 전진선택 - 효과가 없는 모델에서 시작합니다. 각 설명변수에 대해 이 방법은 변수가 포함되었을 때 모델에 대한 변수의 기여도를 반영하는 F 통계량을 계산합니다. F 통계량에 대한 p-값은 모델에서 변수를 포함하도록 지정된 유의수준과 비교됩니다. 기본적으로 이 값은 0.05입니다. 유의수준을 변경하려면 모형에 추가되는 기준 텍스트 상자에 값을 입력합니다.

옵션 이름	설명
	이 값보다 큰 유의수준을 가지는 F 통계량이 없으면 전진선택법은 중단됩니다. 그렇지 않으면 전진선택법은 가장 큰 F 통계량을 가진 변수를 모델에 추가합니다. 그런 다음, 모델 외부에 남아 있는 변수에 대해 F 통계량을 다시 계산하여 평가 절차를 반복합니다. 결국, 유의한 F 통계량을 생성하는 변수가 없을 때까지 변수는 모델에 하나씩 추가됩니다. 변수가 모형에 추가되면 변수는 모델에서 계속 유지됩니다. ■ 후진제거 - 모든 설명변수를 포함하여 모델에 대한 F 통계량을 계산함으로써 시작됩니다. 모형에 남아 있는 모든 변수가 유의한 F 통계량을 계산할 때까지 변수가 모델에서 하나씩 삭제됩니다. 유의수준은 모형에 남는 기준 텍스트 상자에 지정됩니다. 기본적으로 이 값은 0.05입니다. 각 단계에서 모델에 대한 최소 분포를 보여 주는 변수가 삭제됩니다. ■ 단계별 선택 - 전진선택법의 변형입니다. 단계별 방법에서 모델에 이미 있는 변수가 반드시 유지될 필요가 없습니다. 전진선택법은 F 통계량이 유의하면 모델에 변수를 하나씩 추가합니다. 유의수준은 모델에 추가되는 기준 텍스트 상자에 지정됩니다. 그러나, 단계별 선택법은 변수가 추가된 후 이미 모델에 포함된 변수를 모두 확인하여 유의한 F 통계량을 생성하지 않은 변수를 모두 삭제합니다. 유의수준은 모형에 남는 기준 텍스트 상자에 지정됩니다. 이러한 확인이 끝나고 해당 변수가 삭제된 후에만 다른 변수가 모델에 추가될 수 있습니다. 단계별 프로세스는 다음 조건 중 하나에 충족되면 종료됩니다. □ 모델 외부의 변수가 유의한 F 통계량을 가지지 않고 모델의 모든 변수가 모델에 남도록 지정된 유의수준에서 유의할 때 □ 모델에 추가될 변수가 방금 삭제된 변수일 때 ■ 최적 서브셋 - 알고리즘을 사용하여 가능한 모든 모델 크기에 대해 1,2,3 ...개의 효과 모델에서 모든 양적변수가 포함된 단일 모델에 이르기까지 최고 카이제곱 통계량을 가진 지정된 개수의 모델을 찾습니다.
유의수준	■ 전진선택법에서는 모델에 양적(설명)변수를 추가하기 위한 유의수준을 지정합니다. ■ 후진제거법에서는 양적변수를 모델에 남기기 위한 유의수준을 지정합니다. ■ 단계별 선택법에서는 두 가지 유의수준을 모두 지정합니다.
모형에 포함시킬 효과	모형에 추가할 효과를 선택합니다. 변수가 모형에 포함시킬 효과 상자에 추가됩니다. 변수를 선택하고 화살표 버튼을 클릭하여 이 변수의 순서를 재지정할 수 있습니다. 주: 이 옵션은 모형 선택 방법으로 전체 모형 적합(선택 없음)을 선택하면 사용할 수 없습니다.

로지스틱 회귀: 모델 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

옵션 이름	설명
추정값의 상세 정보	모수 추정값의 상관행렬을 표시할지, 모수 추정값의 공분산 행렬을 표시할지 또는 두 가지 모두를 표시할지를 지정합니다.
모델 적합 평가	다음과 같은 모델 적합통계량을 지정합니다. ■ 영향력 통계량 - 이진 반응 모델에서는 영향력 있는 관측값을 식별하기 위한 진단 측정치를 표시합니다. ■ Hosmer와 Lemeshow 적합도 검정 - 이진 반응 모델에서는 Hosmer와 Lemeshow 적합도 검정을 표시합니다. ■ 이탈도 및 피어슨 적합도 검정 - 피어슨 카이제곱 통계량, 이탈도, 자유도, 각 통계량을 해당 자유도로 나눈 값 및 해당 p-값을 표시합니다. ■ 일반화 R-제곱 - 적합 모델에 대한 일반화 결정계수를 표시합니다. 주: 영향력 통계량, Hosmer와 Lemeshow 적합도 검정 및 이탈도 및 피어슨 적합도 검정 옵션은 종속변수 역할에 이진 변수를 할당할 때만 사용할 수 있습니다.
분류표 표시	예측 이벤트 확률이 범위 (0,1)의 절단점 값 이상인지, 이하인지에 따라 입력 이진 반응 관측값을 분류합니다. 예측 이벤트 확률이 절단점 값을 초과하면 관측값은 이벤트로 예측됩니다. 기본 테이블은 가장 작은 추정 확률(가장 가까운 0.02로 반올림)부터 가장 높은 추정 확률(가장 가까운 0.02로 반올림)까지 0.02단위로 확률 범위에 대한 분류를 표시합니다. 기본 리스트에 없는 절단점을 사용하려면 임계확률값(절단점) 텍스트 상자에 하나 이상의 값을 입력합니다. 각 값은 0에서 1 사이여야 합니다. 주: 이 옵션은 종속변수 역할에 할당한 변수가 이진 변수일 때만 사용할 수 있습니다.
신뢰한계	모수 영역에서 모수에 대한 신뢰구간을 지정하고 조건부 오즈비 영역에서 오즈비를 지정합니다. ■ 왈드 - 개별 왈드 검정을 기반으로 신뢰구간을 계산합니다. ■ 프로파일 우도 - 프로파일 우도를 기반으로 신뢰구간을 계산합니다. 이 구간은 일반화 우도비 검정에 대한 근사 카이제곱 분포에서 파생되어 구성됩니다.

옵션 이름	설명
	신뢰수준 드롭다운 리스트에서 값을 선택하여 회귀모수 또는 오즈비에 대한 유의수준을 설정할 수도 있습니다.
모델 적합 방법	테이블 적합 방법을 지정합니다. 기본적으로 작업은 모델에 대한 최량 적합 방법을 자동으로 결정합니다. 적합 기법 드롭다운 리스트에서 회귀모수 추정을 위한 Fisher 스코어링 또는 Newton-Raphson 최적화 기법을 선택할 수 있습니다. 주: 반응 유형으로 순서가 없음을 선택했으면 Fisher 스코어링 최적화 기법을 사용할 수 없습니다. Firth의 별점가능도 체크박스를 선택하여 모수 추정값의 편향을 줄이는 Firth의 별점최대우도 추정을 수행합니다. 이 방법은 회귀사건에서 자주 발생하는 분리성에 유용하고 정확한 로지스틱 회귀의 대안입니다. 노트 주: Firth의 별점가능도 체크 박스는 종속변수 역할에 이진 변수를 할당할 때만 사용할 수 있습니다.

로지스틱 회귀: 도표 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

기본적으로 현재 선택된 데이터에 적용 가능한 모든 도표가 출력에 포함됩니다. 그러나 **사용자 정의 도표 리스트** 옵션을 선택하여 출력에 포함할 도표를 선택할 수 있습니다. 다음 옵션 중에서 선택할 수 있습니다.

도표 이름	설명
오즈비 도표	오즈비에 대한 프로파일 우도 신뢰한계와 오즈비에 대한 왈드의 신뢰한계를 도표로 생성합니다. 이 옵션은 모델 > 옵션 패널에서 조건부 오즈비 옵션을 선택할 때만 사용할 수 있습니다.
ROC 도표	ROC(Receiver Operating Characteristic) 도표를 생성합니다. ROC 곡선 아래 영역은 "예측 확률과 관측된 반응변수의 결합" 테이블의 통계량 <i>c</i> 로 결정됩니다.
영향력 도표	다음 도표를 표시합니다. ■ 모델로 제대로 설명되지 않은 관측값을 식별하는 피어슨(카이제곱) 잔차 ■ 적합이 잘 되지 않은 관측값 지레값을 식별하는 이탈도 ■ 지레값 ■ 개별 관측값이 회귀분석 추정값에 미치는 영향력의 측도를 제공하는 신뢰구간 이동 진단

도표 이름	설명
	■ 개별 관측값을 제외한 전체 회귀분석 추정값 변경 사항의 측도를 제공하는 신뢰구간 이동 진단 ■ 개별 관측값 삭제로 인한 카이제곱 적합도 통계량의 변경 사항 ■ 개별 관측값 삭제로 인한 이탈도의 변경 사항
DFBETAS 도표	모수 추정값에서 관측값을 제외한 표준화된 차이에 대한 도표를 표시합니다. 개별 관측값이 적합 모델의 각 추정 모수에 미치는 영향을 평가하는 데 유용합니다.
Phat 도표	다음 도표를 표시합니다. ■ 개별 관측값 삭제로 인한 카이제곱 적합도 통계량의 변경 사항 ■ 개별 관측값 삭제로 인한 이탈도의 변경 사항 ■ 신뢰구간 이동 ■ 지레값 대 예측 사건 확률
지레값 도표	다음 도표를 표시합니다. ■ 개별 관측값 삭제로 인한 카이제곱 적합도 통계량의 변경 사항 ■ 개별 관측값 삭제로 인한 이탈도의 변경 사항 ■ 신뢰구간 이동 ■ 예측 확률 대 지레값
DPC 도표	다음 도표를 표시합니다. ■ 개별 관측값 삭제로 인한 카이제곱 적합도 통계량의 변경 사항 대 예측 사건 확률 ■ 개별 관측값 삭제로 인한 이탈도의 변경 사항 대 예측 사건 확률 이러한 도표에서 데이터 표식은 신뢰구간 이동 값에 따라 색상이 지정됩니다.
효과 도표	종속변수값에 대한 예측 확률 대 모델 효과를 도표로 표시합니다.

주: ROC 도표, 영향력 도표, DfBETAS 도표, 지레값 도표, Phat 도표 및 DPC 도표는 이진 변수가 종속변수 역할에 할당될 때만 사용할 수 있습니다.

로지스틱 회귀: 예측값 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **예측값**을 클릭합니다.

옵션 이름	설명
예측할 데이터	출력 데이터를 예측하는 데 사용할 데이터 소스를 식별합니다. 추가 데이터 체크 박스를 선택하면 이 데이터가 포함됩니다.
출력 데이터 저장	예측통계량 및 진단통계량을 출력 데이터셋에 영구히 저장할지를 지정합니다. ■ 예측값 - 각 반응레벨의 예측 확률을 저장합니다. 세 개의 레벨 1, 2, 3이 있는 반응변수 Y에 대한 개별 확률은 $Pr(Y=1)$, $Pr(Y=2)$, $Pr(Y=3)$입니다. ■ 진단 통계량 - DFBETAS 진단 통계량, 신뢰구간 변위 진단(C 및 CBAR), H 행렬, 표준오차 추정값을 가진 선형예측변수 추정값을 저장합니다. 주: SAS Enterprise Guide는 옵션 창의 작업 > 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
추가적인 통계량	■ 잔차를 선택하여 다음과 같은 통계량을 저장합니다. □ 이탈도 잔차 □ 피어슨(카이제곱) 잔차 □ 개별 관측값 삭제로 인한 이탈도의 변경 사항 □ 개별 관측값 삭제로 인한 카이제곱 적합도 통계량의 변경 사항 ■ 예측한계를 선택하면 각 예측값에 대한 신뢰구간의 상한 및 하한값과 종속변수의 예측값 평균이 출력에 포함됩니다.
출력 및 도표 표시	결과를 표시합니다. 결과를 출력하지 않으려면 체크 박스의 선택을 취소합니다. 이 체크 박스를 선택 취소하면 도표 패널의 옵션을 사용할 수 없습니다. 결과는 기본적으로 표시됩니다. 예측값을 결과에 포함하려면 예측값 표시 체크 박스를 선택합니다. 이 옵션은 예측할 데이터 영역의 데이터 소스 및 출력 및 도표 표시 체크 박스를 선택할 때만 사용할 수 있습니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

.....
주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.
.....

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

.....
주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.
.....

지도 그래프

지도 그래프 작업 정보	293
지도 그래프: 지도 유형 선택	294
지도 그래프: 분석 역할에 변수 할당	294
지도 그래프: 모양 옵션 설정	295
지도 옵션 설정	295
레이아웃 지정	296
색상 사용자 정의	296
추가 모양 옵션	297
제목 및 각주 지정	297
지도 그래프: 결과 미리 보기	297
속성 보기	298

지도 그래프 작업 정보

지도 그래프 작업은 국가, 도, 군과 같은 다른 지리적 영역에 대한 반응변수값의 변량을 보여주는 2D(2차원 지도) 또는 3D(블록 및 프리즘) 컬러 지도를 생성합니다.

다음 예에서는 각 지도 유형의 차이를 보여 줍니다. 미국 지도에서는 각 주의 유해 폐기물 지역 수를 표시할 수 있습니다.

- 2차원 지도에서는 각 주에 있는 지역 수(반응 변수)는 여러 가지 색상의 빨간색으로 표시됩니다. 유해 폐기물 지역 수가 가장 많은 주는 짙은 빨간색으로 표시되고, 유해 폐기물 지역이 없는 주는 흰색으로 표시됩니다.
- 블록 지도에서는 각 주에 있는 지역의 수가 블록의 높이와 색상으로 표시되며 각 지도 영역의 가운데에 표시됩니다.
- 프리즘 지도에서 각 주의 사이트 수는 각 지도 영역에 대한 다면체(양각된 다각형)의 높이와 색상으로 표시됩니다.

자세한 내용은 “그래픽 출력형식 선택” (5페이지)에서 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GMAP
필수 SAS 제품	Base SAS, SAS/GRAPH
추가 SAS 제품(권장)	없음

지도 그래프: 지도 유형 선택

이 옵션에 액세스하려면 선택 영역에서 **지도 그래프**를 클릭합니다.

역할 이름	설명
2D 지도	지정된 반응변수값이 다양한 색상으로 표현되는 2차원 지도를 생성합니다. 지도 데이터셋을 빠르게 보려면 빈 지도 생성 체크 박스를 선택합니다. 이렇게 하면 결과물로 단순 2차원 지도가 생성됩니다.
수직기둥 지도	지정된 반응변수의 크기 레벨이 다양한 높이 및 색상을 지닌 블록으로 표현되는 3차원 블록 지도를 생성합니다.
3D 프리즘 지도	지정된 반응변수의 크기 레벨이 다양한 높이 및 색상을 지닌 다면체(세워진 다각형)로 표현되는 3차원 프리즘 지도를 생성합니다.

지도 그래프: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

지도 그래프 작업에는 두 개의 입력 데이터셋이 필요합니다.

- 지도에 중첩된 정보가 포함된 반응 데이터 소스입니다. 반응 데이터의 예로는 판매 수치 또는 인구 통계를 들 수 있습니다. 기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 반응 데이터 소스입니다.
- 지도에 대한 X 좌표 및 Y 좌표가 포함된 지도 데이터 소스입니다. 작업을 연 후에는 지도 데이터 소스를 지정해야 합니다.

지도 데이터 소스 및 **반응 데이터 소스** 필드에는 각 입력 데이터 소스의 이름이 나타납니다. 빈 지도를 생성할 때에는 지도 데이터 소스와 반응 데이터 소스가 동일합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
ID	지도 영역을 정의하는 지도 및 반응 데이터셋의 변수를 식별합니다. ID 변수 역할에 할당한 모든 칼럼은 지도 및 반응 데이터셋 모두에 나타나야 합니다. ID 변수는 숫자 또는 문자일 수 있으며 지도 및 반응 데이터셋 모두에서 이름, 유형 및 길이가 같아야 합니다.
반응	지도에 표시될 반응 값이 포함된 반응 데이터셋의 변수를 식별합니다. 반응 변수의 값이 표현되는 방식은 지도의 유형에 따라 다릅니다. 한 개의 반응 변수만 할당할 수 있습니다. ■ 2차원 지도에서는 반응 변수의 값은 지도 영역에 채워지는 여러 가지 다른 색상으로 표현됩니다. ■ 블록 지도에서는 반응 변수의 값은 다양한 높이, 패턴 또는 색상으로 표현됩니다. ■ 프리즘 지도에서는 반응 변수의 값은 세워진 지도 영역으로 표현됩니다.
그룹 분석 칼럼	이 역할에 할당하는 칼럼의 값에 따라 생성되는 그래프의 수가 결정됩니다. 각 고유타에 대해 별도의 그래프가 생성됩니다.

지도 그래프: 모양 옵션 설정

지도 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **매핑**을 클릭합니다.

옵션 이름	설명
지도 데이터셋에서 모든 지도 영역 포함	출력에 데이터가 없는 지리적 영역을 포함할지를 지정합니다.
반응변수의 여러 값에 대해 개별 반응 레벨 생성	반응변수의 각 고유타에 여러 다른 색상을 사용할지를 지정합니다.
반응변수에 결측값 사용	반응변수의 레벨로 결측값을 포함할지를 지정합니다.
반응 값을 전체에 대한 백분율로 수집	모든 반응 값 또는 해당 통계량을 수집할 것인지와 각 지역의 전체에 대한 백분율을 그래프로 표시할 것인지를 지정합니다. 이 백분율 계산에 사용될 통계량을 지정하려면 그래프로 표시할 통계량 지

옵션 이름	설명
	정 옵션을 선택합니다. 그래프로 표시할 통계량을 지정하지 않으면 각 지역에서 첫 번째 관측값의 반응 값만 계산됩니다.
그래프로 표시할 통계량 지정	반응변수를 계산하는 데 사용할 통계량을 지정합니다. 기본적으로 각 지역에서 첫 번째 관측값의 반응 값만 계산됩니다. 하지만 반응 변수에 문자값이 포함되면 첫 번째 관측값 또는 빈도를 그래프로 표시하도록 선택할 수 있습니다. 반응변수에 숫자값이 포함되면 반응변수의 첫 번째 관측값, 합, 빈도 또는 평균 값을 그래프로 표시하도록 선택할 수 있습니다.
그래프로 표시할 반응 레벨 수 지정	지도에 나타나는 데이터 값을 표시하는 데 사용할 색상의 수를 지정합니다.
사용되는 지도 관측값의 밀도 지정	지도 포인트와 다른 지도 포인트의 근접도를 지정합니다. 여러 지도 포인트가 근접해 있으면(고밀도) 지도의 품질을 현저히 저하시키지 않고 여러 포인트를 지을 수 있습니다. 이 옵션에는 0과 6 사이의 정수를 지정할 수 있습니다. 값을 지정하지 않으면 지도 데이터셋의 모든 관측값이 표시됩니다. 주: 이 옵션은 지도 데이터셋에 DENSITY라는 변수가 포함된 경우에만 사용할 수 있습니다.

레이아웃 지정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **레이아웃**을 클릭합니다.

모든 지도 유형은 지도의 다음 영역에 대해 색상을 지정할 수 있습니다.

- 값이 없는("빈 값") 지리적 영역의 내부
- 값이 없는 지리적 영역의 외곽선
- 값이 있는 지리적 영역의 외곽선

또한 외곽선의 너비도 지정할 수 있습니다.

블록 지도를 생성하는 중이면 지도에 사용할 블록의 모양, 블록 외곽선의 색상 및 블록의 크기(또는 너비)도 지정할 수 있습니다.

색상 사용자 정의

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **색상**을 클릭합니다.

지도 그래프 작업에는 SAS에서 제공하는 색상 구성이 사용됩니다. 그러나 **사용자 정의 색상 사용** 체크 박스를 선택하고 사용할 색상을 선택하여 사용자 정의 색상 집합을 생성할 수 있습니다. 이 작업에서는 사용자 정의 색상 리스트에 나타나는 순서대로 색상을 사용합니다.

추가 모양 옵션

일부 옵션은 여러 그래프 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- [“범례 옵션 설정” \(596페이지\)](#)
- [“그래프 영역 사용자 정의” \(597페이지\)](#)

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

.....
 주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

지도 그래프: 결과 미리 보기

결과 미리 보기를 보려면 선택 영역에서 **미리 보기**를 클릭합니다.

이 결과 미리 보기에서 결과를 생성하기 전에 추가로 변경 사항을 적용할지를 결정할 수 있습니다. 지도 데이터셋이 커질 수 있으므로 이 미리 보기를 사용하면 시간을 절감할 수 있습니다. 작업에서 옵션을 변경하려면 **새로 고침**을 클릭하여 새로운 결과 미리 보기를 생성합니다. 결과 생성 중 오류가 발생하면 SAS 로그가 나타납니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

평균 및 범위관리도

평균 및 범위관리도 정보	299
평균 및 범위관리도: 분석 역할에 변수 할당	300
관리도: 관리한계 지정	301
관리도: 수행할 검정 선택	302
관리도: 도표 모양 옵션 설정	303
관리도: 분석 결과 저장	303
제목 및 각주 지정	304
속성 보기	304

평균 및 범위관리도 정보

Shewhart 관리도는 공정이 통계적 관리 상태에 있는지 여부를 결정하기 위한 그래픽 및 분석 도구입니다. 평균 및 범위관리도 작업은 하위그룹 평균 및 범위에 대한 평균과 범위관리도를 생성합니다. 이 그래프는 공정의 중심 경향 및 변동성을 분석하는 데 유용합니다.

이 작업을 사용하여 품질 관리를 분석할 수 있습니다. 예를 들어 실리콘 제조 과정에서 5개의 회로 배치 표본이 추출되었고 직경(단위: 밀리미터)을 측정한다고 가정합니다. 25개 배치에 대한 측정값이 SAS 데이터셋에 저장되어 평균 및 범위 관리도 생성에 사용됩니다. 평균 관리도의 각 포인트는 특정 배치 측정값의 평균을 나타내고 범위관리도의 각 포인트는 특정 배치의 측정값 범위를 나타냅니다. 범위관리도의 각 포인트는 특정 배치에 대한 측도의 범위를 나타냅니다. 모든 점이 관리 한계 내에 있으면 프로세스가 통계적 관리 상태에 있다는 결론을 내릴 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	SHEWHART
필수 SAS 제품	Base SAS, SAS/QC, SAS/GRAPH

요구 사항 이름	프로시저 및 제품 이름
추가 SAS 제품(권장)	SAS/STAT

평균 및 범위관리도: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
공정 측정 변수	분석할 공정 변수를 지정합니다. 이 변수에는 개별 측정값이 포함됩니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
하위그룹 식별자	데이터의 하위그룹을 식별하는 변수를 지정합니다. 변수값은 입력 테이블의 관측값이 논리적 하위그룹으로 배열되는 방법을 나타냅니다. 일반적으로 값은 다음과 같습니다. ■ 하위그룹 표본의 수집 순서를 지정하는 지수 ■ 하위그룹 표본이 수집된 날짜 또는 시간 ■ 하위그룹 표본을 고유하게 식별하는 레이블 측정값 표시 변수는 필수 변수이며 이 역할에는 단 하나의 변수만 할당할 수 있습니다. 식별된 하위그룹으로 데이터를 정렬하려면 하위그룹별 정렬 체크 박스를 선택합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
블록 변수	데이터를 연속된 하위그룹 블록으로 그룹화하는 변수를 최대 두 개까지 지정합니다. 이 블록은 범례에서 레이블로 표시되며 각 블록 변수는 범례에서 한 종류의 레이블이 됩니다.

관리도: 관리한계 지정

이 옵션에 액세스하려면 선택 영역에서 **관리한계**를 클릭합니다.

옵션 이름	설명
Sigma 한계	관리도에 그려지는 요약 하위그룹 통계량의 표준오차(sigma)의 배수로 관리한계의 너비를 지정합니다. 너비는 양수여야 합니다. 기본 배수는 3입니다.
계산 방법 선택	관리한계의 계산 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 활성 데이터에서 관리한계 계산 ■ 선택한 데이터에서 관리한계 계산 관리한계 또는 관리한계를 계산할 수 있는 모수를 포함하는 데이터셋을 지정하려면 찾아보기를 클릭합니다. 사용할 데이터셋을 선택하고 열기를 클릭합니다. 주: 선택한 데이터셋은 LIMITS= 데이터셋이어야 합니다. LIMITS= 데이터셋에 필요한 변수에 대한 자세한 내용은 생성 중인 그래프의 프로시저 문서를 참조하십시오. 모든 하위 그룹에 대한 요약통계량을 도표화하려면 하위 그룹 표본 크기가 명목 관리한계 표본 크기와 동일한지 관계없이 모든 하위그룹에 대한 요약통계량 도표를 생성합니다. 체크 박스를 선택합니다. 거의 모든 하위 그룹의 표본 크기가 공통이고 Nominal 표본 크기 n에 대응하는(가변이 아닌) 고정된 관리한계를 표시하려고 할 때 이 통계량을 도표화할 수 있습니다. 이 옵션을 선택하고 하위 그룹에 대한 표본 크기가 유의하게 다르면 관리한계 해석은 표본 크기가 n인 하위 그룹에 대해서만 의미가 있습니다. 표본 크기가 n과 같지 않은 데이터 포인트에 대해 특수 표식을 추가하려면 표본 크기에 해당하는 포인트에 대해 특수 표식을 추가하며, 표본 크기는 고정 관리한계의 명목표본 크기와 동일하지 않습니다. 체크 박스를 선택합니다. ■ 관리한계 지정 상한, 중심 및 하한 필드에 관리상한, 중심선 및 관리하한을 각각 지정합니다. 하위그룹의 표본 크기가 다양하면, 명목표본 크기를 지정해야 합니다. 주: 그룹 분석변수 역할에 변수를 할당하면 관리한계 값을 입력할 수 없습니다.

관리도: 수행할 검정 선택

이 옵션에 액세스하려면 선택 영역에서 **검정**을 클릭합니다.

주: 이 검정은 3 Sigma 한계가 아니거나 하위그룹 표본 크기에 따라 달라지는 관리한계에는 적용되지 않습니다.

옵션 이름	설명
검정 선택	체크 박스를 사용하여 런 검정, 패턴 검정 및 Western Electric 규칙이라고도 하는 특수 원인에 대한 검정을 하나 이상 요청합니다. 이러한 검정은 주관리도에 그려진 포인트에서 특정 비임의 패턴을 찾아냅니다. 이러한 비임의 패턴의 발생을 신호라고 합니다. 이러한 검정을 통해 특수한 변동의 원인이 존재하는지를 알 수 있습니다. 각 검정을 선택하면 해당 검정에 대한 설명이 설명 상자에 나타납니다.
레이블	텍스트 상자를 사용하여 특수 원인에 대한 검정이 양성으로 나타난 포인트에 대한 레이블을 입력합니다. 레이블의 문자 수는 16자를 초과할 수 없습니다. 각 레이블은 검정에서 신호가 발생한 모든 점에 표시됩니다.
레이블 색상	■ 드롭다운 리스트를 사용하여 지정한 레이블의 색상을 선택합니다. ■ 영역 선 표시 체크 박스를 사용하여 주관리도에 A, B, C 영역을 구분하는 선을 표시합니다. 이 영역은 관리한계 사이의 간격을 여섯 개의 동일한 간격을 지닌 하위 간격으로 분할함으로써 구성됩니다. ■ 3 sigma 한계 무시 체크 박스를 사용하여 기본 배수 3 이외의 sigma 관리한계를 지정했을 때 특수한 원인을 찾는 검정을 수행할 수 있습니다. sigma 한계를 변경하려면 선택 영역에서 관리한계를 클릭합니다. ■ 모든 점을 검정의 시작점으로 사용 체크 박스를 사용하여 중복되는 점 패턴에 특수 원인 검정을 적용합니다.

관리도: 도표 모양 옵션 설정

일부 모양 옵션은 여러 관리도 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- “축 옵션 설정” (605페이지)
- “참조선 추가” (605페이지)
- “기타 도표 옵션 설정” (606페이지)
- “블록 변수 옵션 설정” (608페이지)

관리도: 분석 결과 저장

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

분석 결과를 저장하기 위해 다양한 통계량을 포함하는 출력 데이터셋을 생성할 수 있습니다. 영구적으로 저장하려는 출력 데이터의 체크 박스를 선택합니다.

역할 이름	설명
하위그룹 통계량과 관리한계 데이터셋	관리도에 그려지는 정보를 포함하는 출력 데이터셋을 생성합니다. 여기에는 하위그룹 식별자값 및 해당 요약통계량과 관리한계가 포함됩니다. 주: 이 옵션은 OUTTABLE= 데이터셋을 생성합니다. SHEWHART 프로시저에서 TABLE= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.
하위그룹 통계량 출력 데이터셋	하위그룹 요약통계량을 포함하는 출력 데이터셋을 생성합니다. 주: 이 옵션은 OUTHISTORY= 데이터셋을 생성합니다. SHEWHART 프로시저에서 HISTORY= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.
관리한계 출력 데이터셋	관리한계를 저장하는 출력 데이터셋을 생성합니다. 추후에 이 작업 사용 시 관리한계의 계산 방법을 선택할 때 이 데이터셋을 사용할 수 있습니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식

■ 결과의 그래픽 출력에 대한 출력형식

.....

주: 기존 작업의 레이블을 변경하는 경우 속성 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

.....

평균 및 표준편차관리도

평균 및 표준편차관리도 정보	307
평균 및 표준편차관리도: 분석 역할에 변수 할당	308
관리도: 관리한계 지정	309
관리도: 수행할 검정 선택	310
관리도: 도표 모양 옵션 설정	311
관리도: 분석 결과 저장	311
제목 및 각주 지정	312
속성 보기	312

평균 및 표준편차관리도 정보

Shewhart 관리도는 공정이 통계적 관리 상태에 있는지 여부를 결정하기 위한 그래픽 및 분석 도구입니다. 평균 및 표준편차관리도 작업은 하위그룹 평균 및 하위그룹 표준편차에 대한 평균 및 표준편차관리도를 생성합니다. 이 그래프는 공정의 중심 경향 및 변동성을 분석하는데 유용합니다.

이 작업을 사용하여 출력 분포를 확인하고 프로세스가 통계적 관리 상태에 있는지 여부를 확인할 수 있습니다. 예를 들어 석유 회사가 터빈으로 물을 가열한다고 가정합니다. 가열한 물은 증기가 되어 땅속으로 주입됨으로써 석유의 점성이 낮아지고 추출이 쉬워집니다. 이러한 프로세스는 매일 20회 수행되면 물을 희망 온도까지 가열하는 데 사용되는 전력량(단위: 킬로와트)을 기록합니다. 평균관리도의 각 포인트는 특정 날짜의 측정 평균을 나타냅니다. 평표준편차관리도의 각 포인트는 특정 날짜에 대한 측정값의 표준편차를 나타냅니다. 모든 포인트가 관리한계 내에 위치하면 공정이 통계적 관리 상태에 있다고 결론을 내릴 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
----------	--------------

사용되는 SAS 프로시저	SHEWHART
---------------	----------

요구 사항 이름	프로시저 및 제품 이름
필수 SAS 제품	Base SAS, SAS/QC, SAS/GRAPH
추가 SAS 제품(권장)	SAS/STAT

평균 및 표준편차관리도: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
공정 측정 변수	분석할 공정 변수를 지정합니다. 이 변수에는 개별 측정값이 포함됩니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
하위그룹 식별자	데이터의 하위그룹을 식별하는 변수를 지정합니다. 변수값은 입력 테이블의 관측값이 논리적 하위그룹으로 배열되는 방법을 나타냅니다. 일반적으로 값은 다음과 같습니다. ■ 하위그룹 표본의 수집 순서를 지정하는 지수 ■ 하위그룹 표본이 수집된 날짜 또는 시간 ■ 하위그룹 표본을 고유하게 식별하는 레이블 측정값 표시 변수는 필수 변수이며 이 역할에는 단 하나의 변수만 할당할 수 있습니다. 식별된 하위그룹으로 데이터를 정렬하려면 하위그룹별 정렬 체크 박스를 선택합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.

역할 이름	설명
블록 변수	데이터를 연속된 하위그룹 블록으로 그룹화하는 변수를 최대 두 개 까지 지정합니다. 이 블록은 범례에서 레이블로 표시되며 각 블록 변수는 범례에서 한 종류의 레이블이 됩니다.

관리도: 관리한계 지정

이 옵션에 액세스하려면 선택 영역에서 **관리한계**를 클릭합니다.

옵션 이름	설명
Sigma 한계	관리도에 그려지는 요약 하위그룹 통계량의 표준오차(sigma)의 배수로 관리한계의 너비를 지정합니다. 너비는 양수여야 합니다. 기본 배수는 3입니다.

계산 방법 선택 관리한계의 계산 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다.

- **활성 데이터에서 관리한계 계산**
- **선택한 데이터에서 관리한계 계산**

관리한계 또는 관리한계를 계산할 수 있는 모수를 포함하는 데이터셋을 지정하려면 **찾아보기**를 클릭합니다. 사용할 데이터셋을 선택하고 **열기**를 클릭합니다.

주: 선택한 데이터셋은 LIMITS= 데이터셋이어야 합니다. LIMITS= 데이터셋에 필요한 변수에 대한 자세한 내용은 생성 중인 그래프의 프로시저 문서를 참조하십시오.

모든 하위 그룹에 대한 요약통계량을 도표화하려면 **하위그룹 표본 크기가 명목 관리한계 표본 크기와 동일하지 관계없이 모든 하위그룹에 대한 요약통계량 도표를 생성합니다**. 체크 박스를 선택합니다. 거의 모든 하위 그룹의 표본 크기가 공통이고 Nominal 표본 크기 n 에 대응하는(가변이 아닌) 고정된 관리한계를 표시하려고 할 때 이 통계량을 도표화할 수 있습니다. 이 옵션을 선택하고 하위 그룹에 대한 표본 크기가 유의하게 다르면 관리한계 해석은 표본 크기가 n 인 하위 그룹에 대해서만 의미가 있습니다.

표본 크기가 n 과 같지 않은 데이터 포인트에 대해 특수 표식을 추가하려면 **표본 크기에 해당하는 포인트에 대해 특수 표식을 추가하며, 표본 크기는 고정 관리한계의 명목표본 크기와 동일하지 않습니다**. 체크 박스를 선택합니다.

- **관리한계 지정**

옵션 이름	설명
	상한, 중심 및 하한 필드에 관리상한, 중심선 및 관리하한을 각각 지정합니다. 하위그룹의 표본 크기가 다양하면, 명목표본 크기를 지정해야 합니다. 주: 그룹 분석변수 역할에 변수를 할당하면 관리한계 값을 입력할 수 없습니다.

관리도: 수행할 검정 선택

이 옵션에 액세스하려면 선택 영역에서 **검정**을 클릭합니다.

주: 이 검정은 3 Sigma 한계가 아니거나 하위그룹 표본 크기에 따라 달라지는 관리한계에는 적용되지 않습니다.

옵션 이름	설명
검정 선택	체크 박스를 사용하여 런 검정, 패턴 검정 및 Western Electric 규칙이라고도 하는 특수 원인에 대한 검정을 하나 이상 요청합니다. 이러한 검정은 주관리도에 그려진 포인트에서 특정 비임의 패턴을 찾아냅니다. 이러한 비임의 패턴의 발생을 신호라고 합니다. 이러한 검정을 통해 특수한 변동의 원인이 존재하는지를 알 수 있습니다. 각 검정을 선택하면 해당 검정에 대한 설명이 설명 상자에 나타납니다.
레이블	텍스트 상자를 사용하여 특수 원인에 대한 검정이 양성으로 나타난 포인트에 대한 레이블을 입력합니다. 레이블의 문자 수는 16자를 초과할 수 없습니다. 각 레이블은 검정에서 신호가 발생한 모든 점에 표시됩니다.
레이블 색상	■ 드롭다운 리스트를 사용하여 지정한 레이블의 색상을 선택합니다. ■ 영역 선 표시 체크 박스를 사용하여 주관리도에 A, B, C 영역을 구분하는 선을 표시합니다. 이 영역은 관리한계 사이의 간격을 여섯 개의 동일한 간격을 지닌 하위 간격으로 분할함으로써 구성됩니다. ■ 3 sigma 한계 무시 체크 박스를 사용하여 기본 배수 3 이외의 sigma 관리한계를 지정했을 때 특수한 원인을 찾는 검정을 수행할 수 있습니다. sigma 한계를 변경하려면 선택 영역에서 관리한계를 클릭합니다.

옵션 이름	설명
	■ 모든 점을 검정의 시작점으로 사용 체크 박스를 사용하여 중복되는 점 패턴에 특수 원인 검정을 적용합니다.

관리도: 도표 모양 옵션 설정

일부 모양 옵션은 여러 관리도 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- “축 옵션 설정” (605페이지)
- “참조선 추가” (605페이지)
- “기타 도표 옵션 설정” (606페이지)
- “블록 변수 옵션 설정” (608페이지)

관리도: 분석 결과 저장

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

분석 결과를 저장하기 위해 다양한 통계량을 포함하는 출력 데이터셋을 생성할 수 있습니다. 영구적으로 저장하려는 출력 데이터의 체크 박스를 선택합니다.

역할 이름	설명
하위그룹 통계량과 관리한계 데이터셋	관리도에 그려지는 정보를 포함하는 출력 데이터셋을 생성합니다. 여기에는 하위그룹 식별자값 및 해당 요약통계량과 관리한계가 포함됩니다. 주: 이 옵션은 OUTTABLE= 데이터셋을 생성합니다. SHEWHART 프로시저에서 TABLE= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.
하위그룹 통계량 출력 데이터셋	하위그룹 요약통계량을 포함하는 출력 데이터셋을 생성합니다. 주: 이 옵션은 OUTHISTORY= 데이터셋을 생성합니다. SHEWHART 프로시저에서 HISTORY= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.

역할 이름	설명
관리한계 출력 데이터셋	관리한계를 저장하는 출력 데이터셋을 생성합니다. 추후에 이 작업 사용 시 관리한계의 계산 방법을 선택할 때 이 데이터셋을 사용할 수 있습니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버

- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

혼합모형

혼합모형 작업 정보	315
혼합모형: 분석 역할에 변수 할당	316
혼합모형: 고정 효과 모형 생성	317
혼합모형: 고정 효과 모형 옵션 설정	317
혼합모형: 임의효과 지정	318
혼합모형: 반복 효과 지정	320
혼합모형: 최소제곱평균 지정	321
혼합모형: 도표 생성	322
혼합모형: 예측값 옵션 설정	323
제목 및 각주 지정	324
속성 보기	324

혼합모형 작업 정보

혼합선형모델은 일반화된 표준선형모델으로서 데이터에 상관관계의 존재와 비상수 변이를 허용합니다. 혼합선형모델을 사용하면 데이터의 평균은 물론, 분산 및 공분산 모형도 개발할 수 있습니다.

이 작업을 사용하여 두 가지 주효과 사이의 상호관계를 모델로 만들 수 있습니다. 예를 들어 18명의 키를 측정하고 이 18명은 Family와 Gender에 따라 분류된다고 가정합니다. Height를 종속변수로, Gender를 고정분류효과로, Family를 임의분류효과로 사용하면, 혼합 모형 작업을 사용하여 혼합선형모델을 이 데이터에 고정할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
----------	--------------

사용되는 SAS 프로시저	MIXED
---------------	-------

요구 사항 이름	프로시저 및 제품 이름
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

혼합모형: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
종속변수	종속(반응)변수로 사용할 변수를 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다.
양적변수	연속 독립변수(설명변수)로 사용할 변수를 지정합니다.
분류변수	이산형 독립변수 효과로 사용할 변수를 지정합니다. 이 역할에는 숫자 또는 문자 변수를 할당할 수 있으나 이러한 변수는 한정된 수의 이산 값을 가지고 있어야 합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 종속변수나 독립변수로 이미 선택한 변수로는 분석을 그룹화할 수 없습니다.
상대 가중값 변수	이 역할에 지정된 각 변수값은 반응변수로 선택된 변수의 대응값에 가중값을 적용하는 데 사용됩니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.
ID 레이블	예측값 테이블에 포함할 추가 변수를 지정합니다.

혼합모형: 고정 효과 모형 생성

이 옵션에 액세스하려면 선택 영역에서 **고정 효과 모형**을 클릭합니다.

여러 SAS 작업에 모델을 생성할 수 있습니다. 자세한 내용은 [부록 3, “모델 생성” \(609페이지\)](#)에서 참조하십시오.

혼합모형: 고정 효과 모형 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **고정 효과 모형 옵션**을 클릭합니다.

옵션 이름	설명
절편에 대한 검정 표시	절편에 대해 수행한 가설 검정의 결과를 포함하려면 이 옵션을 선택합니다.
가설 검정 유형	각 가설 검정은 하나 이상의 추정 가능 함수를 사용하여 가설을 검정합니다. 추정 가능 함수는 모델 모수의 선형 조합입니다. 기본 가설 검정은 Type 3입니다. 해당 체크 박스를 선택하여 가설 검정을 추가 또는 제거할 수 있습니다.
모수 추정값 표시	모수 추정값은 선택한 추정 방법에 따라 결정됩니다.
모수 추정값의 신뢰 한계	신뢰구간은 미리 지정된 신뢰수준으로 추정되는 합리적인 모수값의 범위를 정하는 상한 및 하한 경계입니다. 신뢰수준 값은 직접 입력하거나 드롭다운 리스트에서 선택할 수 있습니다. 기본값은 95% 신뢰수준입니다.
추정 방법	추정 방법은 모델의 분산과 공분산 행렬 부분의 요소를 추정하는 데 사용되는 기술입니다. 기본 방법은 잔차 최대 가능도 입니다. 드롭다운 메뉴에서 다른 추정 방법을 선택할 수도 있습니다.
자유도 방법	자유도는 합에 포함되는 독립적인 정보의 수를 참조하는 통계량입니다. 드롭다운 메뉴에서 다음 방법 중 하나를 선택할 수 있습니다. 기본 방법은 기본 입니다. 이 옵션을 선택하면, 혼합모형 작업은 모델 특징 및 모델이 임의효과와 반복 효과를 포함하는지 여부에 따라 기본 방법을 선택합니다. 다음 리스트에는 임의 또는 반복효과의 유무에 따라 기본 방법이 선택되는 방법이 나와 있습니다. ■ 모델에 임의효과가 포함되어 있으면, 기본 방법은 봉쇄 모델입니다.

옵션 이름	설명
	■ 모델이 반복 효과만 포함하는 경우 기본 방법은 사이 및 대상 내 부분 방법입니다. ■ 모델에 임의 및 반복효과가 포함되어 있으면, 기본 방법은 봉쇄 모델입니다. ■ 모델에 임의 또는 반복효과가 포함되어 있지 않으면, 기본 방법은 잔차모델입니다.
추정가능함수	Type I, Type II 및 Type III 체크 박스를 선택하여 각 모델효과에 대해 Type I - Type III 추정가능함수의 행렬 계수를 표시합니다. 특정 추정가능함수에 대한 행렬 계수를 표시하려면 유형 옆의 체크 박스를 선택합니다. 주: 절편에 대한 검정 표시 체크 박스를 선택하면 절편에 대한 추정 가능 함수를 표시할 수 있습니다.
조율모수 사용	주: 이러한 조율모수는 문제가 불량조건일 때만 유용합니다. Type III 함수의 민감도를 변경하려면 Type III 함수의 민감도 상자에 새로운 값을 입력합니다. 지정 값보다 작은 절대값을 가진 추정 가능 함수의 모든 요소는 0으로 설정됩니다. 기본값은 1E-08입니다. 행렬을 지울 때 사용되는 민감도를 변경하려면 비정칙성 확인값 상자에 원하는 값을 입력합니다. 대각피벗원소가 사용자가 지정한 값을 곱한 행렬의 원래 대각원소보다 작을 때 관련 칼럼은 이전 칼럼에 대하여 1차 종속인 것으로 선언되고 관련 모수는 0으로 설정됩니다. 기본값은 1E-07입니다.

혼합모형: 임의효과 지정

이 옵션에 액세스하려면 선택 영역에서 **임의효과**를 클릭합니다.

- 1 **추가**를 클릭하여 효과를 **추정할 임의효과** 상자에 추가합니다.
- 2 리스트의 각 효과에 대해 **임의효과와 옵션** 상자에서 옵션을 지정합니다. 다음과 같은 옵션을 사용할 수 있습니다.

옵션 이름	설명
임의효과	혼합모형에 임의효과를 지정합니다. 기본적으로 임의효과는 사용되지 않습니다. 모형에서 임의효과로 절편을 포함할 수도 있습니다. 효과 빌더 대화 상자를 사용하여 혼합모형에 대한 임의효과를 지정합니다. 효과를 지정하려면 한 개 이상의 변수를 양적변수 또는 분류변수 역할에 할당해야 합니다. 변수 조합을 선택하여 교차, 지분, 팩토리

옵션 이름	설명
	얼 또는 다항 효과를 생성할 수 있습니다. 임의효과를 지정한 후, 확인을 클릭하여 변경 사항을 저장하고 혼합모형 작업으로 돌아갑니다. 모형을 생성하는 단계는 “혼합모형: 고정 효과 모형 생성” (317페이지)을 참조하십시오.
신뢰한계	각 임의효과 추정치에 대해 t-유형 신뢰한계를 표시할지를 지정합니다.
모델 식별자	분석에 포함할 서브젝트 식별자를 지정합니다. 전체 서브젝트에 대해 완전한 독립성을 가정합니다. 이 옵션은 G 행렬에서 동일한 블록을 가진 블록-대각 구조를 생성합니다. Z 행렬이 이 블록-대각을 수용하도록 수정됩니다. 효과 빌더 대화 상자를 사용하여 혼합모형에 대한 서브젝트 식별자를 지정합니다. 효과를 지정하려면 한 개 이상의 변수를 분류변수 역할에 할당해야 합니다. 변수 조합을 선택하여 교차, 지분, 팩토리얼 또는 다항 효과를 생성할 수 있습니다. 지정한 효과가 서브젝트 식별자 상자에 표시됩니다. 이 효과를 제거하려면 효과 제거를 클릭합니다. 확인을 클릭하여 이 서브젝트 식별자를 저장하고 혼합모형 작업으로 돌아갑니다.
그룹 효과	분석에 포함할 그룹 식별자를 지정합니다. 이 옵션으로 G 행렬의 공분산 구조에서 이질성을 지정하는 효과를 정의합니다. 동일한 그룹 효과를 가진 관측값은 공분산 모수가 모두 같습니다. 새로운 레벨의 그룹 효과는 각각 원본 그룹과 동일한 구조를 갖는 새로운 공분산 모수 집합을 생성합니다. 효과 빌더 대화 상자를 사용하여 혼합모형에 대한 그룹 식별자를 지정합니다. 효과를 지정하려면 한 개 이상의 변수를 분류변수 역할에 할당해야 합니다. 변수 조합을 선택하여 교차, 지분, 팩토리얼 또는 다항 효과를 생성할 수 있습니다. 지정한 효과가 그룹 식별자 상자에 표시됩니다. 이 효과를 제거하려면 효과 제거를 클릭합니다. 확인을 클릭하여 이 그룹 식별자를 저장하고 혼합모형 작업으로 돌아갑니다.
공분산 구조	G 행렬의 공분산 구조를 지정합니다.
추가 옵션	분석에 G 행렬을 표시할지 여부와 사용할 G 행렬 옵션을 지정합니다. 다음과 같은 옵션을 사용할 수 있습니다. ■ G 행렬 표시 - 추정 G 행렬을 표시합니다. ■ G 행렬의 Cholesky 근 표시 - 추정 G 행렬의 하한 삼각 Cholesky 근을 표시합니다. ■ G 행렬의 역 Cholesky 근 표시 - 추정 G 행렬의 역-Cholesky 근을 표시합니다. ■ G 행렬의 역행렬 표시 - 추정 G 행렬의 역행렬을 표시합니다. ■ 상관행렬 표시 - 추정 G 행렬에 대응하는 상관행렬을 표시합니다. ■ 추정값 포함 - 임의효과 모수에 대한 해결책을 표시합니다.

혼합모형: 반복 효과 지정

이 옵션에 액세스하려면 선택 영역에서 **반복 효과**를 클릭합니다.

리스트의 각 효과에 대해 **반복 효과와 옵션** 상자에서 옵션을 지정합니다. 다음과 같은 옵션을 사용할 수 있습니다.

옵션 이름	설명
사용할 효과	분석에 사용할 서브젝트 내 효과를 선택합니다. 서브젝트 내 효과는 분류변수 역할에 지정한 변수입니다. 데이터가 서브젝트 내에서 올바른 반복 측도로 정렬되지 않으면 서브젝트 내 효과를 지정할 수 없습니다. 예를 들어 시간이 반복 효과이고 데이터가 각 서브젝트 내에서 시간으로 정렬되어 있지 않으면 모델에서 데이터를 제대로 해석하려면 서브젝트 내 효과로 시간을 지정해야 합니다. 기본적으로 서브젝트 내 효과는 사용되지 않습니다.
모델 식별자	분석에 포함할 서브젝트 식별자를 지정합니다. 대부분의 반복 측도 모델에서는 반복 효과가 필요하지 않습니다. 서브젝트 식별자를 지정하면 반복 문을 사용하여 혼합모형의 R 행렬을 지정합니다. 서브젝트 식별자는 R 행렬의 블록을 정의합니다. 전체 서브젝트에 대해 완전한 독립성을 가정합니다. 효과 빌더 대화 상자를 사용하여 혼합모형에 대한 서브젝트 식별자를 지정합니다. 효과를 지정하려면 한 개 이상의 변수를 분류변수 역할에 할당해야 합니다. 변수 조합을 선택하여 교차, 지분, 팩토리얼 또는 다항 효과를 생성할 수 있습니다. 지정한 효과가 서브젝트 식별자 상자에 표시됩니다. 이 효과를 제거하려면 효과 제거를 클릭합니다. 확인을 클릭하여 이 서브젝트 식별자를 저장하고 혼합모형 작업으로 돌아갑니다.
공분산 구조	R 행렬의 공분산 구조를 지정합니다.
그룹 효과	분석에 포함할 그룹 식별자를 지정합니다. 이 옵션으로 G 행렬의 공분산 구조에서 이질성을 지정하는 효과를 정의합니다. 동일한 그룹 효과를 가진 관측값은 공분산 모수가 모두 같습니다. 새로운 레벨의 그룹 효과는 각각 원본 그룹과 동일한 구조를 갖는 새로운 공분산 모수 집합을 생성합니다. 효과 빌더 대화 상자를 사용하여 혼합모형에 대한 그룹 식별자를 지정합니다. 효과를 지정하려면 한 개 이상의 변수를 분류변수 역할에 할당해야 합니다. 변수 조합을 선택하여 교차, 지분, 팩토리얼 또는 다항 효과를 생성할 수 있습니다.

옵션 이름	설명
	지정한 효과가 그룹 식별자 상자에 표시됩니다. 이 효과를 제거하려면 효과 제거 를 클릭합니다. 확인 을 클릭하여 이 그룹 식별자를 저장하고 혼합모형 작업으로 돌아갑니다.
R 행렬 옵션	분석에 <i>R</i> 행렬을 표시할지 여부와 사용할 <i>R</i> 행렬 옵션을 지정합니다. 다음과 같은 옵션을 사용할 수 있습니다. ■ R 행렬 표시 - 추정 <i>R</i> 행렬을 표시합니다. ■ 상관행렬 표시 - 추정 <i>R</i> 행렬에 대응하는 상관행렬을 표시합니다. ■ R 행렬의 역행렬 표시 - 추정 <i>R</i> 행렬의 역행렬을 표시합니다. ■ R 행렬의 Cholesky 근 표시 - 추정 <i>R</i> 행렬의 하한 삼각 Cholesky 근을 표시합니다.

혼합모형: 최소제곱평균 지정

선택 영역에서 **최소제곱 Post Hoc 검정**을 클릭합니다.

- 1 **추가**를 클릭하여 효과를 **추정할 효과** 상자에 추가합니다.
- 2 리스트의 각 효과에 대해 **최소제곱평균 검정 및 옵션** 상자에서 옵션을 지정합니다. 다음과 같은 옵션을 사용할 수 있습니다.

옵션 이름	설명
사용할 효과	분석에서 사용할 분류변수 효과를 선택합니다. 분류변수 효과는 분류 변수 역할에 지정한 변수입니다. 기본적으로 분류변수 효과는 사용되지 않습니다.
평균 비교	분석에 사용할 비교 방법을 지정합니다. ■ 차이에 대한 P-값 표시 - 최소제곱평균의 차이를 표시합니다. 다음과 같은 옵션을 사용할 수 있습니다. □ 없음 - 차이를 표시하지 않습니다. □ 기본값 - 모든 쌍별 차이를 표시합니다. □ 모든 쌍별 차 - 모든 쌍별 차이를 표시합니다. 기본값입니다. □ 첫 번째 레벨을 사용하는 대조군 - 지정된 각 최소제곱평균 효과의 첫 번째 레벨을 사용하여 차이를 계산합니다. 양측 검정 및 신뢰한계는 이러한 유형의 차이와 관련됩니다. □ 대조군보다 작다는 가설 검정 - 비관리 레벨이 관리 레벨보다 유의하게 작은지를 검정합니다. 단측 결과에 이 검정을 사용할 수 있습니다.

옵션 이름	설명
	□ 대조군보다 크다는 가설 검정 - 비관리 레벨이 관리 레벨보다 유의하게 크지를 검정합니다. 단측 결과에 이 검정을 사용할 수 있습니다. ■ 비교를 위한 수정 방법 - 최소제곱평균의 차이에 대한 p-값 및 신뢰 한계에 대해 다중비교 수정을 수행할 수 있습니다.
신뢰한계	최소제곱평균에 대해 구성된 t -유형의 신뢰한계를 포함할지를 지정합니다.
공변량	계산에 공변량을 포함할지 여부와 L 행렬 계수를 표시할지를 지정합니다. ■ 공변량 값 - 공변량 값을 계산하는 방법을 지정합니다. 다음과 같은 옵션을 사용할 수 있습니다. □ 평균값 - 모든 공변량 효과를 표준 최소제곱평균 계산을 위한 해당 평균값으로 설정합니다. □ 평균값과 수정된 교차제곱합 - 공변량 효과를 해당 평균값으로 설정하고 공변량의 교차제곱합을 수정합니다. □ 변수별 값 지정 - 공변량에 임의의 값을 지정할 수 있습니다. 출력에서 추가되는 칼럼은 공변량의 값을 나타냅니다. ■ L 행렬 계수 표시 - 모든 최소제곱평균 효과에 대한 L 행렬 계수를 표시합니다.

혼합모형: 도표 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

기본적으로 현재 선택된 데이터에 적용 가능한 모든 도표가 출력에 포함됩니다. 그러나 **사용자 정의 도표 리스트** 옵션을 선택하여 출력에 포함할 도표를 선택할 수 있습니다. 다음 옵션 중에서 선택할 수 있습니다.

옵션 이름	설명
상자 도표	분류 효과만으로 구성된 모델의 효과에 대한 상자 도표를 요청합니다. 이 효과에는 두 개 이상의 분류변수(교호작용 및 지분 효과)가 포함될 수 있지만 연속 변수는 포함될 수 없습니다. 기본적으로 상자 도표는 고정 효과 모형, 임의 효과 모형 및 반복 효과 모형의 제한적 효과에 대한 원잔차(raw residual)를 기반으로 생성됩니다.
거리 도표	우도 또는 제한된 우도 거리의 도표를 생성합니다.

옵션 이름	설명
영향력 추정 도표	영향력 분석의 삭제 추정값 패널을 생성합니다.
영향력 통계량 패널	영향력 통계량 패널을 생성합니다. 반복 영향도 분석의 경우 패널에 고정 효과 및 공분산 모수에 대한 Cook의 D 및 CovRatio 통계량이 표시되므로 두 추정값 유형 모두에 대한 추정값 및 정밀도에 대한 영향을 측정할 수 있습니다. 비반복 분석에서는 고정 효과에 대한 통계량만 도표화됩니다.
원잔차 도표	원잔차(raw residual) 패널을 생성합니다. 기본적으로 조건부 잔차가 산출됩니다.
스튜던트화 잔차 도표	스튜던트화 잔차 패널을 생성합니다. 기본적으로 조건부 잔차가 산출됩니다.
피어슨 잔차 도표	피어슨 잔차 패널을 생성합니다. 기본적으로 조건부 잔차가 산출됩니다.
PRESS 잔차 도표	PRESS 잔차 또는 PRESS 통계량의 도표를 생성합니다. 주변 평균의 "leave-one-out" 또는 "leave-set-out" 예측을 기반으로 합니다.
척도화 잔차 도표	척도화 잔차(주변분산-공분산 행렬의 역-Cholesky 근으로 척도화된 반응 및 주변잔차)를 기반으로 잔차 그래프 패널을 생성합니다.

혼합모형: 예측값 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **예측값**을 클릭합니다.

옵션 이름	설명
예측할 데이터	출력 데이터를 예측하는 데 사용할 데이터 소스를 식별합니다. 추가 데이터 체크 박스를 선택하면 찾아보기 를 클릭하여 추가 데이터의 위치를 지정합니다.
출력 데이터 저장	예측 및 예측 평균을 출력 데이터셋에 저장할지를 지정합니다. 주: SAS Enterprise Guide는 옵션 창의 작업 > 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기 를 클릭합니다.

옵션 이름	설명
예측값 표시	결과에 예측값을 표시합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

모델 스코어링

모델 스코어링 마법사 정보	327
모델 스코어링: 데이터 선택	328
모델 스코어링: 스코어링 모델 선택	328
모델 스코어링: 필수 모델 입력에 데이터 변수 매핑	328
모델 스코어링: 출력 데이터 칼럼 선택	329
모델 스코어링: 출력 데이터 저장	329
모델 스코어링: 선택 확인	329
모델 스코어링: 결과 해석	330

모델 스코어링 마법사 정보

모델 스코어링 마법사를 사용하여 기존 **SAS Enterprise Miner** 예측 모델에 대한 데이터셋을 스코어링할 수 있습니다. 원시 데이터셋의 순위화나 스코어링은 대부분의 데이터 마이닝 문제의 목표입니다. 예를 들어 예측 모형을 사용하여 대출 신청자 리스트, 직접 마케팅 응답자 또는 잠재적 사기 사례 등을 포함하는 데이터셋 스코어링을 수행할 수 있습니다. 이 작업을 사용하여 입력 데이터셋에 대한 하나 이상의 조각 변수를 생성할 수 있습니다.

스코어링 코드를 생성하는 모델이 먼저 생성되고 **SAS Enterprise Miner**에서 검증되어야 **모델 스코어링** 마법사의 모델에 액세스할 수 있습니다. 모형으로 생성된 스코어링 코드는 SAS 메타데이터 레파지토리에도 저장되어야 합니다.

이 마법사를 사용하여 소비자 속성의 데이터 테이블을 평가할 수 있습니다. 이 분석에서 구매 가능성이 가장 높은 고객의 메일 리스트를 생성할 수 있습니다. 다른 모델의 스코어 코드를 사용하여 하나 이상의 속성 스코어를 기반으로 각 고객을 시장 세그먼트에 할당할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	없음

요구 사항 이름	프로시저 및 제품 이름
필수 SAS 제품	SAS Enterprise Miner
추가 SAS 제품(권장)	없음

모델 스코어링: 데이터 선택

기본적으로 **모델 스코어링** 마법사에서는 프로젝트에서 활성 데이터를 사용합니다.

다음을 클릭하여 스코어링 모델을 선택합니다.

모델 스코어링: 스코어링 모델 선택

- 1 데이터를 선택해야 스코어링 모델을 선택할 수 있습니다.
- 2 **찾아보기**를 클릭하여 스코어링 모델을 선택합니다. **모형 스코어링** 창이 나타납니다. **찾아보기** 탭에서 사용하려는 모형을 선택합니다.

모형을 선택한 후 **확인**을 클릭합니다. 모형에 대한 디렉터리 경로가 **스코어링 모형** 상자에 나타나고 해당 속성이 **모형 속성** 상자에 나타납니다. 모형으로 생성된 SAS 코드를 보려면 **소스**를 클릭합니다.
- 3 **다음**을 클릭하여 필수 모델 입력에 데이터 변수를 매핑합니다.

모델 스코어링: 필수 모델 입력에 데이터 변수 매핑

- 1 스코어링 모델을 선택해야 필수 모델 입력에 데이터 변수를 매핑할 수 있습니다.
- 2 각 필수 모델 입력에서는 데이터 변수를 선택하여 소스 데이터에서 사용합니다. 기본적으로 동일한 이름 및 유형을 가진 모형 입력 및 데이터 변수가 매핑됩니다. 모형 입력으로 매핑된 데이터 유형을 변경하려면 변수를 선택하고 **편집**을 클릭합니다. 나타나는 드롭다운 리스트에서 사용할 데이터 변수를 선택합니다.

- 3 다음을 클릭하여 출력 칼럼을 지정합니다.

주: 다음 버튼은 데이터 변수가 모든 필수 입력으로 매핑된 때에만 사용할 수 있습니다.

모델 스코어링: 출력 데이터 칼럼 선택

- 1 필수 입력에 데이터 변수를 매핑해야 출력 데이터 칼럼을 선택할 수 있습니다.
- 2 **데이터 변수** 탭과 **모델 출력** 탭에서 항목을 선택하여 출력에 포함할 수 있습니다. **데이터 변수** 탭에는 입력 데이터셋의 모든 변수가 포함됩니다. **모형 출력** 탭에는 **Date**, **DateTime** 및 **ModelKey** 키가 포함되어 있습니다. 각 탭에서 출력에 포함하려는 항목을 선택하고 **선택한 칼럼** 영역으로 끌어 놓습니다. **선택한 칼럼** 영역의 변수 모형 출력이 출력에 나타납니다. 변수 및 출력의 순서를 지정하려면 위쪽/아래쪽 화살표를 사용합니다.
- 3 다음을 클릭하여 출력 데이터를 저장합니다.

모델 스코어링: 출력 데이터 저장

출력 데이터 칼럼을 선택해야 출력 데이터를 저장할 수 있습니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

다음을 클릭하여 선택을 확인합니다.

출력 데이터셋에는 원본 데이터와 더불어 예측값, 잔차 및 분류 결과와 같은 스코어가 포함됩니다. 출력 데이터셋의 변수에 대한 자세한 내용은 결과 해석을 참조하십시오.

모델 스코어링: 선택 확인

선택을 확인하기 전에 출력 데이터 칼럼을 선택해야 합니다.

마지막 단계에 선택 요약이 포함됩니다. 이러한 선택 사항을 변경하려면 **뒤로**를 클릭합니다. 작업을 실행하려면 **마침**을 클릭합니다.

모델 스코어링: 결과 해석

출력 데이터셋에는 원본 데이터와 더불어 예측값, 잔차 및 분류 결과와 같은 스코어가 포함됩니다. 일반적으로 타겟 변수, 입력 변수 또는 입력 데이터셋의 이름에 접두어를 추가하면 스코어 변수 이름이 생성됩니다.

다음 테이블에는 이러한 각 접두어가 리스트에 나열되고, 접두어에 대한 간략한 설명이 제공되고, 타겟 변수가 필요한지를 지정하며, 이 변수가 의사결정 행렬에 있는지를 지정합니다. 결과 해석에 대한 자세한 내용은 **SAS Enterprise Miner**의 도움말을 참조하십시오.

접두어	설명	타겟이 필요합 니까?	의사결 정 행 렬입니 까?
AOV_	16개의 동일한 간격의 그룹으로 나뉘지는 구간 변수의 BIN 번호를 구별합니다.	아니요	아니요
B_	Bin 변수	예	아니요
BL_	의사결정에 대한 가능한 최적 손실분, $-B(i)$	예	예
BP_	의사결정에 대한 가능한 최적 이익분, $B(i)$	예	예
CL_	타겟값에서 계산된 손실분, $-C(i)$	예	예
CP_	타겟값에서 계산된 이익분, $C(i)$	예	예
D_	모델로 선택한 의사결정 레이블	아니요	예
EL_	모델이 선택한 의사결정에 대한 기대 손실분, $-E(i)$	아니요	예
EM_CCF	평균 신용 비용 요인	아니요	아니요
EM_CLASSIFICATION	사례가 분류된 정규화된 범주	아니요	예
EM_DECISION	모델로 선택한 의사결정 레이블	아니요	예
EM_EVENTPROBABILITY	타겟 이벤트의 사후 확률	아니요	아니요
EM_EXPOSURE	평균 적발치	아니요	아니요

접두어	설명	타겟이 필요합 니까?	의사결 정 행 렬입니 까?
EM_LGD	평균 LGD(Loss Given Default)	아니요	아니요
EM_PD	평균 예측값	아니요	아니요
EM_PREDICTION	예측값	아니요	아니요
EM_PROBABILITY	EM_CLASSIFICATION 변수에 할당된 레벨 에 해당하는 최대 사후 확률	아니요	아니요
EM_PROFITLOSS	예측 이익분 또는 손실분	아니요	예
EM_SEGMENT	조각 변수	아니요	아니요
EP_	모델로 선택한 의사결정에 대한 예측 이익 분, $E(i)$	아니요	예
F_	사례가 포함된 정규화된 범주	예	아니요
G_	관측값이 속한 그룹의 식별 번호	아니요	아니요
GRP_	관측값이 속한 그룹의 식별 번호	아니요	아니요
I_	사례가 분류된 정규화된 범주	아니요	아니요
IC_	투자비용, $IC(i)$	아니요	예
IMP_	결측값 처리된 변수	아니요	아니요
LBL	테이블 및 도표에 사용할 변수에 대한 텍스트 레이블	아니요	아니요
M_	결측값 처리된 관측값을 식별하는 지표 변수 또는 필터링된 관측값을 식별하는 지표 변수	—	아니요
_N	이웃 변수; 이웃 관측값 나열	예	아니요
NODE	의사결정 트리 출력 데이터셋의 각 리프 (leaf)에 대한 숫자 ID	예	아니요
P_	예측값 및 사후 확률	아니요	아니요
R_	일반 잔차, 타겟 - 예측	예	아니요
REPL_	대체 변수	예	아니요

접두어	설명	타겟이 필요합 니까?	의사결 정 행 렬입니 까?
ROI_	투자수익률, $ROI(i)$	예	예
SCORECARD_BIN	각 관측값에 할당된 Bin	아니요	아니요
SCORECARD_POINTS	각 개별 항목에 대한 총 스코어카드 포인트	아니요	아니요
SEGMENT	조각 변수	아니요	아니요
SOM_DIMENSION1	SOM(self-organizing map)의 행 식별	아니요	아니요
SOM_DIMENSION2	SOM의 행 식별	아니요	아니요
SOM_ID	SOM 변수	아니요	아니요
SOM_SEGMENT	SOM으로 생성된 클러스터 식별	아니요	아니요
U_	사례가 분류된 서식화되지 않은 범주	아니요	아니요
WOE_	관측값이 속한 그룹의 식별 번호	아니요	아니요

스코어 출력 데이터셋에는 예측값을 계산하거나 의사결정을 내리는 데 문제가 있음을 나타내는 '_WARN_'이라는 변수도 포함됩니다. _WARN_은 비어 있거나(문제가 없음을 나타냄) 다음과 같은 문자 코드 중 하나 이상을 포함하고 있는 문자 변수입니다.

Code	의미
C	결측 비용 변수
M	결측 입력값
P	올바르지 않은 사후확률(예: <0 또는 >1)
U	인식할 수 없는 입력 범주

비선형회귀

비선형회귀 작업 정보	333
비선형회귀: 분석 역할에 변수 할당	334
비선형회귀: 모델 옵션 설정	334
비선형회귀: 계산 옵션 설정	335
반복 옵션 지정	335
스텝-크기 검색 방법 지정	335
비선형회귀: 도표 생성	336
비선형회귀: 결과 옵션 설정	337
제목 및 각주 지정	338
속성 보기	338

비선형회귀 작업 정보

비선형회귀 작업에서는 비선형모델의 모수에 대해 최소제곱 또는 가중최소제곱추정값을 생성합니다. 비선형모델은 선형모델에 비해 지정 및 추정하기가 어렵습니다. 단순히 회귀변수를 나열하는 것이 아니라 회귀식을 선택하고 모수 이름을 선언한 다음 모수 초기값을 제공해야 합니다. 일부 모델은 적합시키기가 어려워 프로시저가 모델에 정확히 적합하지 못할 수도 있습니다.

실험에 대한 비선형 모델을 생성해야 할 때 이 작업을 사용할 수 있습니다. 예를 들어 특정 효소와 기질쌍에 대해 농도와 성능 간의 관계를 연구한다고 가정합니다. 다른 기질 농도에서 관측된 반응율(속도)을 레코드하고 SAS 데이터셋의 값을 저장합니다. 비선형회귀 작업으로 각 파라미터, 관련 근사 표준오차 및 근사 신뢰구간에 대한 관측값을 계산할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
----------	--------------

사용되는 SAS 프로시저	NLIN
---------------	------

요구 사항 이름	프로시저 및 제품 이름
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	없음

비선형회귀: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
종속변수	종속(반응)변수로 사용할 변수를 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다.
설명변수	독립(설명)변수로 사용할 변수를 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 다른 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.

비선형회귀: 모델 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **모델**을 클릭합니다.

드롭다운 리스트에서 비선형모델식을 선택합니다. 이 식은 예측값을 평가하는 데 사용됩니다.

선택한 모델의 모수 상자에서 a, b, c 모수 형식을 단일 값 또는 값의 범위로 지정합니다. 모수 형식으로 단일 값을 지정하면 모수는 **단일 값** 상자의 값으로 지정됩니다. 모수 형식으로 값의 범위를 지정하면 시작값과 끝값, 증분을 지정할 수도 있습니다.

모수의 제한 또는 경계를 설정하려면 **모수 경계**를 클릭하고 **모수 경계** 창에서 경계를 지정합니다.

주: **모수 경계** 창은 사용자가 입력한 모수 경계의 구문만 검증할 뿐, 경계의 수학적 타당성을 검증하지는 않습니다. 경계가 수학적 의미가 있는지 확인하지 않습니다. 예를 들어 모델이 수렴하지 못하더라도 $2 < b < 1$ 과 같은 경계를 입력할 수 있습니다.

비선형회귀: 계산 옵션 설정

반복 옵션 지정

이 옵션에 액세스하려면 선택 영역에서 **계산**을 클릭합니다.

도표 이름	설명
방법	방법 드롭다운 리스트에서 방법을 선택합니다. Gauss-Newton , Marquardt , Newton 및 Steepest descent (Gradient) 옵션을 사용할 수 있습니다.
최대 반복	작업이 수렴 시도를 중단하기 전에 수행할 반복 횟수를 지정합니다. 이 값은 양의 정수이어야 합니다. 기본 최대 반복 횟수는 100 입니다.
모수 추정에 Moore-Penrose 역의 역 사용	모수 추정에 Moore-Penrose 역을 사용하려면 이 체크 박스를 선택합니다.
평균제곱오차 바꾸기	추정값의 표준오차 계산을 위한 평균제곱오차를 바꾸려면 체크 박스를 선택하고 텍스트 상자에 값을 지정합니다.
최소화 튜닝	모델의 수렴 기준 또는 비정칙성 기준을 텍스트 상자에 지정합니다. 모든 반복법에서 수렴 결정을 위해 Bates와 Watts의 상대 오프셋 수렴 측도가 기본적으로 사용됩니다. 기본값은 1.0E-5 입니다. 특이점 기준은 Hessian 또는 근사를 Hessian으로 반전할 때 허용되는 가장 작은 피벗값의 절대 크기입니다. 기본값은 1.0E-8 입니다.

스텝-크기 검색 방법 지정

이 옵션에 액세스하려면 선택 영역에서 **계산**을 클릭합니다.

도표 이름	설명
스텝-크기 검색 방법	검색 방법을 선택합니다. 다음과 같은 옵션을 사용할 수 있습니다. ■ 스텝-크기 검색 안 함 - 매 반복 시, 목표 값이 감소해야 한다는 제한이 적용되지 않습니다. 지정된 범위 내에 모수 추정값을 유지하고 평가 가능한 관측값 수가 줄어들지 않도록 하기 위해 반씩 스텝은 계속 사용됩니다. 이 옵션은 최소제곱문제를 반복적으로 가중값을 부여하는 데 유용합니다. 기본값입니다. ■ 반씩 스텝 - 이 방법은 최소 오차제곱합(SSE)을 검색합니다. ■ 3차 보간 - 이 방법은 3차 보간을 사용하여 스텝 크기를 추정합니다. 추정된 스텝 크기로 인해 SSE에서 감소가 발생하지 않으면 반씩 스텝이 사용됩니다. ■ Golden 섹션 탐색 - Golden 섹션 탐색으로 스텝 크기를 결정합니다. Tau 모수는 검색할 시작 구간의 길이를 결정하며, Rho 모수는 검색의 정밀도를 지정합니다. Rho 모수는 검색의 정밀도를 지정합니다.
반씩 스텝의 최대 횟수	수행할 반씩 스텝의 최대 횟수를 지정합니다. 이 값은 양의 정수이어야 합니다.

비선형회귀: 도표 생성

선택 영역에서 이 옵션은 **도표** 아래에 표시됩니다. 생성할 도표 유형을 선택합니다. 각 유형의 도표를 생성할 수 있습니다.

주: 표시할 도표 요약 상자에 요청한 도표가 모두 나열됩니다.

도표 이름	설명
예측값	각 종속변수의 관측값 대 예측값에 대한 산점도와 각 종속변수의 관측값 대 각 설명변수(독립변수)에 대한 산점도를 생성할 수 있습니다. 관측값 대 독립변수 도표에서는 모든 예측값 도표에 신뢰한계 또는 예측한계를 표시하도록 선택할 수 있습니다. ■ 신뢰한계를 지정하면 도표에는 평균에 대한 % 신뢰구간이 표시됩니다. ■ 예측한계를 지정하면 도표에는 개별 반응에 대한 예측구간이 표시됩니다. 두 가지 도표에서 alpha는 .05입니다.

도표 이름	설명
잔차	각 종속변수의 예측값 또는 각 설명(독립)변수값에 대해 각 종속변수의 일반, 표준화, 스튜던트화 잔차 도표를 그릴 수 있습니다. 잔차는 다음 공식을 사용하여 계산됩니다. ■ 일반잔차(ordinary residuals)는 관측값과 예측값의 차이입니다. ■ 표준화잔차는 잔차를 표준오차로 나눈 값입니다. ■ 스튜던트화 잔차는 현재 관측값을 제외한 잔차를 표준오차로 나눈 값입니다.
영향력	각 종속변수의 예측값 또는 설명(독립)변수값에 대해 지레값을 도표로 표시할 수 있습니다. 관측값에 대한 지레값은 H 행렬에서 대응되는 대각원소의 값입니다.

비선형회귀: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

통계량이나 모수 추정값을 영구 저장할 수 있습니다. 모수 추정값 저장을 선택하면 반복 횟수가 제한을 초과했을 때 마지막 추정값이 저장되도록 선택할 수도 있습니다. 최대 반복 옵션에서 반복 횟수 제한을 지정할 수 있습니다.

주: SAS Enterprise Guide는 **옵션 창**의 **작업>출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

또한 결과에 반복 히스토리, Hougaard의 왜도 등이 표시되도록 선택할 수 있습니다. Hougaard의 왜도 측도를 통해 모수가 선형에 가까운지 또는 모수에 상당한 정도의 비선형성이 포함되어있는지 평가할 수 있습니다.

주: 이 옵션은 **모든 출력 숨기기** 체크 박스를 선택할 때에는 사용할 수 없습니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

비모수 일원분산분석

비모수 일원분산분석 작업 정보	339
비모수 일원분산분석: 분석 역할에 변수 할당	340
비모수 일원분산분석: 분석 선택	340
비모수 일원분산분석: 정확 p-값 계산	341
비모수 일원분산분석: 결과 옵션 설정	341
제목 및 각주 지정	342
속성 보기	342

비모수 일원분산분석 작업 정보

비모수 일원분산분석 작업은 일원 분류에서의 위치 및 척도 차이에 대한 비모수적 검정으로 구성됩니다. 또한 이 작업은 원 데이터에 대한 표준 분산분석과 경험적 분포함수 기반의 통계량도 제공합니다.

이 작업을 사용하여 대립가설에 대한 귀무가설을 검정할 수 있습니다. 예를 들어 임상 실험에 참여한 류마티스 관절염을 앓고 있는 59명의 여자 환자가 두 개의 그룹인 '활성'과 '위약'으로 지정되었다고 가정합니다. 각 환자의 응답 상태(아주 좋음=5, 좋음=4, 중간=3, 나쁨=2, 아주 나쁨=1)가 기록됩니다. 비모수 일원분산분석 작업을 사용하여 환자의 응답 상태가 두 치료 그룹간에 달라지는 대립가설에 대해 두 치료 그룹 간의 응답 상태에 차이가 없는 귀무가설을 검증할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	NPAR1WAY
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

비모수 일원분산분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
종속변수	종속(반응)변수로 사용할 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
독립변수	독립(설명)변수로 사용할 변수를 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다. 이 리스트의 변수에 의해 정의된 하위그룹에 대해 개별 분석이 수행됩니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 종속변수나 독립변수로 이미 선택한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수 값입니다.

비모수 일원분산분석: 분석 선택

이 옵션에 액세스하려면 선택 영역에서 **분석**을 클릭합니다.

옵션 이름	설명
검정 점수	실행할 분석을 선택합니다.
경험적 분포함수	경험적 분포함수를 수행하려면 경험적 분포함수(EDF) 통계량 계산 체크 박스를 선택합니다.
결측값	체크 박스를 선택하여 결측값을 정식 분류 레벨로 포함할 수 있습니다.
연속성 상관계수	Wilcoxon 점수와 Siegel-Tukey 점수는 작업에서 표준화된 검정 통계량을 계산할 때 연속성 상관계수를 사용합니다. 검정에서 연속성 상관계수를 사용하지 않으려면 연속성 상관계수 생략 체크 박스를 선택합니다.

비모수 일원분산분석: 정확 p-값 계산

이 옵션에 액세스하려면 선택 영역에서 **정확 p-값** 계산을 클릭합니다.

옵션 이름	설명
정확 p-값	기본적으로 작업은 정확 p-값이 아닌 근사값을 계산합니다. 해당 체크 박스를 선택하여 다양한 분석에 대한 정확 p-값을 계산할 수 있습니다. 주: 검정 점수를 선택해야만 정확 p-값 영역의 체크 박스가 활성화됩니다. 검정 점수를 선택하려면 선택 영역에서 분석을 클릭합니다. 모든 정확 p-값 계산을 취소하려면 재설정을 클릭합니다.
계산 시간 제한	계산 시간 제한 영역에서 각 정확 p -값을 계산하기 위한 시간 제한을 지정하거나 시간 제한을 무제한 으로 지정합니다. 정확 p -값 계산에는 시간과 메모리가 많이 소요될 수 있습니다.

비모수 일원분산분석: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
출력 데이터 저장	결과의 일부 또는 전체를 출력 테이블에 저장합니다. 출력에 포함할 통계량의 체크 박스를 선택합니다. 주: 검정 점수를 선택해야만 포함할 통계량 체크 박스가 활성화됩니다. 검정 점수를 선택하려면 선택 영역에서 분석을 클릭합니다. 주: SAS Enterprise Guide는 옵션 창의 작업> 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
리포트 생략	결과를 표시하지 않습니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블

- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 속성 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

불량개수관리도

불량개수관리도 작업 정보	345
불량개수관리도: 분석 역할에 변수 할당	346
관리도: 관리한계 지정	347
관리도: 수행할 검정 선택	348
관리도: 도표 모양 옵션 설정	349
관리도: 분석 결과 저장	349
제목 및 각주 지정	350
속성 보기	350

불량개수관리도 작업 정보

Shewhart 관리도는 공정이 통계적 관리 상태에 있는지 여부를 결정하기 위한 그래픽 및 분석 도구입니다. 불량개수관리도 작업은 하위그룹 표본의 불량(결함) 수에 대한 불량개수관리도를 생성합니다.

이 작업을 사용하여 제조 공정에서의 결점수를 모니터링할 수 있습니다. 예를 들어 전자 회사에서 회로를 500개의 배치로 생산하고 불량개수관리도를 사용하여 부적합한 회로 수를 모니터링한다고 가정합니다. 30개의 배치를 검사하고 각 배치의 고장 수를 계산합니다. 고장 수는 불량개수관리도를 생성하는 데 사용되는 SAS 데이터셋에 저장됩니다. 불량개수관리도의 각 점은 특정 하위 그룹에 있는 부적합 항목의 수를 나타냅니다. 예를 들어 첫 번째 배치에 대해 도표화된 값은 5입니다. 모든 포인트가 관리한계(기본값 3 sigma) 내에 위치하면 공정이 통계적 관리 상태에 있다고 결론을 내릴 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	SHEWHART
필수 SAS 제품	Base SAS, SAS/QC, SAS/GRAPH

요구 사항 이름	프로시저 및 제품 이름
추가 SAS 제품(권장)	SAS/STAT

불량개수관리도: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
결점 수	결점수를 갖는 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
하위그룹 식별자	데이터의 하위그룹을 식별하는 변수를 지정합니다. 변수값은 입력 테이블의 관측값이 논리적 하위그룹으로 배열되는 방법을 나타냅니다. 일반적으로 값은 다음과 같습니다. ■ 하위그룹 표본의 수집 순서를 지정하는 지수 ■ 하위그룹 표본이 수집된 날짜 또는 시간 ■ 하위그룹 표본을 고유하게 식별하는 레이블 하위그룹 식별자는 필수 변수이며 이 역할에는 변수를 하나만 할당할 수 있습니다. 식별된 하위그룹으로 데이터를 정렬하려면 하위그룹별 정렬 체크 박스를 선택합니다.
하위그룹 표본 크기	하위그룹 표본 크기를 변수값으로 지정합니다. 하위그룹 표본 크기는 반드시 입력해야 하며 이 역할에는 하나의 숫자 변수만 할당할 수 있습니다. 하위그룹 표본 크기가 고정되어 있고 크기에 상수값을 지정해야 할 때 하위그룹 표본 크기로 (변수 대신) 숫자값 사용 체크 박스를 선택합니다. 하위그룹 표본 크기 상자에 사용할 값을 입력합니다. 하위 그룹 표본 크기는 프로세스 변수의 최댓값과 99999(포함) 사이의 값이어야 합니다. 주: 옵션 영역에서 하위그룹 표본 크기로 (변수 대신) 숫자값 사용 체크 박스를 선택하면 이 역할은 필요하지 않습니다.

역할 이름	설명
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
블록 변수	데이터를 연속된 하위그룹 블록으로 그룹화하는 변수를 최대 두 개까지 지정합니다. 이 블록은 범례에서 레이블로 표시되며 각 블록 변수는 범례에서 한 종류의 레이블이 됩니다.

관리도: 관리한계 지정

이 옵션에 액세스하려면 선택 영역에서 **관리한계**를 클릭합니다.

옵션 이름	설명
Sigma 한계	관리도에 그려지는 요약 하위그룹 통계량의 표준오차(sigma)의 배수로 관리한계의 너비를 지정합니다. 너비는 양수여야 합니다. 기본 배수는 3입니다.
계산 방법 선택	관리한계의 계산 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 활성 데이터에서 관리한계 계산 ■ 선택한 데이터에서 관리한계 계산 관리한계 또는 관리한계를 계산할 수 있는 모수를 포함하는 데이터셋을 지정하려면 찾아보기를 클릭합니다. 사용할 데이터셋을 선택하고 열기를 클릭합니다. 주: 선택한 데이터셋은 LIMITS= 데이터셋이어야 합니다. LIMITS= 데이터셋에 필요한 변수에 대한 자세한 내용은 생성 중인 그래프의 프로시저 문서를 참조하십시오. 모든 하위 그룹에 대한 요약통계량을 도표화하려면 하위그룹 표본 크기가 명목 관리한계 표본 크기와 동일한지 관계없이 모든 하위그룹에 대한 요약통계량 도표를 생성합니다. 체크 박스를 선택합니다. 거의 모든 하위 그룹의 표본 크기가 공통이고 Nominal 표본 크기 n에 대응하는(가변이 아닌) 고정된 관리한계를 표시하려고 할 때 이 통계량을 도표화할 수 있습니다. 이 옵션을 선택하고 하위 그룹에 대한 표본 크기가 유의하게 다르면 관리한계 해석은 표본 크기가 n인 하위 그룹에 대해서만 의미가 있습니다.

옵션 이름	설명
	표본 크기가 n과 같지 않은 데이터 포인트에 대해 특수 표식을 추가하려면 표본 크기에 해당하는 포인트에 대해 특수 표식을 추가하며, 표본 크기는 고정 관리한계의 명목표본 크기와 동일하지 않습니다. 체크 박스를 선택합니다. ■ 관리한계 지정 상한, 중심 및 하한 필드에 관리상한, 중심선 및 관리하한을 각각 지정합니다. 하위그룹의 표본 크기가 다양하면, 명목표본 크기를 지정해야 합니다. 주: 그룹 분석변수 역할에 변수를 할당하면 관리한계 값을 입력할 수 없습니다.

관리도: 수행할 검정 선택

이 옵션에 액세스하려면 선택 영역에서 **검정**을 클릭합니다.

주: 이 검정은 3 Sigma 한계가 아니거나 하위그룹 표본 크기에 따라 달라지는 관리한계에는 적용되지 않습니다.

옵션 이름	설명
검정 선택	체크 박스를 사용하여 런 검정, 패턴 검정 및 Western Electric 규칙이라고도 하는 특수 원인에 대한 검정을 하나 이상 요청합니다. 이러한 검정은 주관리도에 그려진 포인트에서 특정 비임의 패턴을 찾아냅니다. 이러한 비임의 패턴의 발생을 신호라고 합니다. 이러한 검정을 통해 특수한 변동의 원인이 존재하는지를 알 수 있습니다. 각 검정을 선택하면 해당 검정에 대한 설명이 설명 상자에 나타납니다.
레이블	텍스트 상자를 사용하여 특수 원인에 대한 검정이 양성으로 나타난 포인트에 대한 레이블을 입력합니다. 레이블의 문자 수는 16자를 초과할 수 없습니다. 각 레이블은 검정에서 신호가 발생한 모든 점에 표시됩니다.
레이블 색상	■ 드롭다운 리스트를 사용하여 지정한 레이블의 색상을 선택합니다. ■ 영역 선 표시 체크 박스를 사용하여 주관리도에 A, B, C 영역을 구분하는 선을 표시합니다. 이 영역은 관리한계 사이의 간격을 여섯 개의 동일한 간격을 지닌 하위 간격으로 분할함으로써 구성됩니다.

옵션 이름	설명
	■ 3 sigma 한계 무시 체크 박스를 사용하여 기본 배수 3 이외의 sigma 관리한계를 지정했을 때 특수한 원인을 찾는 검정을 수행할 수 있습니다. sigma 한계를 변경하려면 선택 영역에서 관리한계를 클릭합니다. ■ 모든 점을 검정의 시작점으로 사용 체크 박스를 사용하여 중복되는 점 패턴에 특수 원인 검정을 적용합니다.

관리도: 도표 모양 옵션 설정

일부 모양 옵션은 여러 관리도 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- “축 옵션 설정” (605페이지)
- “참조선 추가” (605페이지)
- “기타 도표 옵션 설정” (606페이지)
- “블록 변수 옵션 설정” (608페이지)

관리도: 분석 결과 저장

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

분석 결과를 저장하기 위해 다양한 통계량을 포함하는 출력 데이터셋을 생성할 수 있습니다. 영구적으로 저장하려는 출력 데이터의 체크 박스를 선택합니다.

역할 이름	설명
하위그룹 통계량과 관리한계 데이터셋	관리도에 그려지는 정보를 포함하는 출력 데이터셋을 생성합니다. 여기에는 하위그룹 식별자값 및 해당 요약통계량과 관리한계가 포함됩니다. 주: 이 옵션은 OUTTABLE= 데이터셋을 생성합니다. SHEWHART 프로시저에서 TABLE= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.
하위그룹 통계량 출력 데이터셋	하위그룹 요약통계량을 포함하는 출력 데이터셋을 생성합니다.

역할 이름	설명
	주: 이 옵션은 OUTHISTORY= 데이터셋을 생성합니다. SHEWHART 프로시저에서 HISTORY= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.
관리한계 출력 데이터셋	관리한계를 저장하는 출력 데이터셋을 생성합니다. 추후에 이 작업 사용 시 관리한계의 계산 방법을 선택할 때 이 데이터셋을 사용할 수 있습니다.

주: SAS Enterprise Guide는 **옵션 창**의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 속성 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

일원분산분석

일원분산분석 작업 정보	353
일원분산분석: 분석 역할에 변수 할당	354
일원분산분석: 검정 선택	354
일원분산분석: 그룹 평균 비교	355
일원분산분석: 정량적 변수에 대한 통계량 선택	356
일원분산분석: 도표 생성	356
일원분산분석: 결과 옵션 설정	357
제목 및 각주 지정	358
속성 보기	358

일원분산분석 작업 정보

일원분산분석(ANOVA)에서는 둘 이상의 처리 레벨을 갖는 하나의 처리 인수를 다룹니다. 일원분산분석의 목적은 각 레벨 평균 간의 차이를 검정하고 이러한 차이의 값을 구하는 데 있습니다. 처리 레벨이 두 개이면, 이 분석은 두 그룹의 평균을 비교하는 t 검정과 동일합니다.

ANOVA 작업을 사용하여 다음 작업을 수행할 수 있습니다.

- 붉은색 토끼풀 식물의 질소 함유에 영향을 미치는 박테리아 연구. 요인은 박테리아 품종이며 여섯 개의 레벨로 구성됩니다.
- 랜덤화 완비 블록 설계 분석. 예를 들어 처리에 대하여 세 가지 서로 다른 유형이 특정 작물의 산출 및 가치에 영향을 미치는지에 관심이 있다고 가정합니다. 실험 단위가 동질적이지 않으므로 블록 요인을 도입하여 각 블록 내에서 실험 단위가 동질이 되도록 합니다. 처리에 대한 세 가지 서로 다른 유형은 각 블록에서 임의로 할당됩니다.
- 세 가지 브랜드 전지의 수명 비교. 요인은 브랜드이며 세 개의 레벨로 구성됩니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	ANOVA
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

일원분산분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
종속변수	종속(반응)변수로 사용할 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
독립변수	독립(설명)변수로 사용할 변수를 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 종속변수나 독립변수로 이미 선택한 변수로는 분석을 그룹화할 수 없습니다.

일원분산분석: 검정 선택

이 옵션에 액세스하려면 선택 영역에서 **검정**을 클릭합니다.

독립변수 역할에 할당된 변수가 두 개 이상의 개별값을 가질 때 일원분산분석 검정을 사용할 수 있습니다. 등분산 가정에 대해 모델을 검정할 수 있습니다.

옵션 이름	설명
Welch의 분산 가중 ANOVA	가중 분산을 사용하여 그룹 평균을 검정합니다. 등분산 가정이 기각되면, 이 검정을 사용할 수 있습니다.
등분산 검정	다음 검정 중 하나를 선택합니다. ■ Bartlett 검정은 정규이론 우도비 검정을 수정한 것입니다. 이 검정은 데이터가 정규분포를 따를 때 제1종 오류 비율을 정확하게 계산합니다. ■ Brown-Forsythe 검정은 Levene 검정의 변형입니다. 등분산성은 그룹 중위수의 절대편차를 사용하여 결정됩니다. 이 검정은 분산의 차이를 알아내는 데 유용하지만 데이터가 여러 개의 큰 그룹으로 되어 있으면, 계산이 복잡하여 많은 컴퓨팅 리소스를 필요로 합니다. ■ Levene의 검정 - 분산의 동질성을 검사하는 표준 검정입니다. 이 검정은 잔차제곱을 계산하여 등분산성을 결정합니다.

일원분산분석: 그룹 평균 비교

이 옵션에 액세스하려면 선택 영역의 평균 머리글 아래에서 **평균 비교**를 클릭합니다.

출력 데이터셋에 사용할 주효과(독립변수)에 대한 다중비교검정을 선택합니다. **신뢰수준** 드롭다운 리스트를 사용하여 선택한 모든 검정에 대한 신뢰수준을 선택합니다. 기본값은 95% 신뢰수준입니다.

다음과 같은 검정을 사용할 수 있습니다.

- Bonferroni t 검정
- Tukey의 스튜던트화 범위검정(HSD)
- Duncan의 다중범위검정
- Dunnett의 t 검정
- Fisher의 최소 유의차 검정
- Gabriel의 다중비교절차
- Student-Newman-Keuls 다중범위검정
- Waller-Duncan k-ratio 검정
- Scheffe의 다중비교절차
- Ryan-Einot-Gabriel-Welsch 다중범위검정

일원분산분석: 정량적 변수에 대한 통계량 선택

이 옵션에 액세스하려면 선택 영역의 평균 머리글 아래에서 **통계량 계산**을 클릭합니다.
독립변수의 각 레벨에서 종속변수에 대해 계산 및 표시할 기술통계량을 선택합니다.

옵션 이름	설명
평균	표본변수값을 더한 후, 그 값을 관측값 개수로 나누어 계산한 산술평균입니다.
표준편차	데이터 값 그룹의 변동성에 대한 통계적 측도입니다. 이 측도는 빈도분석에서 산포의 측도로 가장 많이 사용되며, 분산에 대한 양의 제곱근과 같습니다.
표준오차	표본평균의 표준편차입니다. 표준오차는 표본크기의 제곱근에 대해 표본 표준편차 비율로 정의됩니다.
분산	데이터 값의 산포에 대한 통계적인 측도입니다. 이 측도는 각 관측값과 표본평균 차의 제곱에 대한 평균입니다.
비결측 관측값 수	결측값이 없는 관측값의 수입니다.
결측값 수	값이 입력되지 않은 관측값의 개수입니다.
최솟값	관측값의 최솟값입니다.
최댓값	관측값의 최댓값입니다.

일원분산분석: 도표 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.
체크 박스를 사용하여 상자 도표, 평균 도표 또는 두 가지 모두를 생성할 수 있습니다.

도표 이름	설명
상자수염도	상자 도표는 그룹 관측값의 평균, 사분위수, 최댓값 및 최솟값을 표시합니다. 이 도표는 다음 통계량을 나타냅니다. ■ 위쪽 수염의 끝점은 최댓값을 나타냅니다. ■ 상자의 맨 위는 제3 사분위수(또는 75번째 백분위수)를 나타냅니다. ■ 상자의 길이는 사분위범위(제3 사분위수와 제1 사분위수 사이의 거리)를 나타냅니다. ■ 상자 안의 점은 평균을 나타냅니다. ■ 상자 안의 수평선은 중위수(또는 50번째 백분위수)를 나타냅니다. ■ 상자의 맨 아래는 제1 사분위수(또는 25번째 백분위수)를 나타냅니다. ■ 아래쪽 수염의 끝점은 최솟값을 나타냅니다.
평균	독립변수에 대한 평균 및 표준오차를 표시한 평균 도표입니다. 그래프의 각 점은 분산 막대에 둘러싸여 있습니다. 평균을 선택하면 다음과 같은 옵션도 선택할 수 있습니다. ■ 평균 또는 표준편차의 표준오차를 사용하여 막대를 그릴지를 선택합니다. 평균의 표준오차는 평균의 표준편차와 동일합니다. 표준편차는 표본분산의 제곱근입니다. ■ 막대 길이를 계산할 때 합동분산을 사용할지를 선택합니다. 합동분산은 모집단이 동일한 두 독립표본의 결합분산입니다. ■ 막대의 높이를 1, 2 또는 3 표준 단위로 선택합니다. ■ 세로축의 시작점을 0으로 할 것인지를 선택합니다.

일원분산분석: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
통계량 출력 데이터 저장	결과를 출력 데이터셋에 저장합니다. 출력 데이터셋에는 모델의 각 효과에 대한 제곱합, 자유도, F 통계량 및 p-값이 포함됩니다. 주: SAS Enterprise Guide는 옵션 창의 작업 > 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.

옵션 이름	설명
모든 출력 숨기기	결과를 표시하지 않습니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

일원빈도분석

일원빈도분석 작업 정보	361
일원빈도분석: 분석 역할에 변수 할당	362
일원빈도분석: 통계량 옵션 설정	362
일원빈도분석: 도표 생성	363
일원빈도분석: 결과 옵션 설정	363
제목 및 각주 지정	364
속성 보기	365

일원빈도분석 작업 정보

일원빈도분석 작업을 사용하여 데이터에서 빈도 테이블을 생성할 수 있으며 이항 및 카이제곱 검정을 수행할 수도 있습니다.

이 작업을 사용하여 새로운 약품의 효능을 분석할 수 있습니다. 예를 들어 의학 연구원 집단에서 피부 상태에 따른 새로운 치료의 효능 평가에 대해 관심이 있다고 가정합니다. 참여 클리닉의 피부과 전문의가 연구를 수행하고 상태를 평가할 수 있도록 교육을 받았습니다. 훈련 후, 두 피부과 전문의는 준비 조사에서의 피부 상태를 만족하는 환자를 검사하고 평가합니다. 일원빈도분석 작업은 진단에 대한 일치도를 평가하는 데 사용됩니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	FREQ
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

일원빈도분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분석변수	분석할 변수를 지정합니다. 이 역할에 지정한 각 변수에 대해 일원빈도 테이블이 생성됩니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 다른 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.

일원빈도분석: 통계량 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **통계량**을 클릭합니다.

옵션 이름	설명
빈도 테이블 옵션	일원빈도 테이블에 포함할 통계량을 선택합니다. 다음과 같은 옵션을 사용할 수 있습니다. ■ 누적빈도와 누적백분율 - 분석변수의 각 값에 대한 빈도, 누적빈도, 전체 빈도의 백분율 및 누적백분율이 포함된 테이블을 생성합니다. ■ 빈도와 누적빈도 - 분석변수의 각 값에 대한 빈도와 누적빈도가 포함된 테이블을 생성합니다.

옵션 이름	설명
	■ 빈도와 백분율 - 분석변수의 각 값에 대한 빈도와 전체 빈도의 백분율이 포함된 테이블을 생성합니다. ■ 빈도만 - 분석변수의 각 값에 대한 빈도만 포함된 테이블을 생성합니다.
결측값	결측값을 빈도 테이블에 포함하려면 빈도 표시 체크 박스를 선택합니다. 결측값 빈도를 이항 또는 카이제곱 검정에 포함하려면 계산에 포함 체크 박스를 선택합니다.
이항비	근사 검정, 정확 p -값 계산 또는 두 가지를 모두 선택합니다. 이항비 검정에서는 검정 비율(귀무가설 비율 값) 및 신뢰수준을 지정합니다.
카이제곱 적합도 검정	근사 검정, 정확 p -값 계산 또는 두 가지를 모두 선택합니다.
정확 계산	정확 p-값 계산에는 시간과 메모리가 많이 소요될 수 있습니다. 정확 계산 영역에서 계산 시간 제한 체크 박스를 선택하고 각 교차 테이블의 p-값 계산에 대한 시간 제한을 초 단위로 지정합니다. 기본값은 900초입니다. 정확 p-값을 직접 계산하지 않고 정확 p-값의 Monte Carlo 추정값을 계산하려면 Monte Carlo 추정 사용 체크 박스를 선택합니다. Monte Carlo 추정은 정확한 계산을 위해 상당한 시간과 메모리를 요구하는 복잡한 문제에 유용합니다.

일원빈도분석: 도표 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

체크 박스를 사용하여 빈도 막대 그래프를 표시할 수 있습니다. 수평 막대 그래프, 수직 막대 그래프 또는 두 가지 모두를 표시할 수 있습니다.

일원빈도분석: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
출력 데이터 저장	빈도 테이블을 출력 데이터셋에 저장하려면 빈도 및 백분율을 포함한 데이터셋 생성 체크 박스를 선택합니다. 출력 데이터에 일원 테이블의 누적빈도와 누적백분율을 포함하려면 누적 빈도 및 백분율 데이터셋에 추가 체크 박스를 선택합니다. 데이터셋에 포함하려는 분석변수를 선택한 다음, 데이터셋을 저장하려는 위치를 지정합니다. 데이터셋의 경로는 <code>\\서버\라이브러리\멤버</code> 형식이어야 합니다. 찾아보기를 클릭하여 위치를 선택할 수도 있습니다. 이항비 또는 카이제곱 검정을 선택했으면, 데이터셋에 이항비와 카이제곱 통계량 추가 체크 박스를 선택하여 출력 데이터셋에 빈도 테이블을 영구적으로 저장할 수 있습니다. 데이터셋을 저장하려는 위치를 지정합니다. 데이터셋의 경로는 <code>\\서버\라이브러리\멤버</code> 형식이어야 합니다. 찾아보기를 클릭하여 위치를 선택할 수도 있습니다. 주: 이항비와 카이제곱 검정통계량은 분석변수 역할의 마지막 변수에 대해서만 영구적으로 저장할 수 있습니다.
모든 출력 숨기기	결과를 표시하지 않습니다. 주: 작업을 실행하려면 출력을 표시하거나 결과를 데이터셋에 저장해야 합니다. 빈도 및 백분율을 포함한 데이터셋 생성 체크 박스의 선택을 취소하고 모든 출력 및 도표 숨기기 체크 박스를 선택하면 작업을 실행하거나 저장할 수 없습니다.
출력 데이터의 순서 선택	테이블의 값 순서를 지정하려면 출력 데이터 순서 드롭다운 리스트에서 출력 데이터 순서를 선택합니다. 다음과 같은 옵션을 사용할 수 있습니다. ■ 데이터셋 순서 - 입력 데이터셋의 순서에 따라 값을 정렬합니다. ■ 서식화된 값 - 서식화된 값의 오름차순으로 값을 정렬합니다. ■ 빈도 내림차순 - 빈도변수의 내림차순으로 값을 정렬합니다. ■ 서식화되지 않은 값 - 서식화되지 않은 값의 오름차순으로 값을 정렬합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.

- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

불량률관리도

불량률관리도 정보	367
불량률관리도: 분석 역할에 변수 할당	368
관리도: 관리한계 지정	369
관리도: 수행할 검정 선택	370
관리도: 도표 모양 옵션 설정	371
관리도: 분석 결과 저장	371
제목 및 각주 지정	372
속성 보기	372

불량률관리도 정보

Shewhart 관리도는 공정이 통계적 관리 상태에 있는지 여부를 결정하기 위한 그래픽 및 분석 도구입니다. 불량률관리도 작업은 하위그룹 표본에 존재하는 불량품의 비율에 대한 불량률관리도를 생성합니다.

이 작업을 사용하여 제조 공정에서의 결점 비율을 모니터링할 수 있습니다. 예를 들어 전자 회사에서 회로를 500개의 배치로 생산하고 불량률관리도를 사용하여 부적합한 회로 비율을 모니터링한다고 가정합니다. 30개의 배치를 검사하고 각 배치의 고장 수를 계산합니다. 고장 수는 불량률관리도를 생성하는 데 사용되는 SAS 데이터셋에 저장됩니다. 불량률관리도의 각 점은 특정 하위 그룹에서 부적합 항목의 비율을 나타냅니다. 예를 들어 첫 번째 순환의 고장 수가 5이면 첫 번째 배치에 대해 도표화된 값은 $5/500 = 0.01$ 입니다. 모든 포인트가 관리한계(기본값 3 sigma) 내에 위치하면 공정이 통계적 관리 상태에 있다고 결론을 내릴 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
----------	--------------

사용되는 SAS 프로시저	SHEWHART
---------------	----------

요구 사항 이름	프로시저 및 제품 이름
필수 SAS 제품	Base SAS, SAS/QC, SAS/GRAPH
추가 SAS 제품(권장)	SAS/STAT

불량률관리도: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
불량 비율	결점수를 갖는 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
하위그룹 식별자	데이터의 하위그룹을 식별하는 변수를 지정합니다. 변수값은 입력 테이블의 관측값이 논리적 하위그룹으로 배열되는 방법을 나타냅니다. 일반적으로 값은 다음과 같습니다. ■ 하위그룹 표본의 수집 순서를 지정하는 지수 ■ 하위그룹 표본이 수집된 날짜 또는 시간 ■ 하위그룹 표본을 고유하게 식별하는 레이블 하위그룹 식별자는 필수 변수이며 이 역할에는 변수를 하나만 할당할 수 있습니다. 식별된 하위그룹으로 데이터를 정렬하려면 하위그룹별 정렬 체크 박스를 선택합니다.
하위그룹 표본 크기	하위그룹 표본 크기를 변수값으로 지정합니다. 하위그룹 표본 크기는 반드시 입력해야 하며 이 역할에는 하나의 숫자 변수만 할당할 수 있습니다. 하위그룹 표본 크기가 고정되어 있고 크기에 상수값을 지정해야 할 때 하위그룹 표본 크기로 (변수 대신) 숫자값 사용 체크 박스를 선택합니다. 하위그룹 표본 크기 상자에 사용할 값을 입력합니다. 하위그룹 표본 크기는 프로세스 변수의 최댓값과 99999(포함) 사이의 값이어야 합니다. 주: 옵션 영역에서 하위그룹 표본 크기로 (변수 대신) 숫자값 사용 체크 박스를 선택하면 이 역할은 필요하지 않습니다.

역할 이름	설명
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
블록 변수	데이터를 연속된 하위그룹 블록으로 그룹화하는 변수를 최대 두 개 까지 지정합니다. 이 블록은 범례에서 레이블로 표시되며 각 블록 변수는 범례에서 한 종류의 레이블이 됩니다.

관리도: 관리한계 지정

이 옵션에 액세스하려면 선택 영역에서 **관리한계**를 클릭합니다.

옵션 이름	설명
Sigma 한계	관리도에 그려지는 요약 하위그룹 통계량의 표준오차(sigma)의 배수로 관리한계의 너비를 지정합니다. 너비는 양수여야 합니다. 기본 배수는 3입니다.
계산 방법 선택	관리한계의 계산 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 활성 데이터에서 관리한계 계산 ■ 선택한 데이터에서 관리한계 계산 관리한계 또는 관리한계를 계산할 수 있는 모수를 포함하는 데이터셋을 지정하려면 찾아보기를 클릭합니다. 사용할 데이터셋을 선택하고 열기를 클릭합니다. 주: 선택한 데이터셋은 LIMITS= 데이터셋이어야 합니다. LIMITS= 데이터셋에 필요한 변수에 대한 자세한 내용은 생성 중인 그래프의 프로시저 문서를 참조하십시오. 모든 하위 그룹에 대한 요약통계량을 도표화하려면 하위그룹 표본 크기가 명목 관리한계 표본 크기와 동일한지 관계없이 모든 하위그룹에 대한 요약통계량 도표를 생성합니다. 체크 박스를 선택합니다. 거의 모든 하위 그룹의 표본 크기가 공통이고 Nominal 표본 크기 n에 대응하는(가변이 아닌) 고정된 관리한계를 표시하려고 할 때 이 통계량을 도표화할 수 있습니다. 이 옵션을 선택하고 하위 그룹에 대한 표본 크

옵션 이름	설명
	기가 유의하게 다르면 관리한계 해석은 표본 크기가 n인 하위 그룹에 대해서만 의미가 있습니다. 표본 크기가 n과 같지 않은 데이터 포인트에 대해 특수 표식을 추가하려면 표본 크기에 해당하는 포인트에 대해 특수 표식을 추가하며, 표본 크기는 고정 관리한계의 명목표본 크기와 동일하지 않습니다. 체크 박스를 선택합니다. ■ 관리한계 지정 상한, 중심 및 하한 필드에 관리상한, 중심선 및 관리하한을 각각 지정합니다. 하위그룹의 표본 크기가 다양하면, 명목표본 크기를 지정해야 합니다. 주: 그룹 분석변수 역할에 변수를 할당하면 관리한계 값을 입력할 수 없습니다.

관리도: 수행할 검정 선택

이 옵션에 액세스하려면 선택 영역에서 **검정**을 클릭합니다.

주: 이 검정은 3 Sigma 한계가 아니거나 하위그룹 표본 크기에 따라 달라지는 관리한계에는 적용되지 않습니다.

옵션 이름	설명
검정 선택	체크 박스를 사용하여 런 검정, 패턴 검정 및 Western Electric 규칙이라고도 하는 특수 원인에 대한 검정을 하나 이상 요청합니다. 이러한 검정은 주관리도에 그려진 포인트에서 특정 비임의 패턴을 찾아냅니다. 이러한 비임의 패턴의 발생을 신호라고 합니다. 이러한 검정을 통해 특수한 변동의 원인이 존재하는지를 알 수 있습니다. 각 검정을 선택하면 해당 검정에 대한 설명이 설명 상자에 나타납니다.
레이블	텍스트 상자를 사용하여 특수 원인에 대한 검정이 양성으로 나타난 포인트에 대한 레이블을 입력합니다. 레이블의 문자 수는 16자를 초과할 수 없습니다. 각 레이블은 검정에서 신호가 발생한 모든 점에 표시됩니다.
레이블 색상	■ 드롭다운 리스트를 사용하여 지정한 레이블의 색상을 선택합니다. ■ 영역 선 표시 체크 박스를 사용하여 주관리도에 A, B, C 영역을 구분하는 선을 표시합니다. 이 영역은 관리한계 사이의

옵션 이름	설명
	간격을 여섯 개의 동일한 간격을 지닌 하위 간격으로 분할함으로써 구성됩니다.
	■ 3 sigma 한계 무시 체크 박스를 사용하여 기본 배수 3 이외의 sigma 관리한계를 지정했을 때 특수한 원인을 찾는 검정을 수행할 수 있습니다. sigma 한계를 변경하려면 선택 영역에서 관리한계를 클릭합니다. ■ 모든 점을 검정의 시작점으로 사용 체크 박스를 사용하여 중복되는 점 패턴에 특수 원인 검정을 적용합니다.

관리도: 도표 모양 옵션 설정

일부 모양 옵션은 여러 관리도 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- “축 옵션 설정” (605페이지)
- “참조선 추가” (605페이지)
- “기타 도표 옵션 설정” (606페이지)
- “블록 변수 옵션 설정” (608페이지)

관리도: 분석 결과 저장

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

분석 결과를 저장하기 위해 다양한 통계량을 포함하는 출력 데이터셋을 생성할 수 있습니다. 영구적으로 저장하려는 출력 데이터의 체크 박스를 선택합니다.

역할 이름	설명
하위그룹 통계량과 관리 한계 데이터셋	관리도에 그려지는 정보를 포함하는 출력 데이터셋을 생성합니다. 여기에는 하위그룹 식별자값 및 해당 요약통계량과 관리한계가 포함됩니다. 주: 이 옵션은 OUTTABLE= 데이터셋을 생성합니다. SHEWHART 프로시저에서 TABLE= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.

역할 이름	설명
하위그룹 통계량 출력 데이터셋	하위그룹 요약통계량을 포함하는 출력 데이터셋을 생성합니다. 주: 이 옵션은 OUTHISTORY= 데이터셋을 생성합니다. SHEWHART 프로시저에서 HISTORY= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.
관리한계 출력 데이터셋	관리한계를 저장하는 출력 데이터셋을 생성합니다. 추후에 이 작업 사용 시 관리한계의 계산 방법을 선택할 때 이 데이터셋을 사용할 수 있습니다.

주: SAS Enterprise Guide는 **옵션 창**의 **작업 > 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

파레토 그래프

파레토 그래프 작업 정보	375
파레토 그래프: 분석 역할에 변수 할당	376
파레토 그래프: 범주 개수 제한	377
파레토 그래프: 도표 모양 옵션 설정	378
막대 옵션 설정	378
파레토 그래프: 축 옵션 설정	378
파레토 그래프: 참조선 추가	379
파레토 그래프: 인셋 추가	379
기타 도표 옵션 설정	380
제목 및 각주 지정	380
속성 보기	381

파레토 그래프 작업 정보

파레토 그래프는 미국과 일본의 산업계에서 품질 향상을 위해 사용하고 있는 7가지 기본 도구 중 하나입니다. 파레토 그래프는 공정의 상대적인 문제 발생 빈도를 크기의 역순에 따라 막대(왼쪽에서 오른쪽으로)로 표시합니다. 수평 파레토 그래프는 이러한 정보를 길이의 역순에 따라(위쪽에서 아래쪽) 막대로 표시합니다.

파레토 그래프는 문제의 작은 부분인 중요한 소수가 나머지 사소한 다수보다 훨씬 빈번하게 발생한다는 파레토 원칙을 적용한 도구입니다. 파레토 그래프는 가장 중요한 문제로서 해결 효과가 가장 큰 문제를 식별하여 품질 향상 작업의 우선순위를 정하는 데 도움을 줍니다.

파레토 그래프 작업을 사용하여 이벤트 전후의 결과를 비교할 수 있는 비교 파레토 그래프를 생성할 수 있습니다. 예를 들어 산화금속 반도체 축전기 공정 과정에서 불량 원인이 확산로의 튜브를 청소한 전후로 기록됩니다. 출력에서 두 개의 파레토 그래프가 세로로 배열되어 청소 프로세스가 불량 수를 늘이거나 줄이는지를 결정할 수 있습니다.

요구 사항 이름	프로시저 및 제품 이름
----------	--------------

사용되는 SAS 프로시저 PARETO

요구 사항 이름	프로시저 및 제품 이름
필수 SAS 제품	Base SAS, SAS/QC, SAS/GRAPH
추가 SAS 제품(권장)	SAS/STAT

파레토 그래프: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
공정변수	분석할 공정변수를 지정합니다. 이 변수는 그래프의 파레토 범주를 결정합니다. 공정변수는 필수 변수이며 이 역할에는 단 하나의 변수만 할당할 수 있습니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
상대 가중값 변수	상대 가중값 변수로 사용할 변수를 지정합니다. 이 역할에 변수를 지정하면 막대의 높이는 범주의 가중 빈도를 나타내게 됩니다. 상대 가중값 변수는 발생 빈도가 아닌 비용 기반의 파레토 분석을 수행하는 데 유용합니다. 이 역할에는 변수를 하나만 할당할 수 있습니다.
분류변수	분류변수로 사용할 하나 이상의 변수를 지정합니다. 이 역할에 변수를 지정하면 여러 분류 레벨의 빈도를 비교할 수 있는 비교 파레토 그래프가 생성됩니다.

파레토 그래프: 범주 개수 제한

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

옵션 이름	설명
범주 표시	그래프에 표시되는 범주의 수를 제한할 수 있습니다. 기본적으로 공정의 모든 범주는 파레토 그래프의 막대로 표시됩니다. 표시되는 범주의 수를 수정하려면 다음 옵션 중 하나를 선택합니다. ■ 큰 값을 가지는 N개 - n개의 큰 값을 갖는 파레토 범주만 표시되는 그래프를 생성합니다. 기본 n 값은 5입니다. ■ 전체 빈도의 N퍼센트까지 - 누적백분율이 n퍼센트와 같거나 작은 파레토 범주가 표시되는 그래프를 생성합니다. 기본 n 값은 90입니다. ■ 전체 빈도의 최소 N퍼센트 - 백분율이 n퍼센트와 같거나 큰 파레토 범주가 표시되는 그래프를 생성합니다. 기본 n 값은 5입니다. 옵션을 선택한 후 N 상자에 n 값을 입력합니다. 제외할 모든 범주를 포괄하는 새로운 범주를 생성하려면 위에 포함되지 않는 범주 표시 체크 박스를 선택합니다. 이 범주의 기본 레이블은 기타입니다. 레이블 상자에서 다른 범주 레이블(32자 이하)을 지정할 수 있습니다.
공정변수의 결측값 포함	결측값을 범주로 취급하려면 이 체크 박스를 선택합니다. 이 범주는 파레토 그래프에 막대로 표시됩니다. 공정변수가 문자 변수이면 결측값은 공백의 내부값(변수의 입력값)으로 정의됩니다. 공정변수가 숫자이면 결측값은 SAS 결측값 중 하나로 정의됩니다. 기본적으로 결측값은 분석에서 제외됩니다.
데이터셋에 결과 저장	출력 데이터셋에 결과를 저장하려면 이 체크 박스를 선택합니다. 주: SAS Enterprise Guide는 옵션 창의 작업>출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.

파레토 그래프: 도표 모양 옵션 설정

막대 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **막대**를 클릭합니다.

옵션 이름	설명
막대 방향 및 모양	막대를 수직으로 표시하려면 수직 을 선택합니다. 기본 방향은 세로입니다. 막대를 수평으로 표시하려면 수평 을 선택합니다. 막대의 색상을 변경하려면 막대 드롭다운 리스트를 사용합니다. 막대 외곽선의 색상을 변경하려면 외곽선 드롭다운 리스트를 사용합니다.
막대 레이블	막대에 레이블을 표시할 수 있습니다. 막대 높이 를 선택하여 각 막대의 높이에 레이블을 지정합니다. 누적 백분율 을 선택하여 각 막대의 누적 백분율에 레이블을 지정합니다. 레이블은 기본적으로 표시되지 않습니다.
곡선 표시	파레토 그래프에는 기본적으로 누적백분율 곡선이 표시됩니다. 곡선의 색상을 변경하려면 곡선 색상 드롭다운 리스트를 사용합니다. 곡선의 점에 값을 레이블로 표시하려면 값을 레이블로 사용 체크 박스를 선택합니다. 이 곡선과 두 번째 축이 나타나지 않게 하려면 곡선 표시 체크 박스의 선택을 취소합니다.
막대 강조 표시	가장 높음 체크 박스를 선택하여 가장 높은 n 값에 대한 막대를 강조 표시합니다. 최솟값 을 선택하여 가장 낮은 n 값에 대한 막대를 강조 표시합니다. 강조 표시 색상을 변경하려면 색상 드롭다운 리스트를 사용합니다. 두 개 이상의 막대를 강조 표시하려면 막대 개수 상자에서 위쪽/아래쪽 화살표를 사용하여 n 값을 지정합니다. n 값은 프로세스 범주의 수와 같거나 작아야 합니다.

파레토 그래프: 축 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **축**을 클릭합니다.

막대의 방향에 따라 축 옵션이 기본 세로축(Y)에 적용될지 또는 기본 가로축(X)에 적용될지가 결정됩니다.

다음 옵션을 설정할 수 있습니다.

- 축 색상 및 너비
- 기본 축에 대한 레이블
- 파레토 그래프 및 배경 채우기 색상 주위에 프레임이 그려집니다.
- 기본 축에 대한 척도 다음 옵션 중에서 선택할 수 있습니다.
 - **백분율** - 비율이 총 빈도의 백분율로 표시됩니다. 상대 가중값 변수를 할당하면 비율은 총 가중값의 백분율이 됩니다. 기본값입니다.
 - **개수** - 비율이 빈도 단위(개수)로 표시됩니다. 상대 가중값을 할당하면 이 옵션을 사용할 수 없습니다.
 - **가중값** - 비율이 상대 가중값 변수와 같은 단위로 표시됩니다. 이 옵션은 상대 가중값 변수를 할당할 때만 사용할 수 있습니다.

파레토 그래프: 참조선 추가

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **참조선**을 클릭합니다.

참조선은 막대의 방향에 따라 세로축 또는 가로축을 따라 나타납니다. 참조선을 사용하려면 **참조선 사용** 체크 박스를 선택합니다.

파레토 그래프: 인셋 추가

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **인셋**을 클릭합니다.

요약통계량의 상자 또는 테이블(인셋이라고 함)을 그래프에 추가하여 파레토 그래프를 보강할 수 있습니다. 비교 파레토 그래프를 작성하면 각 성분 그래프에 인셋이 표시됩니다.

주: 파레토 그래프에 인셋을 추가하면 **모양 > 옵션** 패널에서 **표본 크기 범례 표시** 체크 박스의 선택을 취소해야 합니다. 파레토 그래프에 인셋과 표본 크기 범례가 포함되면 인셋 및 범례에 대한 상자가 출력에서 겹칠 수 있습니다. 따라서 인셋만 포함하거나 아니면 표본 크기 범례만 표시하는 것이 좋습니다.

인셋 포함 체크 박스를 선택하여 요약통계량을 파레토 그래프에 추가합니다. 인셋에 포함할 통계량을 선택합니다. 이전에 선택한 옵션에 따라 사용할 수 있는 통계량이 다릅니다.

- **표본 크기**는 인셋에 기본적으로 표시되는 유일한 통계량입니다.
- **상대 가중값 변수** 역할에 변수를 할당하면 **가중값 합**을 선택할 수 있습니다.
- 범주의 수를 제한할 때 다음과 통계량을 사용할 수 있습니다.
 - **제외된 관측값** - 제외된 범주를 나타내는 **기타** 범주가 없을 때 사용할 수 있습니다.
 - **OTHER= 내 관측값 수** - 제외된 범주를 나타내는 **기타** 범주가 있을 때 사용할 수 있습니다.
 - **OTHER= 범주 수** - 제외된 범주를 나타내는 **기타** 범주가 있을 때 사용할 수 있습니다.

기타 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **옵션**을 클릭합니다.

옵션 이름	설명
표본크기 범례	표본 크기 범례를 표시하려면 표본 크기 범례 표시 체크 박스를 선택합니다. 범례는 기본적으로 $N=n$으로 표시되며, 여기서 n은 파레토 범주의 총수가 됩니다. 비교 파레토 그래프를 작성하면 각 성분 그래프에 범례가 표시됩니다. 막대의 방향에 따라 그래프에서의 범례 위치가 결정됩니다. 주: 파레토 그래프에 인셋을 추가하면 표본 크기 범례 표시 체크 박스의 선택을 취소해야 합니다. 파레토 그래프에 인셋과 표본 크기 범례가 포함되면 인셋 및 범례에 대한 상자가 출력에서 겹칠 수 있습니다. 따라서 인셋만 포함하거나 아니면 표본 크기 범례만 표시하는 것이 좋습니다. 범례 주위에 프레임을 생성하려면 배경 드롭다운 리스트에서 색상을 선택합니다. 이 색상은 프레임을 채우는 데 사용됩니다. 표본크기 범례에 레이블을 표시하려면 레이블 텍스트 상자에 최대 32자의 문자열을 입력합니다.
비교 옵션	하나 이상의 분류변수를 분석 역할에 지정하면 페이지당 개수 상자를 사용하여 비교 파레토 그래프의 칼럼과 행 정렬 방식을 변경할 수 있습니다. 한 개의 분류변수를 할당하면 페이지당 한 개의 칼럼과 두 개의 행이 기본 정렬 방식이 됩니다. 두 개의 분류변수를 할당하면 페이지당 두 개의 칼럼과 두 개의 행이 기본 정렬 방식이 됩니다. 레이블 배경 드롭다운 리스트를 사용하여 행 레이블과 칼럼 레이블의 프레임 영역을 채우는 데 사용할 색상을 지정할 수 있습니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

원 그래프

원 그래프 작업 정보	383
원 그래프: 그래프 유형 선택	384
원 그래프: 분석 역할에 변수 할당	384
원 그래프: 모양 옵션 설정	385
원 그래프 옵션 설정	385
레이아웃 지정	386
고급 옵션 설정	387
추가 모양 옵션	387
제목 및 각주 지정	387
속성 보기	387

원 그래프 작업 정보

원 그래프 작업은 데이터를 부채꼴 모양의 원 "조각"으로 표시하여 전체에 대한 상대적 크기를 나타내는 단순, 그룹형 또는 누적형 그래프를 생성합니다. 각 조각은 하나의 데이터 범주를 나타냅니다. 조각의 크기는 전체 그래프 통계량에 대해 해당 데이터가 차지하는 비율을 나타냅니다. 예를 들어 체인의 전체 매출 중 각 점포의 매출을 원 그래프로 표시할 수 있습니다.

자세한 내용은 [“그래픽 출력형식 선택” \(5페이지\)](#)에서 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GCHART
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

원 그래프: 그래프 유형 선택

이 옵션에 액세스하려면 선택 영역에서 **원 그래프**를 클릭합니다.

옵션 이름	설명
단순 원	전체에 대한 부분의 상대적인 크기를 나타내는 그래프를 생성합니다. 데이터가 부채꼴 모양의 원 "조각"으로 표시됩니다. 각 조각은 하나의 데이터 범주를 나타냅니다. 조각의 크기는 전체 그래프 통계량에 대해 해당 데이터가 차지하는 비율을 나타냅니다.
그룹형 원	그룹화 칼럼 역할에 할당한 변수의 각 고윳값에 대해 개별 원 그래프를 생성합니다.
누적형 원	그래프를 누적화 칼럼 역할에 할당한 변수값에 따라 동심원 형태의 고리로 분할합니다. 고리의 너비는 각 하위그룹마다 동일하며 원의 반지름으로 결정됩니다.
그룹형/누적형 원	그룹화 칼럼 역할에 할당한 변수의 각 고윳값에 대해 개별 원 그래프를 생성합니다. 각 그래프는 누적화 칼럼 역할에 할당한 변수값에 따라 동심원 형태의 고리로 분할됩니다.

원 그래프: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

주: 선택한 그래프 유형에 따라 사용할 수 있는 역할이 다릅니다.

역할 이름	설명
그래프 칼럼	이 역할에 할당하는 칼럼의 값에 따라 조각이 결정됩니다. 칼럼은 문자 또는 숫자가 될 수 있습니다.

역할 이름	설명
그룹화 칼럼	이 역할에 할당하는 칼럼의 각 고윳값에 대해 원 그래프가 생성됩니다. 모든 그래프는 동일한 페이지에 표시됩니다.
누적화 칼럼	이 역할에 할당하는 칼럼의 각 고윳값에 대해 동심원(누적)이 생성됩니다. 이 칼럼은 하위그룹 변수라고 불립니다. 주: 그래픽 출력형식으로 Java를 선택했으면 하위그룹 변수의 값은 동심원에 표시되지 않습니다. 대신, 결과 그래프에서 원 조각으로 드릴 다운해서 하위그룹에 대한 값을 표시합니다. 그러나, 그래프 출력형식으로 Java 이미지를 선택하면 그래프가 정적 이미지가기 때문에, Java 그래픽 출력형식에서 사용할 수 있는 드릴다운 기능을 사용하지 못합니다.
합계 칼럼	이 역할에 할당하는 칼럼에 따라 조각의 크기가 결정됩니다.
그룹 분석 칼럼	이 역할에 할당하는 칼럼의 값에 따라 생성되는 그래프의 수가 결정됩니다. 각 고윳값에 대해 별도의 그래프가 생성됩니다. 각 원 그래프는 새로운 페이지에 나타납니다.

원 그래프: 모양 옵션 설정

원 그래프 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **파이**를 클릭합니다.

옵션 이름	설명
색상 지정	조각 색상을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 기본 색상 구성 - 조각에 기본 색상 구성을 사용합니다. 기본 색상 구성은 현재 스타일에 따라 결정됩니다. 스타일 관리자나 결과에 대한 속성 대화 상자를 사용하여 스타일을 지정할 수 있습니다. SAS Enterprise Guide의 옵션 대화 상자에서 각 결과 유형에 대한 스타일을 지정할 수도 있습니다. SAS Add-In for Microsoft Office에서 SAS Add-In 옵션 대화 상자를 사용하여 SAS 리포트 및 HTML 출력형식에 스타일을 적용할 수 있습니다. ■ 기존 색상(이전 버전에서 지정됨) - 이전 버전의 SAS Enterprise Guide 또는 SAS Add-In for Microsoft Office에서 지정한 색상을 사용할 수 있습니다. 이 옵션은 이전 릴리스에서 마이그레이션된 SAS 콘텐츠에 대해서만 사용 가능합니다.

옵션 이름	설명
	■ 사용자 정의 색상 - 조각에 사용할 사용자 정의 색상을 지정할 수 있습니다. 드롭다운 리스트를 사용하여 막대 그래프에 최대 12개의 색상을 지정할 수 있습니다. 그래프에 12개가 넘는 색상을 사용해야 할 때는 현재의 기본 색상 구성에서 추가 색상이 선택됩니다. 사용자 정의 색상을 기본값으로 돌리려면 재설정을 클릭합니다. 주: 도넛 그래프 유형에 따라 일부 사용자 정의 색상을 사용하지 못할 수도 있습니다.
조각 수 지정	그래프의 조각 수를 지정합니다. 기본적으로 조각 수는 자동으로 결정됩니다. 연속 칼럼에서는 값이 여러 범위로 나뉘며 각 범위마다 조각이 한 개씩 표시됩니다. 전체 4% 이하를 나타내는 조각은 "기타"라는 조각으로 합쳐집니다. 그래프에 표시되는 조각 수를 지정하려면 조각 개수 지정 체크 박스를 선택합니다.

레이아웃 지정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **레이아웃**을 클릭합니다.

각 조각의 레이블에 이름, 백분율, 통계값 또는 이 세 가지의 조합 중 어떤 것을 사용할지를 선택합니다. 레이블의 각 요소에 대해 다음 레이블 위치 중 하나를 선택합니다.

- 조각 내부
- 원 또는 도넛 외부
- 조각 외부(선으로 연결 표시)

조각 레이블의 글꼴 색상을 조각과 동일한 색상으로 표시되도록 선택할 수 있습니다. 조각 내부에 레이블을 표시할 때는 이 옵션을 선택하지 마십시오.

"기타" 조각에 대해 백분율 및 레이블을 지정할 수 있습니다. 이 조각은 특정 백분율 값 이하인 그래프 통계량 값을 가지는 중간점을 모은 것입니다. 총합 백분율의 기본값은 4입니다. 그러므로 전체 중 4% 미만을 나타내는 조각은 "기타" 조각에 넣을 수 있습니다.

"기타" 조각의 레이블에는 최대 16자의 텍스트 문자열을 지정할 수 있습니다. 기본 레이블은 기타입니다. "기타" 조각은 조각 순서에 관계없이 원의 마지막 조각입니다.

하나의 중간점만 "기타" 범주에 속하면 해당 중간점에 대한 조각은 원의 원래 위치와 원래 레이블을 유지합니다. 예를 들어 음료의 백분율 판매를 표시하는 원에 다음 조각 및 백분율 값이 있다고 가정합니다. 합계 백분율이 5이면, Soda에 대한 조각이 마지막 조각이 되는 대신에 원에서 세 번째 조각으로 남아 있고, 이 조각은 "기타" 대신에 "Soda"라는 레이블이 붙게 됩니다.

그룹 도표는 **그룹 레이아웃** 영역을 사용하여 그래프를 정렬합니다. 출력 영역의 칼럼 수 또는 행 수를 지정합니다.

고급 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **고급**을 클릭합니다.

그래프의 조각 계산에 사용할 통계량을 드롭다운 리스트에서 선택합니다. **합계 칼럼** 역할에 변수를 할당하지 않으면 각 조각의 빈도 또는 백분율을 계산하도록 선택할 수 있습니다. 기본적으로 빈도가 계산됩니다. **합계 칼럼** 역할에 변수를 할당하면 각 조각의 합계 또는 평균을 계산하도록 선택할 수 있습니다. 기본적으로 합이 계산됩니다.

추가 모양 옵션

일부 옵션은 여러 그래프 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- [“범례 옵션 설정” \(596페이지\)](#)
- [“그래프 영역 사용자 정의” \(597페이지\)](#)

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

.....
주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.
.....

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

원 그래프 마법사

원 그래프 마법사 정보	389
원 그래프 마법사: 데이터 선택	389
원 그래프 마법사: 역할에 변수 할당	390
원 그래프 마법사: 모양 옵션 설정	390
원 그래프 마법사: 제목 및 각주 지정	391

원 그래프 마법사 정보

원 그래프 마법사를 사용하여 2차원이나 3차원의 그룹형/누적형 원 그래프를 생성할 수 있습니다. 각 파이의 조각은 데이터의 범주를 나타냅니다. 조각의 크기는 전체 그래프 통계량에 대해 해당 데이터가 차지하는 비율을 나타냅니다.

원 그래프 마법사: 데이터 선택

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

다음을 클릭하여 역할에 변수를 할당합니다.

원 그래프 마법사: 역할에 변수 할당

- 1 데이터를 선택해야 역할에 변수를 할당할 수 있습니다.
- 2 다음과 같은 역할에 변수를 할당합니다.
 - **조각** - 이 역할에 지정하는 변수의 값에 따라 조각이 결정됩니다. 변수는 문자 또는 숫자가 될 수 있습니다.
다음 옵션을 지정할 수도 있습니다.
 - **각 데이터 값마다 조각 하나** - 문자 변수에서는 이 설정이 적용되지 않습니다. 숫자 변수에서는 각 고유타에 대해 한 개의 조각이 표시됩니다.
 - **조각 자동 생성** - 숫자 변수가 연속된 값으로 간주되고 값이 여러 범위로 나뉘며 각 범위마다 조각이 한 개씩 표시됩니다. 범위 개수는 자동으로 결정됩니다. 전체 4% 이하를 나타내는 조각은 "기타"라는 조각으로 합쳐집니다.
 - **조각 크기** - 이 역할에 지정하는 변수나 통계량에 따라 조각 크기가 결정됩니다.
 - 이 역할에 변수를 할당하면 각 조각의 합계 또는 평균을 계산할 수 있습니다. 기본적으로 합이 계산됩니다.
 - 이 역할에 변수를 할당하지 않으면, 각 조각의 빈도나 백분율을 계산할 수 있습니다. 기본적으로 빈도가 계산됩니다.
- 3 (선택 사항) 다음과 같은 역할에 변수를 할당합니다.
 - **그룹화** - 이 역할에 지정하는 변수의 값은 그룹으로 정리됩니다. 각 고유타에 대해 별도의 파이가 생성됩니다.
 - **누적화** - 이 역할에 지정하는 변수의 값에 따라 원 그래프에 대한 누적 및 동심원 형태의 고리 개수가 결정됩니다. 그래프 출력 유형이 Java이면, 그래프에 동심원 형태의 고리보다는 드릴다운 기능이 있게 됩니다.
 - **그래프 분류** - 각 BY 그룹에 대해 생성된 그래프를 구별합니다. 이 역할에 지정하는 변수의 값에 따라 그룹이 결정됩니다. 각 BY 그룹에 대해 출력이 생성됩니다.
- 4 다음을 클릭하여 모양 옵션을 설정합니다.

원 그래프 마법사: 모양 옵션 설정

원 그래프 마법사는 그래프에서 사용할 색상을 자동으로 결정합니다. 구성 드롭다운 리스트에서 이 색상 구성을 볼 수 있습니다.

다음과 같은 모양 옵션을 지정합니다.

- 2차원 또는 3차원 그래프

- 범례
- 그래프에서 각 조각의 변수 이름
- 그래프에서 각 조각의 백분율
- 그래프에서 각 조각의 데이터 값

다음을 클릭하여 제목과 각주를 지정합니다.

원 그래프 마법사: 제목 및 각주 지정

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

- 1 모양 옵션을 설정해야 제목과 주석을 지정할 수 있습니다.
- 2 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다. 제목과 각주에는 매크로 변수를 사용할 수 있습니다.
- 3 **마침**을 클릭하여 작업을 실행합니다.

P-P 도표

P-P 도표 작업 정보	393
공정능력분석: 분석 역할에 변수 할당	394
P-P 도표: 분포 선택	395
P-P 도표: 분포 사용자 정의	396
공정능력분석: 도표 모양 옵션 설정	396
축 옵션 설정	396
참조선 추가	397
인셋 추가	397
기타 도표 옵션 설정	397
규격 한계 옵션 설정	399
공정능력분석: 분석 유형 지정	399
제목 및 각주 지정	400
속성 보기	400

P-P 도표 작업 정보

공정능력분석은 통계적 관리 상태 공정의 출력분포를 해당 규격한계와 비교하여 규격과 일치하는지를 파악합니다. 확률-확률 도표(P-P 도표 또는 백분율 도표라고도 함)는 변수의 경험적 누적분포함수(ECDF)를 정규과 같은 지정된 이론적 누적분포함수와 비교하기 위해 공정능력분석에 사용됩니다. 두 분포가 일치할 때 도표 상의 점들이 원점을 지나고 단위 기울기를 갖는 직선 형태를 이루게 됩니다. 따라서 P-P 도표를 사용하면 이론분포가 측정값을 얼마나 잘 반영하는지를 파악할 수 있습니다.

P-P 도표를 사용하여 제조 과정을 분석할 수 있습니다. 예를 들어 50개의 철판을 절단하여 두 개의 구멍 사이의 간격을 측정하고 데이터셋에 저장했다고 가정합니다. 절단 과정은 통계적 관리 상태에 있습니다. 거리가 정규분포되어 있는지 확인하려고 합니다. P-도표 작업을 사용하여 정규 누적분포함수를 기반으로 하는 도표를 생성할 수 있습니다.

주: 확률-확률 도표를 지정된 분포의 백분위수를 가진 정렬된 측도 집합을 비교하는 확률 도표와 혼동하지 마십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	CAPABILITY
필수 SAS 제품	Base SAS, SAS/QC, SAS/GRAPH
추가 SAS 제품(권장)	SAS/STAT

공정능력분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분석변수	분석할 숫자 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다. P-P 도표를 제외한 모든 도표에서는 변수 이름의 규격한계 영역에서 각 분석변수에 대한 상한, 목표 값 및 하한을 지정할 수 있습니다. 이 한계를 지정하면 정규성 검정을 보여 주는 테이블, 규격한계를 보여 주는 테이블 및 공정능력지수를 보여 주는 테이블이 자동으로 출력에 포함됩니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n은 해당 행의 빈도변수값입니다.
상대 가중값 변수	상대 가중값으로 사용할 변수를 지정합니다. 이 역할에 변수를 할당하면 각 통계량의 변수값이 가중 통계량 계산에 사용됩니다. 이 역할에는 변수를 하나만 할당할 수 있습니다.

역할 이름	설명
	주: 히스토그램, 확률 도표, P-P 도표, Q-Q 도표 또는 CDF 도표를 생성할 때는 상대 가중값 변수 역할에 변수를 할당하지 않는 것이 좋습니다. 이 역할에 변수를 할당했으면 출력을 해석할 때 주의해야 합니다.
분류변수 (히스토그램 작업에 서만 사용 가능)	분류변수로 사용할 하나 이상의 변수를 지정합니다. 이 역할에 변수를 할당하면 전체 분류 레벨에서 분석변수의 분포를 비교할 수 있는 비교 히스토그램이 생성됩니다. 주: 비교 확률 도표 또는 비교 Q-Q 도표를 생성하려면 분포분석 작업을 사용하십시오.

P-P 도표: 분포 선택

이 옵션에 액세스하려면 선택 영역의 분포 머리글 아래에서 **요약**을 클릭합니다.

P-P 도표에는 하나의 이론분포를 적합시킬 수 있습니다. 분포를 선택한 후 해당 분포의 모수 값 및 선 속성을 지정하려면 선택 영역에서 분포 이름을 선택합니다.

옵션 이름	설명
분포	체크 박스를 사용하여 다음 추정된 밀도 곡선을 하나 이상 선택합니다. ■ 정규 - 평균(μ)과 표준편차(σ)를 가지는 정규밀도함수를 적합시킵니다. ■ 로그 정규 - 척도모수(ζ), 임계모수(θ) 및 형상모수(σ)를 가지는 로그 정규 밀도함수를 적합시킵니다. ■ 지수 - 척도모수(σ) 및 임계모수(θ)를 가지는 지수밀도함수를 적합시킵니다. ■ 와이블 - 척도모수(σ), 임계모수(θ) 및 형상모수(c)를 가지는 3-모수 와이블 밀도함수를 적합시킵니다. ■ 베타 - 척도모수(σ), 임계모수(θ) 및 형상모수(α 및 β)를 가지는 베타 밀도함수를 적합시킵니다. ■ 감마 - 척도모수(σ), 임계모수(θ) 및 형상모수(α)를 가지는 감마 밀도함수를 적합시킵니다.
그래프 스타일	출력에 사용할 그래프 스타일을 선택합니다. 기존 SAS 그래프나 ODS 통계 그래프 중에 선택할 수 있습니다. ODS 그래프에 대한 자세한 내용은 SAS Output Delivery System: User's Guide를 참조하십시오.

P-P 도표: 분포 사용자 정의

이 옵션에 액세스하려면 선택 영역의 분포 머리글 아래에서 사용자 정의할 분포 이름을 클릭합니다.

주: 이러한 분포 옵션은 **요약** 페이지에서 분포를 선택했거나 해당 분포 페이지에서 **분포 이름** 체크 박스를 선택할 때만 사용할 수 있습니다.

정규, 로그 정규, 지수, 와이블, 베타 및 감마 분포에서는 각 분석변수 또는 모든 분석변수에 대해 모수를 지정할 수 있습니다. 기본적으로 작업에서는 모든 분석변수의 모수 값을 추정합니다.

각 분석변수에 대한 모수를 지정하는 방법:

- 1 **모든 변수에 분포 적용** 체크 박스의 선택을 해제합니다.
- 2 **분석변수** 상자에서 변수에 해당되는 체크 박스를 선택합니다.
- 3 모수값을 지정합니다.

주: 추정 모수값을 생성하려면 **추정값 사용**을 클릭합니다.

공정능력분석: 도표 모양 옵션 설정

축 옵션 설정

이 옵션에 액세스하려면 선택 영역의 **모양** ⇨ **축** 머리글 아래에서 **축**을 클릭합니다.

다음 옵션을 설정할 수 있습니다.

- 축 색상
- 눈금 표시 색상
- 가로축 및 세로축 레이블

참조선 추가

이 옵션에 액세스하려면 선택 영역의 **모양** ⇨ **축** 머리글 아래에서 **가로** 및 **세로**를 클릭합니다.

참조선을 사용하려면 **참조선 사용**을 선택합니다. 참조선의 출력형식을 지정하고 각 참조선의 위치를 지정할 수 있습니다.

인셋 추가

이 옵션에 액세스하려면 선택 영역의 **모양** 머리글 아래에서 **인셋**을 클릭합니다.

요약통계량 상자 또는 테이블(인셋)을 그래프에 직접 추가하여 도표를 보기 좋게 만들 수 있습니다. 비교 히스토그램을 생성하면 각각의 작은 히스토그램에 인셋이 표시됩니다.

인셋 포함을 선택하여 도표에 요약통계량을 추가합니다. 인셋에 포함할 통계량을 선택합니다.

주: 공정능력지수를 계산하기 위한 통계량은 **분석변수** 역할에 할당한 변수에 대한 규격한계를 생성할 때만 사용할 수 있습니다. 히스토그램, CDF 도표, 확률 도표 및 Q-Q 도표에 대한 규격한계를 생성할 수 있습니다. 공정능력지수 계산에 대한 자세한 내용은 CAPABILITY 프로시저에 대한 도움말을 참조하십시오.

기타 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **옵션**을 클릭합니다.

사용할 수 있는 도표 모양 옵션은 작업에서 생성되는 도표의 유형에 따라 다릅니다.

역할 이름	설명
배경	도표의 배경 색상을 지정합니다.
막대 외곽선 및 막대 채우기	히스토그램의 막대 색상을 지정합니다. 주: 이 옵션은 비교 히스토그램에서만 사용할 수 있습니다.
축 바꾸기 : 백분위수를 세로축으로 사용	도표에서 축을 반대로 바꿉니다. 이 옵션을 선택하면 도표의 방향이 바뀌므로 선택한 축 옵션 및 참조선에서 지정한 내용은 적용되지 않습니다. 주: 이 옵션은 확률 도표 및 Q-Q 도표에서만 사용할 수 있습니다.

역할 이름	설명
기본 범례 생략	도표에 기본 범례를 표시하지 않습니다.
중간점 사용	히스토그램에 대한 중간점을 지정합니다. 드롭다운 리스트에서 사용할 수 있는 옵션은 히스토그램을 생성하는지 또는 비교 히스토그램을 생성하는지에 따라 달라집니다. 분류변수 역할에 변수를 할당할 때에는 비교 히스토그램을 생성합니다. 비교 히스토그램에서는 다음 옵션을 사용할 수 있습니다. ■ 중간점 지정 - 시작 중간점과 마지막 중간점을 지정할 수 있습니다. 다른 중간점을 계산하는 데 사용할 단계 간격도 지정할 수 있습니다. 중간점은 각 막대로 표시되는 값 범위의 중간 값입니다. ■ 키 - 키 셀의 데이터에 대해 중간점을 결정합니다. 기본적으로 분류변수의 레벨은 첫 번째 분류변수의 내부(서식화되지 않은) 값의 증가 순으로 위에서 아래로(왼쪽에서 오른쪽으로) 표시됩니다. 분류변수를 하나만 지정하면 키 셀은 이러한 순서로 첫 번째에 나타나는 레벨이 됩니다. 분류변수를 두 개 지정하면 키 셀은 이러한 순서로 첫 번째에 나타나는 변수 1과 변수 2의 레벨 결합이 됩니다. 따라서 키 셀의 선택에 따라 모든 셀에 사용되는 균등 수평 축이 결정됩니다. 그런 다음 키 셀에 대한 중간점 리스트에 나머지 셀의 데이터가 포함될 때까지 리스트가 어느 방향이든지 필요한 만큼 확장됩니다. ■ 균등 - 총 표본 크기를 기반으로 중간점 개수를 결정합니다. 분류변수 역할에 변수를 할당하지 않을 때에는 히스토그램을 생성합니다. 히스토그램에서는 다음 옵션을 사용할 수 있습니다. ■ 중간점 지정 - 시작 중간점과 마지막 중간점을 지정할 수 있습니다. 다른 중간점을 계산하는 데 사용할 단계 간격도 지정할 수 있습니다. 중간점은 각 막대로 표시되는 값 범위의 중간 값입니다. ■ 그룹 수 지정 - 데이터에 대한 막대 수(히스토그램 구간이라고도 함)를 지정할 수 있습니다. 기본적으로 표시되는 구간 수는 데이터 범위에 따라 달라집니다. 그러나 표시할 최대 구간 수는 지정할 수 있습니다. 주: 히스토그램 작업을 사용하는 경우에만 중간점을 지정할 수 있습니다.
사용자 정의 그래프 크기 지정(픽셀)	이미지 크기를 지정합니다. 참조할 수 있도록 전체 화면 크기가 너비 및 높이 상자에 제공됩니다.

주: 비교 옵션은 비교 히스토그램에서만 사용할 수 있습니다. 히스토그램 작업에서 **분류변수** 역할에 변수를 할당하여 비교 히스토그램을 생성할 수 있습니다.

역할 이름	설명
행 및 칼럼	분석 역할에 하나 이상의 분류변수를 할당하면, 페이지당 행 수 및 페이지당 칼럼 수 상자를 사용하여 비교 히스토그램에서의 칼럼과 행 배열 방식을 변경할 수 있습니다. 한 개의 분류변수를 할당하면 페이

역할 이름	설명
	지당 한 개의 칼럼과 두 개의 행이 기본 정렬 방식이 됩니다. 두 개의 분류변수를 할당하면 페이지당 두 개의 행과 두 개의 칼럼이 기본 정렬 방식이 됩니다. 옆 프레임 및 위 프레임 드롭다운 리스트를 사용하여 행 레이블과 칼럼 레이블의 프레임 영역을 채울 색상을 지정할 수도 있습니다.
막대 수	기본적으로 표시되는 구간 수는 데이터 범위에 따라 달라집니다. 그러나 표시할 최대 구간 수는 지정할 수 있습니다. 또한 표시되는 구간 수가 데이터 평균 이상 및 이하의 표준편차 값 범위를 기반으로 하도록 표준편차를 지정할 수 있습니다.

규격 한계 옵션 설정

주: 히스토그램, 확률 도표, Q-Q 도표 및 CDF 도표에서 이러한 옵션을 사용할 수 있습니다.

선택 영역의 **모양** 머리글 아래에서 **규격한계**를 클릭하여 규격하한, 목표 값 및 규격상한을 지정합니다.

주: **규격한계** 옵션은 적어도 하나의 분석변수에 대해 하나 이상의 한계를 지정할 때만 사용할 수 있습니다. 분석변수에 대한 규격한계를 지정하려면 선택 영역에서 **데이터**를 클릭하고 분석변수를 선택합니다.

공정능력분석: 분석 유형 지정

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

공정능력분석 리포트에는 기본적으로 기본 신뢰구간 테이블, 기본 측도 테이블, 위치모수에 대한 검정 테이블, 적률 테이블 등이 포함됩니다.

리포트에 테이블을 추가하려면 **테이블** 상자에서 테이블을 선택합니다.

분석 옵션을 지정하려면 **테이블** 상자에서 테이블을 선택합니다. 옵션을 사용할 수 있으면, 옵션을 지정하는 데 사용한 오른쪽 영역에 상자가 나타납니다.

리포트에서 모든 기술통계량 테이블을 숨기려면 **기술통계량 및 공정능력지수 테이블 생략** 체크 박스를 선택합니다. 이 옵션을 선택해도 다음 테이블은 숨겨지지 않습니다.

- 선택한 분포에 의해 생성된 테이블. 해당 분포에 대한 패널에서 **분포테이블 생략** 체크 박스를 선택하여 이러한 테이블을 숨길 수 있습니다.

- INTERVALS 문에 의해 생성된 테이블. NOPRINT 옵션을 지정하면 이러한 테이블을 생략할 수 있습니다. INTERVALS 문에 대한 자세한 내용은 CAPABILITY 프로시저에 대한 도움말을 참조하십시오.

결과를 출력 데이터셋에 저장하려면 **데이터셋에 출력 통계량 저장** 체크 박스를 선택합니다. 데이터셋에는 일반량 통계량 및 공정능력지수가 포함됩니다.

주: SAS Enterprise Guide는 **옵션 창의 작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간

- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 속성 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

시계열 데이터 준비

시계열 데이터 준비 작업 정보	403
시계열 데이터 준비: 분석 역할에 변수 할당	404
시계열 데이터 준비: 출력 빈도 지정	405
시계열 데이터 준비: 시계열 특성 지정	405
시계열 데이터 준비: 각 시계열 변수에 대한 변환 지정	406
시계열 데이터 준비: 결과 옵션 설정	407
속성 보기	408
데이터셋의 결측값에 대한 관측값 행 생성	408
표본 추출 빈도 간의 시계열 데이터 변환	409
입력 데이터의 관측값 특성 변경	410
비주기적 계열을 주기적 추정값으로 변환	410
결측값 보간	411
데이터에 일반 변환 적용	412

시계열 데이터 준비 작업 정보

시계열 데이터 준비 작업을 사용하면 데이터를 다른 시계열 작업을 통한 분석에 더욱 적합하도록 만들 수 있습니다. 또한 다른 작업에서 사용하도록 구성된 데이터에 일반적인 변환을 적용하는 데 사용할 수도 있습니다. 예를 들어 데이터셋의 변수에 수학적 연산, 함수 및 기타 변환을 적용할 수 있습니다.

이 작업을 사용하여 시계열 데이터를 조작할 때, 선택한 옵션에 따라 이후에 데이터에 수행되는 분석의 정확도에 영향을 미칠 수 있으므로 많은 주의를 기울여야 합니다.

이 작업을 사용하여 다음과 같은 작업을 수행할 수 있습니다.

- 연간 시계열로부터 분기 추정값 보간 또는 연간 시계열 생성을 위한 분기값 집계
- 서로 다른 표본 추출 빈도에서 측정된 시계열 결합. 예를 들어 월간 단기 자금주, 분기별 국내 총생산(GDP), 주간 이자율에 대한 데이터가 있고 이러한 변수를 모두 사용하는 모

델에 대한 분석을 수행하려고 합니다. 분석을 수행하려면 먼저 시계열을 공통 빈도로 변환하고 변수를 하나의 데이터셋으로 통합해야 합니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	EXPAND
필수 SAS 제품	Base SAS, SAS/ETS
추가 SAS 제품(권장)	없음

시계열 데이터 준비: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
시간 ID 변수	행에 날짜를 할당하는 데 사용될 변수를 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다. SAS 날짜, 시간 또는 날짜/시간 출력형식을 갖는 첫 번째 변수가 자동으로 이 역할에 할당됩니다. 시간 ID 변수를 생성할 수도 있습니다.
시계열 변수	분석에 사용할 수 있도록 이 작업에서 사용 가능한 수학적 연산을 적용할 변수를 지정합니다. 변수를 지정하지 않으면 입력 데이터의 모든 숫자 변수가 특성 탭에 설정한 옵션에 따라 변환됩니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 다른 역할에 이미 지정한 변수로는 그룹 분석을 수행할 수 없습니다.

시계열 데이터 준비: 출력 빈도 지정

이 옵션에 액세스하려면 선택 영역의 특성 머리글 아래에서 **빈도**를 클릭합니다.

- 출력 빈도가 변경되지 않게 하려면 **입력 빈도와 같음**을 선택합니다.
- 빈도 간격을 변경하려면 **간격** 드롭다운 리스트에서 출력 간격을 선택합니다.
- 출력 빈도에 대한 입력 빈도 비율을 변경하려면 **비율**을 선택합니다. **입력 관측값**과 **출력 관측값** 상자를 사용하여 출력 관측값에 대한 입력 관측값의 비율을 지정합니다.

시계열 데이터 준비: 시계열 특성 지정

이 옵션에 액세스하려면 선택 영역의 특성 머리글 아래에서 **옵션**을 클릭합니다.

옵션 이름	설명
보간법	보간법을 지정합니다. 다음과 같은 방법을 사용할 수 있습니다. ■ 삼차 스플라인 - 3차 스플라인 곡선과 입력값을 적합시킵니다. 3차 스플라인은 서로 결합된 3차 다항 함수를 구성하는 분절된 함수로서, 전체 곡선과 1차 및 2차 미분계수는 연속됩니다. ■ 일차 스플라인 - 연속적인 직선을 연결하여 연속 곡선과 데이터를 적합시킵니다. ■ 스텝함수 - 비연속적이고 구분적인 일정한 곡선을 적합시킵니다. ■ 단순 집계 - 결측값의 보간 없이 시계열의 단순 집계를 수행합니다. ■ 보간 안 함 - 보간이 실행되지 않도록 지정합니다. 주: 빈도 변환을 선택했으면 보간 안 함 옵션을 사용할 수 없습니다. 자세한 내용은 “시계열 데이터 준비: 출력 빈도 지정” (405페이지)에서 참조하십시오. 삼차 스플라인 방법에서는 하단 및 상단 끝점 제약 조건을 선택할 수 있습니다. 끝점 제약 조건 옵션은 다음과 같습니다. ■ Not-a-knot는 처음 두 개의 스플라인 조각이 마지막 두 조각과 마찬가지로 동일한 삼차 곡선의 일부로 제한됨을 의미합니다. 기본값입니다. ■ Natural spline은 내추럴 스플라인 제약 조건을 지정합니다. 스플라인 곡선의 2차 미분계수는 끝점에서 0이 되도록 제한됩니다. ■ Slope - 끝점 스플라인 곡선의 1차 미분계수를 지정합니다. ■ Curvature - 끝점 스플라인 곡선의 2차 미분계수를 지정합니다.

옵션 이름	설명
	Slope 및 Curvature 옵션에서는 각각 1차 미분계수 및 2차 미분계수 값을 지정해야 합니다. 두 미분계수의 기본값은 1입니다.
관측값 특성	입력 및 출력 시계열의 관측값 특성을 지정합니다. 관측값 특성은 시계열에 있는 각 값의 관측값 유형을 정의합니다. 이는 보간이 올바르게 수행될 수 있도록 합니다. 사용할 수 있는 옵션은 다음과 같습니다. ■ Total - 데이터 값이 관측값에 해당하는 시간 구간에 대한 기간 합계를 나타내도록 지정합니다. ■ Average - 데이터 값이 기간 평균을 나타내도록 지정합니다. ■ Beginning - 데이터 값이 주기 시작값을 나타내도록 지정합니다. 기본값입니다. ■ Middle - 데이터 값이 주기 중간점값을 나타내도록 지정합니다. ■ End - 데이터 값이 주기 끝값을 나타내도록 지정합니다. ■ Derivative - 출력 데이터가 입력 데이터에 대한 3차 스플라인 곡선 맞춤의 미분계수가 되도록 지정합니다. 주: Derivative 옵션은 출력 드롭다운 리스트에서만 사용할 수 있습니다. 날짜 정렬과 관측값 특성의 차이를 혼동하지 않도록 해야 합니다. 날짜 정렬 옵션은 각 관측값별로 시간 ID 변수에 지정되는 날짜값을 제어하며 관측값 특성은 보간을 실행할 때 시계열값이 수학적으로 해석되는 방법을 제어합니다. 관측값 특성은 보간이 수행될 때 시계열 값이 수학적으로 해석되는 방법을 제어합니다.
간격내에서 날짜 정렬	각 간격 내에서의 날짜 정렬을 지정합니다. 예를 들어 월간 간격을 지정하면, 시작 옵션은 해당 월의 첫 번째 날에 시간 간격을 시작하고 중간 옵션은 해당 월의 중간 날에 시간 간격을 시작하며 마지막 옵션은 해당 월의 마지막 날에 시간 간격을 시작합니다.

시계열 데이터 준비: 각 시계열 변수에 대한 변환 지정

이 옵션에 액세스하려면 선택 영역의 변환 머리글 아래에서 **보간 전** 또는 **보간 후**를 클릭합니다.

시계열 변수 역할에 할당된 각 변수마다 고유한 변환을 적용할 수 있습니다.

옵션 이름	설명
보간 전/보간 후 연산	추가를 클릭하면 사용 가능한 변환 리스트가 표시됩니다. 추가할 변환의 체크 박스를 선택하고 확인을 클릭합니다. 변환을 제거하려면 제거할 변환을 선택하고 제거를 클릭합니다. 변환은 리스트에 표시된 순서대로 수행됩니다. 변환 순서를 변경하려면 변환을 선택한 다음 위쪽 및 아래쪽 화살표 버튼을 사용하여 리스트에서 해당 변환의 위치를 이동합니다. 예를 들어 $y = 5 + \log(x + 10)$라는 변환을 입력 시계열에 적용하기 위해 다음과 같은 변환을 순서대로 리스트에 입력할 수 있습니다. 1 상수 추가($n = 10$) 2 $\log 3$ 3 상수 추가($n = 5$)
선택한 연산에 대한 옵션	이러한 옵션은 선택한 연산에 대한 옵션 영역에 표시됩니다. 모든 변환 모수가 지정되어야만 실행 및 저장 버튼이 활성화됩니다.

시계열 데이터 준비: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
출력 데이터 저장	변환된 시계열 변수를 포함하는 데이터셋은 항상 저장됩니다. 이 출력 데이터셋의 이름은 수정된 데이터 상자에 표시됩니다. 입력 시계열에 적합한 스플라인 곡선의 계수를 영구적으로 저장하려면 모수 추정값 체크 박스를 선택합니다. 주: SAS Enterprise Guide는 옵션 창의 작업 > 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
출력에 시계열 그래프 포함	변환된 입력 시계열, 변환된 출력 시계열 및 변환된 시계열 도표를 출력에 포함하려면 입출력 시계열 그래프 생성 체크 박스를 선택합니다. 시계열 변수 역할에 할당된 각 변수에 시계열 및 중간 시계열의 도표가 생성됩니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

데이터셋의 결측값에 대한 관측값 행 생성

대부분의 시계열 분석 작업은 시계열 데이터 관측값이 동일한 시간 간격 또는 주기마다 얻어지는 것으로 가정합니다. 특정 시간 간격이 결측이면, 데이터의 간격이 동일하지 않게 되어 분석이 실패하거나 잘못된 결과가 산출될 수 있습니다. 이 데이터는 간격이 일정하지 않습니다.

행	날짜	값
1	JAN10	100
2	FEB10	110
3	MAR10	120

4 MAY10 140

5 JUN10 150

이 데이터에서는 관측값 3과 4 사이의 간격과 관측값 2와 3 사이의 간격이 서로 다를 수 있습니다.

결측된 간격에 대한 관측값을 생성하는 방법

- 1 **시간 ID 변수** 역할에 날짜 변수를 할당합니다.
- 2 간격을 데이터의 빈도로 설정합니다. 예제의 데이터는 **월별** 데이터입니다.
- 3 **실행**을 클릭합니다.

원본 데이터셋의 모든 칼럼을 포함하는 새로운 데이터셋이 생성됩니다. 누락된 간격의 새로운 관측값이 값이 결측으로 설정되어 추가됩니다. 이제 이 데이터를 결측값을 처리하는 작업에서 사용하거나 시계열 데이터 준비 작업을 새로운 데이터셋에 사용하여 결측값을 보간할 수 있습니다. 자세한 내용은 [“결측값 보간” \(411페이지\)](#)에서 참조하십시오.

또한 결측값을 보간하여 결측 관측값의 추정값을 제공할 수도 있습니다. 자세한 내용은 [“비주기적 계열을 주기적 추정값으로 변환” \(410페이지\)](#)에서 참조하십시오.

표본 추출 빈도 간의 시계열 데이터 변환

시계열 데이터 준비 작업을 사용하여 시계열 데이터를 다른 표본 추출 빈도로 변환할 수 있습니다. 예를 들어 주간 데이터를 월간 데이터로 집계하거나 월간 데이터에서 주간 데이터를 보간할 수 있습니다.

시계열 데이터의 빈도를 변환하는 방법

- 1 **시간 ID 변수** 역할에 날짜 변수를 할당합니다. **기존 시간 ID 변수 사용** 영역에서 간격을 입력 데이터의 빈도로 설정합니다.
- 2 변환할 시계열 변수를 **시계열 변수** 역할에 할당합니다.
- 3 선택 영역의 특성 머리글 아래에서 **빈도**를 클릭하고, **출력 빈도** 영역에서 **비율** 또는 **간격**을 선택합니다. 예를 들어 입력 데이터를 주간 빈도로 변환하려면 **간격**을 선택하고 간격 값을 **주로** 설정합니다.
- 4 선택 영역의 특성 머리글 아래에서 **옵션**을 클릭하고 보간법을 데이터에 적합한 방법으로 변경합니다. 예를 들어 **삼차 스플라인**을 선택하여 3차 스플라인 곡선과 입력 데이터를 적합시킵니다.
- 5 입력 데이터 관측값 특성 및 원하는 출력 데이터 관측값 특성을 선택합니다. 예를 들어 월별 데이터를 분기별 평균으로 변환할 수 있습니다.
- 6 원하는 출력 날짜 정렬을 선택합니다.
- 7 **실행**을 클릭합니다.

입력 데이터의 관측값 특성 변경

시계열 데이터 준비 작업은 입력 데이터의 관측값 특성을 변경하는 데 사용할 수 있습니다. 예를 들어 월간 합계 데이터를 월간 평균 데이터로 변환할 수 있습니다.

시계열 데이터의 관측값 특성을 변경하는 방법

- 1 **시간 ID 변수** 역할에 날짜 변수를 할당합니다. **기존 시간 ID 변수 사용** 영역에서 간격을 입력 데이터의 빈도로 설정합니다.
- 2 변환할 시계열 변수를 **시계열 변수** 역할에 할당합니다.
- 3 선택 영역의 특성 머리글 아래에서 **빈도**를 클릭합니다. 데이터의 빈도를 바꾸지 않으려면 **출력 빈도** 영역에서 출력 빈도를 입력 빈도와 동일하게 유지합니다.
- 4 선택 영역의 특성 머리글 아래에서 **옵션**을 클릭하고 보간법을 데이터에 적합한 방법으로 변경합니다. 예를 들어 **삼차 스플라인**을 선택하여 3차 스플라인 곡선과 입력 데이터를 적합시킵니다.
- 5 입력 데이터 관측값 특성 및 원하는 출력 데이터 관측값 특성을 선택합니다. 예를 들어 월별 데이터를 분기별 평균으로 변환할 수 있습니다.
- 6 원하는 출력 날짜 정렬을 선택합니다.
- 7 **실행**을 클릭합니다.

비주기적 계열을 주기적 추정값으로 변환

대부분의 시계열 분석 작업은 시계열 데이터 관측값이 동일한 시간 간격으로 또는 주기적으로 발생하며 누락되는 관측값이 없다고 가정합니다. 따라서 트랜잭션 데이터 등의 비주기적 계열은 먼저 주기적 데이터로 변환되어야 다른 시계열 작업을 통해 분석할 수 있습니다.

비주기적 데이터를 주기적 추정값으로 변환하는 방법

- 1 **시간 ID 변수** 역할에 날짜 변수를 할당합니다. **기존 시간 ID 변수 사용** 영역에서 **비주기적 데이터** 옵션을 선택합니다.
- 2 변환할 시계열 변수를 **시계열 변수** 역할에 할당합니다.
- 3 선택 영역의 특성 머리글 아래에서 **빈도**를 클릭하고 원하는 출력 간격을 선택합니다.
- 4 선택 영역의 특성 머리글 아래에서 **옵션**을 클릭하고 보간법을 데이터에 적합한 방법으로 변경합니다. 예를 들어 **삼차 스플라인**을 선택하여 3차 스플라인 곡선과 입력 데이터를 적합시킵니다.
- 5 원하는 출력 날짜 정렬을 선택합니다.

- 6 입력 데이터 관측값 특성 및 원하는 출력 데이터 관측값 특성을 선택합니다. 예를 들어 비주기적인 시간별 값을 월간 평균으로 변환할 수 있습니다.
- 7 **실행**을 클릭합니다.

결측값 보간

시계열 데이터 준비 작업은 입력 시계열의 결측값을 보간하는 데 사용할 수 있습니다. 이 데이터에는 결측값이 포함되어 있습니다.

행	날짜	값
1	JAN10	100
2	FEB10	110
3	MAR10	120
4	APR10	.
5	MAY10	140
6	JUN10	150
7	JUL10	.
8	AUG10	150

입력 데이터의 결측값을 보간하는 방법

- 1 **시간 ID 변수** 역할에 날짜 변수를 할당합니다. **기존 시간 ID 변수 사용** 영역에서 간격을 입력 데이터의 빈도로 설정합니다.
- 2 변환할 시계열 변수를 **시계열 변수** 역할에 할당합니다.
- 3 선택 영역의 특성 머리글 아래에서 **빈도**를 클릭하고 출력 시계열 빈도로 **입력 빈도와 같음**을 선택합니다.
- 4 선택 영역의 특성 머리글 아래에서 **옵션**을 클릭하고 보간법을 데이터에 적합한 방법으로 변경합니다. 예를 들어 **삼차 스플라인**을 선택하여 3차 스플라인 곡선과 입력 데이터를 적합시킵니다.
- 5 입력 데이터 관측값 특성 및 원하는 출력 데이터 관측값 특성을 선택합니다. 예를 들어 일별 데이터를 월간 평균으로 변환할 수 있습니다.
- 6 원하는 출력 날짜 정렬을 선택합니다.

- 7 **실행**을 클릭합니다.

데이터에 일반 변환 적용

시계열 데이터 준비 작업을 사용하여 입력 시계열에 대해 다양한 수학적 변환을 수행할 수 있습니다. 이러한 변환은 보간 연산과 함께 수행하거나 보간 없이 수행할 수 있습니다.

보간 연산과 함께 데이터 변환을 수행하려면 작업 실행을 제외한 보간 옵션의 모든 단계를 수행합니다. (자세한 내용은 **“결측값 보간” (411페이지)**에서 참조하십시오.) 그런 다음, **변환** 페이지를 사용하여 보간 전 또는 후 데이터에 수행할 변환을 정의합니다. 두 리스트는 비워둘 수 있습니다.

보간 없이 데이터를 변환하는 방법

- 1 **시간 ID 변수** 역할에 날짜 변수를 할당합니다. **기존 시간 ID 변수 사용** 영역에서 간격을 입력 데이터의 빈도로 설정합니다.
- 2 변환할 시계열 변수를 **시계열 변수** 역할에 할당합니다.
- 3 선택 영역의 특성 머리글 아래에서 **옵션**을 클릭하고 보간법으로 **보간 안 함**을 선택합니다.
- 4 원하는 출력 날짜 정렬을 선택합니다.
- 5 선택 영역의 변환 머리글 아래에서 **보간 전**을 클릭하여 수행할 변환을 삽입합니다. 보통 **보간 후** 리스트를 비워 두고 **보간 후** 페이지에서 변환을 정의할 수도 있습니다. 이러한 변환은 **보간 전** 리스트의 변환이 수행된 후에 수행됩니다.
- 6 **실행**을 클릭합니다.

주성분분석

주성분분석 작업 정보	413
주성분분석: 분석 역할에 변수 할당	414
주성분분석: 분석 옵션 설정	414
주성분분석: 도표 생성	415
주성분분석: 결과 옵션 설정	416
제목 및 각주 지정	417
속성 보기	417

주성분분석 작업 정보

주성분분석은 여러 양적변수 간의 관계를 검사하기 위한 다변량분석 기법입니다. 데이터를 요약하거나 선형 관계를 파악하려면 주성분분석을 사용해야 합니다.

너무 많은 변수를 동시에 도표로 나타낼 때 주성분 작업을 사용할 수 있습니다. 예를 들어 1977년에 50개 주 각각에 대해 7개 범주의 인구 100,000명당 범죄율이 있다고 가정합니다. 7개의 숫자 변수가 있으므로 모든 변수를 동시에 도표화하는 것은 불가능합니다. 주성분 분석을 사용하여 데이터를 2 또는 3차원으로 요약하고 데이터를 시각화할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	PRINCOMP
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

주성분분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분석변수	리스트의 각 변수에 대해 분석을 수행합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
편상관변수	부분상관계수 또는 공분산행렬을 분석할 때 부분상관변수로 사용할 변수를 지정합니다. 주성분분석 작업은 이 역할에 지정된 각 분석변수의 예측값으로부터 잔차의 주성분을 계산합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 분석 변수로 선택한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
상대 가중값 변수	분석 변수의 해당 값에 가중값을 부여하는 데 사용할 값을 지정합니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.

주성분분석: 분석 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **분석**을 클릭합니다.

옵션 이름	설명
분석	수행할 분석을 지정합니다. ■ 상관계수 - 상관행렬에서 주성분을 계산합니다.

옵션 이름	설명
	■ 공분산 - 공분산행렬에서 주성분을 계산합니다. 이 옵션을 지정하면 분산 값이 큰 변수는 큰 고윳값을 가진 성분과 더욱 밀접하게 연관되고, 분산 값이 작은 변수는 작은 고윳값을 가진 성분과 밀접하게 연결됩니다. 측정된 변수를 비교할 수 없거나 어떤 방식으로든 표준화된 변수가 아니면 이 옵션을 지정해서는 안 됩니다. ■ 비수정 상관계수 - 상관행렬에서 주성분을 계산합니다. 하지만 모델에서 절편이 생략되므로, 상관행렬과 표준편차가 평균에 대해 수정되지 않을 수도 있습니다. ■ 비수정 공분산 - 공분산행렬에서 주성분을 계산합니다. 하지만 모델에서 절편이 생략되므로, 공분산행렬과 표준편차가 평균에 대해 수정되지 않을 수도 있습니다.
계산되는 주성분	기본 설정은 변수 개수입니다. 위쪽/아래쪽 화살표를 클릭하여 이 숫자를 변경할 수 있습니다.
주성분 명명에 사용되는 접두어	주성분에 대한 접두어를 지정합니다. 지정한 이름은 최대 26자까지 허용됩니다. 접두어를 지정하지 않으면 기본값이 작업에 사용됩니다. 기본적으로 주성분의 이름을 지정하기 위한 접두어는 PRIN1, PRIN2, ..., PRINn입니다.
비정칙성 기준	$0 < p < 1$ 이면 비정칙성 기준을 지정합니다. 부분변수 역할에 할당된 변수의 PARTIAL 문에서 앞에 나열된 변수에서 예측될 때 R^2 이 $1-p$ 인 경우, 해당 변수에는 표준화된 계수 0이 할당됩니다. 기본값은 1E-08 입니다.
분산 계산에 사용할 분모	분산 계산에 사용되는 기본 분모는 자유도입니다. 분산 계산에 사용할 분모 드롭다운 리스트에서 방법을 선택하여 다른 분모를 선택할 수 있습니다. 주: 분할을 위해서는 먼저 편상관변수 역할에 변수를 할당해야 합니다.

주성분분석: 도표 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

옵션 이름	설명
스크리 도표 및 분산 도표 생성	고윳값의 스크리 도표와 분산 비율의 도표를 생성합니다. 스크리 도표를 사용하여 분석에 사용할 성분의 개수를 결정할 수 있습니다. 고윳값은 각 주성분에 해당하며, 표본에서 총 변량이 차지하는 부분을 나타냅니다. 고윳값은 처음 몇 개가 대부분의 변량을 형성하며 큰 값에서 작은 값 순으로 정렬됩니다. 일반적으로 스크리 도표는 처음 몇 개의 고윳값

옵션 이름	설명
	에 대해 급강하가 나타난 후, 나머지 값에서는 평평한 곡선으로 나타납니다. 전이 영역에서 고윳값 개수를 사용하여 포함할 적절한 성분 개수를 결정할 수 있습니다.
주성분 점수의 산점도 행렬 생성	주성분 점수의 산점도 행렬을 생성합니다. 각 성분의 히스토그램은 행렬의 대각 원소에 표시됩니다.
패턴 프로파일 도표 생성	패턴 프로파일 도표를 생성합니다. 각 성분에 대해 하나의 프로파일이 있습니다. Y축 값은 변수와 주성분의 상관관계입니다.
주성분 점수의 산점도 생성	각 주성분 점수의 산점도를 생성합니다. 새로운 관측값의 주성분 점수에 대한 예측 타원을 표시하도록 선택하여 이 도표를 사용자 정의할 수 있습니다. 기본적으로 타원은 표시되지 않습니다.
패턴 성분 도표 생성	쌍별 성분 패턴 도표를 생성합니다. 도표의 각 관측값은 변수와 두 해당 성분의 상관관계입니다. 패턴을 벡터로 표시할 것인지를 선택할 수 있습니다. 벡터를 표시하도록 선택하면 기본적으로 100% 분산 원이 있는 단위 원이 벡터 패턴 도표에 대해 도표화됩니다.

주성분분석: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
원 데이터 및 성분 점수를 포함하는 테이블 생성	원 데이터 및 성분 점수를 포함하는 테이블 생성 옵션은 입력 데이터와 더불어 주성분 점수를 포함하는 테이블을 생성합니다. 주: SAS Enterprise Guide는 옵션 창의 작업> 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
통계량을 가지는 테이블 생성	통계량을 가지는 테이블 생성 옵션을 선택하면 평균, 표준편차, 관측값 개수, 상관계수 또는 공분산, 고윳값 및 고유벡터가 포함된 테이블이 생성됩니다.
주성분 점수 표준화	주성분 점수를 동일한 분산으로 표준화하려면 주성분 점수 표준화 체크박스를 선택합니다. 이 옵션은 원 데이터 및 성분 점수를 포함하는 테이블 생성 상자에서 지정된 테이블에 영향을 줍니다.

옵션 이름	설명
리포트 및 도표 숨기기	결과를 표시하지 않습니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식

■ 결과의 그래픽 출력에 대한 출력형식

.....
주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.
.....

확률 도표

확률 도표 작업 정보	419
공정능력분석: 분석 역할에 변수 할당	420
확률 도표: 분포 선택	421
확률 도표: 분포 사용자 정의	422
공정능력분석: 도표 모양 옵션 설정	422
축 옵션 설정	422
참조선 추가	423
인셋 추가	423
기타 도표 옵션 설정	423
규격 한계 옵션 설정	425
공정능력분석: 분석 유형 지정	425
제목 및 각주 지정	426
속성 보기	426

확률 도표 작업 정보

공정능력분석은 통계적 관리 상태 공정의 출력분포를 해당 규격한계와 비교하여 규격과 일치하는지를 파악합니다. 공정능력분석에서는 확률 도표를 사용하여 변수값을 정규분포와 같은 특정 이론분포의 백분위수와 비교합니다. 데이터 분포와 이론 분포가 일치하면 도표의 점들이 직선 형태를 이룹니다. 따라서 확률 도표를 사용하여 이론 분포가 측정값을 얼마나 잘 반영하는지 파악할 수 있습니다.

주: 확률 도표는 Q-Q 도표와 유사합니다. 확률 도표는 백분위수의 그래픽 추정에 적합한 반면, Q-Q 도표는 분포모수 및 공정능력지수의 그래픽 추정에 적합합니다.

확률 도표를 사용하여 제조 과정을 분석할 수 있습니다. 예를 들어 50개의 강철 봉의 직경을 측정하여 데이터셋에 저장했다고 가정합니다. 봉의 생산 과정은 통계적 관리 상태에 있으며 직경이 정규분포를 이루는지 알아보려고 합니다. 점 패턴의 선형성은 측정값들이 정규분포를 이루는지를 나타냅니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	CAPABILITY
필수 SAS 제품	Base SAS, SAS/QC, SAS/GRAPH
추가 SAS 제품(권장)	SAS/STAT

공정능력분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분석변수	분석할 숫자 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다. P-P 도표를 제외한 모든 도표에서는 변수 이름의 규격한계 영역에서 각 분석변수에 대한 상한, 목표 값 및 하한을 지정할 수 있습니다. 이 한계를 지정하면 정규성 검정을 보여 주는 테이블, 규격한계를 보여 주는 테이블 및 공정능력지수를 보여 주는 테이블이 자동으로 출력에 포함됩니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n은 해당 행의 빈도변수값입니다.
상대 가중값 변수	상대 가중값으로 사용할 변수를 지정합니다. 이 역할에 변수를 할당하면 각 통계량의 변수값이 가중 통계량 계산에 사용됩니다. 이 역할에는 변수를 하나만 할당할 수 있습니다.

역할 이름	설명
	주: 히스토그램, 확률 도표, P-P 도표, Q-Q 도표 또는 CDF 도표를 생성할 때는 상대 가중값 변수 역할에 변수를 할당하지 않는 것이 좋습니다. 이 역할에 변수를 할당했으면 출력을 해석할 때 주의해야 합니다.
분류변수 (히스토그램 작업에 서만 사용 가능)	분류변수로 사용할 하나 이상의 변수를 지정합니다. 이 역할에 변수를 할당하면 전체 분류 레벨에서 분석변수의 분포를 비교할 수 있는 비교 히스토그램이 생성됩니다. 주: 비교 확률 도표 또는 비교 Q-Q 도표를 생성하려면 분포분석 작업을 사용하십시오.

확률 도표: 분포 선택

이 옵션에 액세스하려면 선택 영역의 분포 머리글 아래에서 **요약**을 클릭합니다.

확률 도표에는 하나의 이론분포를 적용할 수 있습니다. 분포를 선택한 후 해당 분포의 모수 값 및 선 속성을 지정하려면 선택 영역에서 분포 이름을 선택합니다.

옵션 이름	설명
분포	체크 박스를 사용하여 다음 추정된 밀도 곡선을 하나 이상 선택합니다. ■ 정규 - 평균(μ)과 표준편차(σ)를 가지는 정규밀도함수를 적합시킵니다. ■ 로그 정규 - 척도모수(ζ), 임계모수(θ) 및 형상모수(σ)를 가지는 로그 정규 밀도함수를 적합시킵니다. ■ 지수 - 척도모수(σ) 및 임계모수(θ)를 가지는 지수밀도함수를 적합시킵니다. ■ 베타 - 척도모수(σ), 임계모수(θ) 및 형상모수(α 및 β)를 가지는 베타 밀도함수를 적합시킵니다. ■ 감마 - 척도모수(σ), 임계모수(θ) 및 형상모수(α)를 가지는 감마 밀도함수를 적합시킵니다. ■ 와이블(3-모수) - 척도모수(σ), 임계모수(θ) 및 형상모수(c)를 가지는 3-모수 와이블 밀도함수를 적합시킵니다. ■ 와이블(2-모수) - 척도모수(σ) 및 형상모수(c)를 가지는 2-모수 와이블 밀도함수를 적합시킵니다. 2-모수 와이블에서는 데이터의 한 임계모수(θ)를 0으로 가정합니다.
그래프 스타일	출력에 사용할 그래프 스타일을 선택합니다. 기존 SAS 그래프나 ODS 통계 그래프 중에 선택할 수 있습니다. ODS 그래프에 대한 자세한 내용은 SAS Output Delivery System: User's Guide 를 참조하십시오.

확률 도표: 분포 사용자 정의

이 옵션에 액세스하려면 선택 영역의 분포 머리글 아래에서 사용자 정의할 분포 이름을 클릭합니다.

주: 이러한 분포 옵션은 **요약** 페이지에서 분포를 선택했거나 해당 분포 페이지에서 **분포 이름** 체크 박스를 선택할 때만 사용할 수 있습니다.

정규, 로그 정규, 지수, 와이블, 베타 및 감마 분포에서는 각 분석변수 또는 모든 분석변수에 대해 모수를 지정할 수 있습니다. 기본적으로 작업에서는 모든 분석변수의 모수 값을 추정합니다.

각 분석변수에 대한 모수를 지정하는 방법:

- 1 **모든 변수에 분포 적용** 체크 박스의 선택을 해제합니다.
- 2 **분석변수** 상자에서 변수에 해당되는 체크 박스를 선택합니다.
- 3 모수값을 지정합니다.

주: 추정 모수값을 생성하려면 **추정값 사용**을 클릭합니다.

공정능력분석: 도표 모양 옵션 설정

축 옵션 설정

이 옵션에 액세스하려면 선택 영역의 **모양** ⇨ **축** 머리글 아래에서 **축**을 클릭합니다.

다음 옵션을 설정할 수 있습니다.

- 축 색상
- 눈금 표시 색상
- 가로축 및 세로축 레이블

참조선 추가

이 옵션에 액세스하려면 선택 영역의 **모양** ⇨ **축** 머리글 아래에서 **가로** 및 **세로**를 클릭합니다.

참조선을 사용하려면 **참조선 사용**을 선택합니다. 참조선의 출력형식을 지정하고 각 참조선의 위치를 지정할 수 있습니다.

인셋 추가

이 옵션에 액세스하려면 선택 영역의 **모양** 머리글 아래에서 **인셋**을 클릭합니다.

요약통계량 상자 또는 테이블(인셋)을 그래프에 직접 추가하여 도표를 보기 좋게 만들 수 있습니다. 비교 히스토그램을 생성하면 각각의 작은 히스토그램에 인셋이 표시됩니다.

인셋 포함을 선택하여 도표에 요약통계량을 추가합니다. 인셋에 포함할 통계량을 선택합니다.

주: 공정능력지수를 계산하기 위한 통계량은 **분석변수** 역할에 할당한 변수에 대한 규격한계를 생성할 때만 사용할 수 있습니다. 히스토그램, CDF 도표, 확률 도표 및 Q-Q 도표에 대한 규격한계를 생성할 수 있습니다. 공정능력지수 계산에 대한 자세한 내용은 CAPABILITY 프로시저에 대한 도움말을 참조하십시오.

기타 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **옵션**을 클릭합니다.

사용할 수 있는 도표 모양 옵션은 작업에서 생성되는 도표의 유형에 따라 다릅니다.

역할 이름	설명
배경	도표의 배경 색상을 지정합니다.
막대 외곽선 및 막대 채우기	히스토그램의 막대 색상을 지정합니다. 주: 이 옵션은 비교 히스토그램에서만 사용할 수 있습니다.
축 바꾸기 : 백분위수를 세로축으로 사용	도표에서 축을 반대로 바꿉니다. 이 옵션을 선택하면 도표의 방향이 바뀌므로 선택한 축 옵션 및 참조선에서 지정한 내용은 적용되지 않습니다. 주: 이 옵션은 확률 도표 및 Q-Q 도표에서만 사용할 수 있습니다.

역할 이름	설명
기본 범례 생략	도표에 기본 범례를 표시하지 않습니다.
중간점 사용	히스토그램에 대한 중간점을 지정합니다. 드롭다운 리스트에서 사용할 수 있는 옵션은 히스토그램을 생성하는지 또는 비교 히스토그램을 생성하는지에 따라 달라집니다. 분류변수 역할에 변수를 할당할 때에는 비교 히스토그램을 생성합니다. 비교 히스토그램에서는 다음 옵션을 사용할 수 있습니다. ■ 중간점 지정 - 시작 중간점과 마지막 중간점을 지정할 수 있습니다. 다른 중간점을 계산하는 데 사용할 단계 간격도 지정할 수 있습니다. 중간점은 각 막대로 표시되는 값 범위의 중간 값입니다. ■ 키 - 키 셀의 데이터에 대해 중간점을 결정합니다. 기본적으로 분류변수의 레벨은 첫 번째 분류변수의 내부(서식화되지 않은) 값의 증가 순으로 위에서 아래로(왼쪽에서 오른쪽으로) 표시됩니다. 분류변수를 하나만 지정하면 키 셀은 이러한 순서로 첫 번째에 나타나는 레벨이 됩니다. 분류변수를 두 개 지정하면 키 셀은 이러한 순서로 첫 번째에 나타나는 변수 1과 변수 2의 레벨 결합이 됩니다. 따라서 키 셀의 선택에 따라 모든 셀에 사용되는 균등 수평 축이 결정됩니다. 그런 다음 키 셀에 대한 중간점 리스트에 나머지 셀의 데이터가 포함될 때까지 리스트가 어느 방향이든지 필요한 만큼 확장됩니다. ■ 균등 - 총 표본 크기를 기반으로 중간점 개수를 결정합니다. 분류변수 역할에 변수를 할당하지 않을 때에는 히스토그램을 생성합니다. 히스토그램에서는 다음 옵션을 사용할 수 있습니다. ■ 중간점 지정 - 시작 중간점과 마지막 중간점을 지정할 수 있습니다. 다른 중간점을 계산하는 데 사용할 단계 간격도 지정할 수 있습니다. 중간점은 각 막대로 표시되는 값 범위의 중간 값입니다. ■ 그룹 수 지정 - 데이터에 대한 막대 수(히스토그램 구간이라고도 함)를 지정할 수 있습니다. 기본적으로 표시되는 구간 수는 데이터 범위에 따라 달라집니다. 그러나 표시할 최대 구간 수는 지정할 수 있습니다. 주: 히스토그램 작업을 사용하는 경우에만 중간점을 지정할 수 있습니다.
사용자 정의 그래프 크기 지정(픽셀)	이미지 크기를 지정합니다. 참조할 수 있도록 전체 화면 크기가 너비 및 높이 상자에 제공됩니다.

주: 비교 옵션은 비교 히스토그램에서만 사용할 수 있습니다. 히스토그램 작업에서 **분류변수** 역할에 변수를 할당하여 비교 히스토그램을 생성할 수 있습니다.

역할 이름	설명
행 및 칼럼	분석 역할에 하나 이상의 분류변수를 할당하면, 페이지당 행 수 및 페이지당 칼럼 수 상자를 사용하여 비교 히스토그램에서의 칼럼과 행 배열 방식을 변경할 수 있습니다. 한 개의 분류변수를 할당하면 페이

역할 이름	설명
	지당 한 개의 칼럼과 두 개의 행이 기본 정렬 방식이 됩니다. 두 개의 분류변수를 할당하면 페이지당 두 개의 행과 두 개의 칼럼이 기본 정렬 방식이 됩니다. 옆 프레임 및 위 프레임 드롭다운 리스트를 사용하여 행 레이블과 칼럼 레이블의 프레임 영역을 채울 색상을 지정할 수도 있습니다.
막대 수	기본적으로 표시되는 구간 수는 데이터 범위에 따라 달라집니다. 그러나 표시할 최대 구간 수는 지정할 수 있습니다. 또한 표시되는 구간 수가 데이터 평균 이상 및 이하의 표준편차 값 범위를 기반으로 하도록 표준편차를 지정할 수 있습니다.

규격 한계 옵션 설정

주: 히스토그램, 확률 도표, Q-Q 도표 및 CDF 도표에서 이러한 옵션을 사용할 수 있습니다.

선택 영역의 **모양** 머리글 아래에서 **규격한계**를 클릭하여 규격하한, 목표 값 및 규격상한을 지정합니다.

주: **규격한계** 옵션은 적어도 하나의 분석변수에 대해 하나 이상의 한계를 지정할 때만 사용할 수 있습니다. 분석변수에 대한 규격한계를 지정하려면 선택 영역에서 **데이터**를 클릭하고 분석변수를 선택합니다.

공정능력분석: 분석 유형 지정

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

공정능력분석 리포트에는 기본적으로 기본 신뢰구간 테이블, 기본 측도 테이블, 위치모수에 대한 검정 테이블, 적률 테이블 등이 포함됩니다.

리포트에 테이블을 추가하려면 **테이블** 상자에서 테이블을 선택합니다.

분석 옵션을 지정하려면 **테이블** 상자에서 테이블을 선택합니다. 옵션을 사용할 수 있으면, 옵션을 지정하는 데 사용한 오른쪽 영역에 상자가 나타납니다.

리포트에서 모든 기술통계량 테이블을 숨기려면 **기술통계량 및 공정능력지수 테이블 생략** 체크 박스를 선택합니다. 이 옵션을 선택해도 다음 테이블은 숨겨지지 않습니다.

- 선택한 분포에 의해 생성된 테이블. 해당 분포에 대한 패널에서 **분포테이블 생략** 체크 박스를 선택하여 이러한 테이블을 숨길 수 있습니다.

- INTERVALS 문에 의해 생성된 테이블. NOPRINT 옵션을 지정하면 이러한 테이블을 생략할 수 있습니다. INTERVALS 문에 대한 자세한 내용은 CAPABILITY 프로시저에 대한 도움말을 참조하십시오.

결과를 출력 데이터셋에 저장하려면 **데이터셋에 출력 통계량 저장** 체크 박스를 선택합니다. 데이터셋에는 일반량 통계량 및 공정능력지수가 포함됩니다.

주: SAS Enterprise Guide는 **옵션 창**의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간

- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 속성 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

비례위험분석

비례위험분석 작업 정보	429
비례위험분석: 분석 역할에 변수 할당	430
비례위험분석: 모델 옵션 설정	431
비례위험분석: 방법 옵션 설정	433
비례위험분석: 도표 생성	434
비례위험분석: 결과 옵션 설정	434
제목 및 각주 지정	435
속성 보기	435

비례위험분석 작업 정보

비례위험분석 작업은 Cox 비례위험분석 모델을 기반으로 생존 데이터에 대한 회귀분석을 수행합니다. 이 준모수적 모형은 설명변수가 생존시간에 미치는 효과를 설명하기 위한 생존 데이터 분석에 널리 사용됩니다.

생존분석은 임상 및 유행(전염)병학 추적 연구에 많이 사용됩니다. 생존분석법을 사용하는 기타 분야로는 사회학, 공학, 및 경제학 등이 있습니다. 분야를 막론하고 생존분석 연구의 공통된 목적은 사건 발생 여부뿐만 아니라 사건이 발생했는지 여부입니다. 예를 들어 수술 후 5년에 사망한 사람에 대한 서브젝트는 수술 후 한 달에 사망한 사람에 대한 서브젝트와 다릅니다. 수술 후 사망자 수에 대한 연구는 생존 시간에 대한 정보를 무시할 수 있습니다. 생존분석은 시간 이외에 결과를 분석할 때에도 사용됩니다. 공학자가 타이어가 고장날 때까지의 주행 거리를 분석하거나, 엔진 수리가 필요할 때까지의 자전거 수를 분석하는 것이 그 예입니다. 이들 연구의 공통점은 사건이 발생할 때까지의 결과를 분석한다는 점입니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
----------	--------------

사용되는 SAS 프로시저	PHREG
---------------	-------

요구 사항 이름	프로시저 및 제품 이름
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

비례위험분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
생존시간	생존시간 변수로 사용할 변수를 지정합니다. 생존시간은 필수이며, 이 역할에는 변수를 하나만 할당할 수 있습니다.
절단 표시 변수	절단 표시 변수로 사용할 변수를 지정합니다. 생존시간을 절단하는 데 사용할 값 리스트를 생성하려면 오른쪽 절단 표시값 영역에서 값을 선택합니다. 사용자 정의 값 입력 상자에서 특정 값을 입력하고 추가를 클릭하면 절단 표시 값 리스트에 값에 포함됩니다. 절단 표시 값 리스트에서 값을 제거하려면 해당 값 옆의 체크 박스의 선택을 해제합니다. 절단 표시값은 숫자여야 하며, 결측값이 아니어야 합니다. 절단 표시 변수는 필수 변수가 아닙니다. 이 역할에는 변수를 하나만 할당할 수 있습니다.
설명변수	모델의 설명변수를 지정합니다.
계층변수	값이 계층 레벨을 결정하는 변수를 지정합니다. 숫자 변수에 대해 끝점을 지정하지 않으면 계층변수 개별값의 레벨 조합에 의해 계층이 결정됩니다. 체크 박스를 클릭하여 결측값이 올바른 계층 레벨을 형성하도록 할 수 있습니다. 그렇지 않으면 결측값이 있는 관측값이 분석에서 제외됩니다. 숫자 계층변수에서는 계층 레벨 영역의 구간 지정을 클릭하여 계층 구간을 지정할 수 있습니다. 그런 다음 끝점을 추가하여 구간을 정의합니다. 기본적으로 고윳값이 계층 레벨을 형성합니다. 이 역할에 여러 변수를 지정하면 각 변수별로 구간을 정의하십시오.

역할 이름	설명
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
ID 변수	생존함수 추정값을 포함하는 출력 데이터셋의 ID 변수로 사용할 하나 이상의 변수를 지정합니다. 출력 데이터셋을 생성하려면 결과 페이지에서 생존함수 추정값을 저장하도록 선택합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.

비례위험분석: 모델 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **모델**을 클릭합니다.

주: 이 옵션은 **설명변수** 역할에 변수를 할당했을 때만 사용할 수 있습니다.

옵션 이름	설명
모형 선택 방법 설정	모형에 나타나는 설명변수를 제어하려면 선택 방법 드롭다운 리스트에서 방법을 선택합니다. 방법을 선택하지 않으면 전체 모형이 사용됩니다.
	전체 모형(선택 없음) 이 방법은 기본값이며 모형 선택 기능을 제공하지 않습니다. 모든 설명변수가 모형에 포함됩니다. 전진선택 전진선택법은 먼저 모델에 들어온 변수의 모수를 추정합니다. 이 변수는 절편이며, 포함할 변수 리스트에 나타나는 처음 n개의 설명변수입니다. 기본 n 값은 0이므로, 모형에 변수가 포함되지 않습니다. 다음으로, 이 방법은 모델에 없는 각 변수에 대해 수정된 카이제곱 통계량을 계산하여 이 통계량의 최댓값을 검사합니다. 이 통계량에 대한 p-값은 모형에서 변수를 포함하도록 지정된 유의수준과 비교됩니다. 기본적으로 이 값은 0.05입니다. 유의수준을 변경하려면 모형에 추가되는 기준 상자에 값을 입력합니다. 이 값보다 큰 유의수준을 가지는 수정된 카이제곱 통계량이 없으면 전진선택은 중단됩니다. 그렇지 않으면, 전진선택법은 가장 큰 수정된 카이제곱 통계량을 가지는 변수를 모형에 추가합니다. 그런 다음, 모델 외부에 남아 있는 변수에 대해 수정된 카이제곱 통계량을 다시 계산하여 평가 절차를 반복합니다. 결국, 유의한 수정된 카이제곱 통계량을 생성하는 변수가 없을 때까지 변수는 모형에 차례로 추가됩니다. 변수가 모형에 추가되면 변수는 모델에서 계속 유지됩니다.

옵션 이름	설명
	후진제거 후진제거법은 먼저 전체 모델에 대해 모수를 추정한 후 포함할 변수 리스트에 설명변수가 없으면 유의하지 않은 변수를 제거합니다. 개별 모수에 대한 Wald 검정 결과를 검토합니다. 모형에 남는 기준 상자에서 지정한 유의수준에 맞지 않는 최소 유의변수는 제거됩니다. (기본적으로 이 값은 0.05입니다.) 변수가 모형에서 제거된 후에는 계속 모형 외부에 남게 됩니다. 이 절차는 지정된 수준과 맞는 변수가 모형에 남지 않을 때까지 반복됩니다. 단계별 선택 단계별 선택법은 전진선택법의 변형으로서, 모델에 이미 있는 변수가 반드시 유지될 필요가 없다는 점이 다릅니다. 변수는 각 전진선택 단계에 이어 후진제거 단계가 한 번 이상 수행되는 방식으로 모형에 추가되거나 모형에서 제거됩니다. 모형에 더 이상 추가되는 변수가 없거나, 모형에 추가된 변수가 이후의 후진제거 단계에서 제거된 유일한 변수가 되면 단계별 선택 과정이 종료됩니다. 최적 서브셋 선택 이 선택 방법은 우도 점수(예측한계제곱) 통계량을 기준으로 모델을 정렬합니다. 하나의 설명변수에서 모든 설명변수를 포함하는 단일 모델에 이르기까지 모든 가능한 모델 크기에 대해 최고 스코어 카이제곱 통계량을 가지는 지정된 모델 수를 식별합니다. 10개 이상의 설명변수가 있으면 각 모형 크기에 대해 선택한 모형의 개수의 최댓값은 설명변수의 총 개수와 같습니다. 설명변수가 10개 이상인 경우 각 모델 크기에 대해 선택된 모델 수는 최대 총 설명변수 수와 같습니다. 가장 작은 모델은 n개의 설명변수를 포함하며, 여기서 n은 포함할 변수 리스트의 설명변수 수에서 모델의 총 설명변수 수까지의 범위입니다. 포함할 변수 리스트가 비어 있으면, n의 값은 1이 됩니다. 주: 절편 모수는 항상 모델에 포함됩니다.
신뢰수준 지정	전진선택법 또는 단계별 선택법에서는 모형에 추가되는 기준 상자에서 지정한 유의수준에 따라 설명변수가 모형에 포함될지 여부가 결정됩니다. 모형에 없는 변수를 검사하여 p-값이 지정된 유의수준과 같거나 작을 때 최소 p-값을 가진 변수가 추가됩니다. 기본 유의수준은 0.05입니다. 후진제거법 또는 단계별 선택법에서는 모형에 남는 기준 상자에서 지정한 유의수준에 따라 설명변수가 모델에 남을지 여부가 결정됩니다. 모형에 있는 변수를 검사하여 최대 p-값이 지정된 유의수준을 초과할 때 최대 p-값을 가진 변수가 제거됩니다. 기본 유의수준은 0.05입니다.
모형에 남는 기준인 변수 선택	변수 선택 과정에서 모형에 포함되거나 제거된 설명변수가 포함할 변수 상자에 표시됩니다. 변수가 항상 모형에 포함되게 하려면 변수 이름 옆의 체크 박스를 선택합니다. 변수가 더 이상 모형에 포함되지 않게 하려면 변수 이름 옆의 체크 박스의 선택을 취소합니다.

비례위험분석: 방법 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **방법**을 클릭합니다.

옵션 이름	설명
위험비에 대한 신뢰 한계 계산	각 설명변수에 대한 신뢰한계를 표시합니다. 신뢰수준 값은 직접 입력하거나 신뢰수준 리스트에서 선택할 수 있습니다. 기본값은 95% 신뢰수준입니다.
모수 추정값의 상관 계수	모수 추정값의 추정된 상관행렬을 표시합니다.
모수 추정값의 공분산	모수 추정값의 추정된 공분산 행렬을 표시합니다.
동일 시간 처리 방법	동일 시간 처리 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ Breslow의 근사 우도법 - Breslow의 근사 우도법을 사용합니다. 동일 값이 많지 않으면 이 방법을 사용할 때 연속적인 시간 척도 모델의 정확 조건부확률 계산에 대한 만족스러운 근사값을 얻을 수 있습니다. 기본값입니다. ■ 이산 로지스틱 모델 - 비례위험모델을 이산 로지스틱 모델로 바꿉니다. 시간 척도가 완전한 이산형이 아니면 이 방법을 사용해야 합니다. 사례-대조 연구 분석에서 매치된 집합에 둘 이상의 사례가 있으면 이 방법이 필요합니다. ■ Efron의 근사 우도법 - Efron의 근사 우도법을 사용합니다. 동일 값이 많지 않으면 이 방법을 사용할 때 연속적인 시간 척도 모델의 정확 조건부확률 계산에 대한 만족스러운 근사값을 얻을 수 있습니다. ■ 정확 조건부확률 계산 - 모든 동일 이벤트 시간은 동일 값의 절단 시간 전이나 큰 값의 절단 시간 전에 발생한다는 비례위험 가정 하에 정확 조건부확률을 계산합니다. 주: 이 방법은 많은 컴퓨터 리소스를 사용할 수 있습니다. 동일 값이 없으면 네 가지 방법에서 모두 동일한 우도가 발생하여 동일한 추정값이 산출됩니다.

비례위험분석: 도표 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

주: 모델 평가 도표를 생성하려면 먼저 **설명변수** 역할에 변수를 할당해야 합니다.

옵션 이름	설명
누적 마팅게일 잔차 도표	각 설명변수에서 이 옵션은 설명변수 값에 대한 관측된 누적 마팅게일 잔차의 도표를 생성합니다(시뮬레이션된 잔차 패턴 20개 포함). 누적 마팅게일 잔차 패널 도표 옵션을 선택하여 네 개의 패널이 있는 도표를 생성할 수 있습니다. 각 패널에는 관측된 누적 마팅게일 잔차와 두 개의 시뮬레이션된 잔차 패턴이 포함됩니다.
표준화점수과정 도표	각 설명변수에서 이 옵션은 후속 시간에 대한 관측된 점수과정 컴포넌트의 도표를 생성합니다(시뮬레이션된 패턴 20개 포함).
생존함수 도표	이 옵션은 설명변수의 참조 집합에 대한 추정생존함수를 도표로 생성합니다. 이 집합은 해당 표본 평균으로 구성됩니다.
누적 위험함수 도표	이 옵션은 설명변수의 참조 집합에 대한 추정누적위험함수를 도표로 생성합니다. 이 집합은 해당 표본 평균으로 구성됩니다. 동일 값이 없으면 네 가지 방법에서 모두 동일한 우도가 발생하여 동일한 추정값이 산출됩니다.

비례위험분석: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
출력 데이터 저장	결과의 일부 또는 전체를 출력 테이블에 저장합니다. 출력에 포함할 통계량의 체크 박스를 선택합니다. ■ 생존함수 추정값 옵션은 생존함수 추정값, 진단 및 추정되는 선형에 측변수를 포함하는 출력 데이터셋을 생성합니다.

옵션 이름	설명
	■ 회귀계수 추정값 옵션은 회귀계수 추정값을 포함하는 출력 데이터셋을 생성합니다. 공분산행렬을 데이터셋에 포함하려면 추정된 공분산행렬 포함 체크 박스를 선택합니다. ■ 기저생존함수 추정값 옵션은 각 계층에 대한 설명변수의 평균과 일치하는 생존함수 추정값을 포함하는 출력 데이터셋을 생성합니다. 데이터셋에는 생존함수 및 선형예측변수, 표준오차, 신뢰구간, 생존의 로그값 및 음의 로그값의 로그값도 포함됩니다. 주: SAS Enterprise Guide는 옵션 창의 작업> 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
모든 출력 및 도표 숨기기	결과를 표시하지 않습니다. 이 체크 박스를 선택하면 도표 패널의 옵션을 사용할 수 없습니다. 결과는 기본적으로 표시됩니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

Q-Q 도표

Q-Q 도표 작업 정보	437
공정능력분석: 분석 역할에 변수 할당	438
Q-Q 도표: 분포 선택	439
Q-Q 도표: 분포 사용자 정의	440
공정능력분석: 도표 모양 옵션 설정	440
축 옵션 설정	440
참조선 추가	441
인셋 추가	441
기타 도표 옵션 설정	441
규격 한계 옵션 설정	443
공정능력분석: 분석 유형 지정	443
제목 및 각주 지정	444
속성 보기	444

Q-Q 도표 작업 정보

공정능력분석은 통계적 관리 상태 공정의 출력분포를 해당 규격한계와 비교하여 규격과 일치하는지를 파악합니다. 공정능력분석에서는 분위수 - 분위수 도표(Q-Q 도표)를 사용하여 순서화된 변수값을 정규분포와 같은 특정 이론분포의 분위수와 비교합니다. 데이터 분포와 이론 분포가 일치하면 도표의 점들이 직선 형태를 이룹니다. 따라서 Q-Q 도표를 사용하여 이론분포가 측정값을 얼마나 잘 반영하는지 파악할 수 있습니다.

주: Q-Q 도표는 확률도표와 유사합니다. Q-Q 도표는 분포모수에 대한 그래프 추정에 적합하며 확률 도표는 백분위수에 대한 그래프 추정에 적합합니다.

Q-Q 도표를 사용하여 제조 과정을 분석할 수 있습니다. 예를 들어 50개의 철판을 절단하여 두 개의 구멍 사이의 간격을 측정하고 데이터셋에 저장했다고 가정합니다. 절단 과정은 통계적 관리 상태에 있으며 제조분포가 정규분포를 이루는지 알아보려고 합니다. Q-Q 도표는 데이터셋의 관측값을 정규분포의 분위수와 비교합니다. 점 패턴의 선형성은 측정값들이 정규분포를 이루는지를 나타냅니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	CAPABILITY
필수 SAS 제품	Base SAS, SAS/QC, SAS/GRAPH
추가 SAS 제품(권장)	SAS/STAT

공정능력분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분석변수	분석할 숫자 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다. P-P 도표를 제외한 모든 도표에서는 변수 이름의 규격한계 영역에서 각 분석변수에 대한 상한, 목표 값 및 하한을 지정할 수 있습니다. 이 한계를 지정하면 정규성 검정을 보여 주는 테이블, 규격한계를 보여 주는 테이블 및 공정능력지수를 보여 주는 테이블이 자동으로 출력에 포함됩니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
상대 가중값 변수	상대 가중값으로 사용할 변수를 지정합니다. 이 역할에 변수를 할당하면 각 통계량의 변수값이 가중 통계량 계산에 사용됩니다. 이 역할에는 변수를 하나만 할당할 수 있습니다.

역할 이름	설명
	주: 히스토그램, 확률 도표, P-P 도표, Q-Q 도표 또는 CDF 도표를 생성할 때는 상대 가중값 변수 역할에 변수를 할당하지 않는 것이 좋습니다. 이 역할에 변수를 할당했으면 출력을 해석할 때 주의해야 합니다.
분류변수 (히스토그램 작업에 서만 사용 가능)	분류변수로 사용할 하나 이상의 변수를 지정합니다. 이 역할에 변수를 할당하면 전체 분류 레벨에서 분석변수의 분포를 비교할 수 있는 비교 히스토그램이 생성됩니다. 주: 비교 확률 도표 또는 비교 Q-Q 도표를 생성하려면 분포분석 작업을 사용하십시오.

Q-Q 도표: 분포 선택

이 옵션에 액세스하려면 선택 영역의 분포 머리글 아래에서 **요약**을 클릭합니다.

Q-Q 도표에는 하나의 이론분포를 적합시킬 수 있습니다. 분포를 선택한 후 해당 분포의 모수값 및 선 속성을 지정하려면 선택 영역에서 분포 이름을 선택합니다.

옵션 이름	설명
분포	체크 박스를 사용하여 다음 추정된 밀도 곡선을 하나 이상 선택합니다. ■ 정규 - 평균(μ)과 표준편차(σ)를 가지는 정규밀도함수를 적합시킵니다. ■ 로그 정규 - 척도모수(ζ), 임계모수(θ) 및 형상모수(σ)를 가지는 로그 정규 밀도함수를 적합시킵니다. ■ 지수 - 척도모수(σ) 및 임계모수(θ)를 가지는 지수밀도함수를 적합시킵니다. ■ 베타 - 척도모수(σ), 임계모수(θ) 및 형상모수(α 및 β)를 가지는 베타 밀도함수를 적합시킵니다. ■ 감마 - 척도모수(σ), 임계모수(θ) 및 형상모수(α)를 가지는 감마 밀도함수를 적합시킵니다. ■ 와이블(3-모수) - 척도모수(σ), 임계모수(θ) 및 형상모수(c)를 가지는 3-모수 와이블 밀도함수를 적합시킵니다. ■ 와이블(2-모수) - 척도모수(σ) 및 형상모수(c)를 가지는 2-모수 와이블 밀도함수를 적합시킵니다. 2-모수 와이블에서는 데이터의 하한 임계모수(θ)를 0으로 가정합니다.
그래프 스타일	출력에 사용할 그래프 스타일을 선택합니다. 기존 SAS 그래프나 ODS 통계 그래프 중에 선택할 수 있습니다. ODS 그래프에 대한 자세한 내용은 SAS Output Delivery System: User's Guide를 참조하십시오.

Q-Q 도표: 분포 사용자 정의

이 옵션에 액세스하려면 선택 영역의 분포 머리글 아래에서 사용자 정의할 분포 이름을 클릭합니다.

주: 이러한 분포 옵션은 **요약** 페이지에서 분포를 선택했거나 해당 분포 페이지에서 **분포 이름** 체크 박스를 선택할 때만 사용할 수 있습니다.

정규, 로그 정규, 지수, 와이블, 베타 및 감마 분포에서는 각 분석변수 또는 모든 분석변수에 대해 모수를 지정할 수 있습니다. 기본적으로 작업에서는 모든 분석변수의 모수 값을 추정합니다.

각 분석변수에 대한 모수를 지정하는 방법:

- 1 **모든 변수에 분포 적용** 체크 박스의 선택을 해제합니다.
- 2 **분석변수** 상자에서 변수에 해당되는 체크 박스를 선택합니다.
- 3 모수값을 지정합니다.

주: 추정 모수값을 생성하려면 **추정값 사용**을 클릭합니다.

공정능력분석: 도표 모양 옵션 설정

축 옵션 설정

이 옵션에 액세스하려면 선택 영역의 **모양** ⇨ **축** 머리글 아래에서 **축**을 클릭합니다.

다음 옵션을 설정할 수 있습니다.

- 축 색상
- 눈금 표시 색상
- 가로축 및 세로축 레이블

참조선 추가

이 옵션에 액세스하려면 선택 영역의 **모양** ⇨ **축** 머리글 아래에서 **가로** 및 **세로**를 클릭합니다.

참조선을 사용하려면 **참조선 사용**을 선택합니다. 참조선의 출력형식을 지정하고 각 참조선의 위치를 지정할 수 있습니다.

인셋 추가

이 옵션에 액세스하려면 선택 영역의 **모양** 머리글 아래에서 **인셋**을 클릭합니다.

요약통계량 상자 또는 테이블(인셋)을 그래프에 직접 추가하여 도표를 보기 좋게 만들 수 있습니다. 비교 히스토그램을 생성하면 각각의 작은 히스토그램에 인셋이 표시됩니다.

인셋 포함을 선택하여 도표에 요약통계량을 추가합니다. 인셋에 포함할 통계량을 선택합니다.

주: 공정능력지수를 계산하기 위한 통계량은 **분석변수** 역할에 할당한 변수에 대한 규격한계를 생성할 때만 사용할 수 있습니다. 히스토그램, CDF 도표, 확률 도표 및 Q-Q 도표에 대한 규격한계를 생성할 수 있습니다. 공정능력지수 계산에 대한 자세한 내용은 CAPABILITY 프로시저에 대한 도움말을 참조하십시오.

기타 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **옵션**을 클릭합니다.

사용할 수 있는 도표 모양 옵션은 작업에서 생성되는 도표의 유형에 따라 다릅니다.

역할 이름	설명
배경	도표의 배경 색상을 지정합니다.
막대 외곽선 및 막대 채우기	히스토그램의 막대 색상을 지정합니다. 주: 이 옵션은 비교 히스토그램에서만 사용할 수 있습니다.
축 바꾸기 : 백분위수를 세로축으로 사용	도표에서 축을 반대로 바꿉니다. 이 옵션을 선택하면 도표의 방향이 바뀌므로 선택한 축 옵션 및 참조선에서 지정한 내용은 적용되지 않습니다. 주: 이 옵션은 확률 도표 및 Q-Q 도표에서만 사용할 수 있습니다.

역할 이름	설명
기본 범례 생략	도표에 기본 범례를 표시하지 않습니다.
중간점 사용	히스토그램에 대한 중간점을 지정합니다. 드롭다운 리스트에서 사용할 수 있는 옵션은 히스토그램을 생성하는지 또는 비교 히스토그램을 생성하는지에 따라 달라집니다. 분류변수 역할에 변수를 할당할 때에는 비교 히스토그램을 생성합니다. 비교 히스토그램에서는 다음 옵션을 사용할 수 있습니다. ■ 중간점 지정 - 시작 중간점과 마지막 중간점을 지정할 수 있습니다. 다른 중간점을 계산하는 데 사용할 단계 간격도 지정할 수 있습니다. 중간점은 각 막대로 표시되는 값 범위의 중간 값입니다. ■ 키 - 키 셀의 데이터에 대해 중간점을 결정합니다. 기본적으로 분류변수의 레벨은 첫 번째 분류변수의 내부(서식화되지 않은) 값의 증가 순으로 위에서 아래로(왼쪽에서 오른쪽으로) 표시됩니다. 분류변수를 하나만 지정하면 키 셀은 이러한 순서로 첫 번째에 나타나는 레벨이 됩니다. 분류변수를 두 개 지정하면 키 셀은 이러한 순서로 첫 번째에 나타나는 변수 1과 변수 2의 레벨 결합이 됩니다. 따라서 키 셀의 선택에 따라 모든 셀에 사용되는 균등 수평 축이 결정됩니다. 그런 다음 키 셀에 대한 중간점 리스트에 나머지 셀의 데이터가 포함될 때까지 리스트가 어느 방향이든지 필요한 만큼 확장됩니다. ■ 균등 - 총 표본 크기를 기반으로 중간점 개수를 결정합니다. 분류변수 역할에 변수를 할당하지 않을 때에는 히스토그램을 생성합니다. 히스토그램에서는 다음 옵션을 사용할 수 있습니다. ■ 중간점 지정 - 시작 중간점과 마지막 중간점을 지정할 수 있습니다. 다른 중간점을 계산하는 데 사용할 단계 간격도 지정할 수 있습니다. 중간점은 각 막대로 표시되는 값 범위의 중간 값입니다. ■ 그룹 수 지정 - 데이터에 대한 막대 수(히스토그램 구간이라고도 함)를 지정할 수 있습니다. 기본적으로 표시되는 구간 수는 데이터 범위에 따라 달라집니다. 그러나 표시할 최대 구간 수는 지정할 수 있습니다. 주: 히스토그램 작업을 사용하는 경우에만 중간점을 지정할 수 있습니다.
사용자 정의 그래프 크기 지정(픽셀)	이미지 크기를 지정합니다. 참조할 수 있도록 전체 화면 크기가 너비 및 높이 상자에 제공됩니다.

주: 비교 옵션은 비교 히스토그램에서만 사용할 수 있습니다. 히스토그램 작업에서 **분류변수** 역할에 변수를 할당하여 비교 히스토그램을 생성할 수 있습니다.

역할 이름	설명
행 및 칼럼	분석 역할에 하나 이상의 분류변수를 할당하면, 페이지당 행 수 및 페이지당 칼럼 수 상자를 사용하여 비교 히스토그램에서의 칼럼과 행 배열 방식을 변경할 수 있습니다. 한 개의 분류변수를 할당하면 페이

역할 이름	설명
	지당 한 개의 칼럼과 두 개의 행이 기본 정렬 방식이 됩니다. 두 개의 분류변수를 할당하면 페이지당 두 개의 행과 두 개의 칼럼이 기본 정렬 방식이 됩니다. 옆 프레임 및 위 프레임 드롭다운 리스트를 사용하여 행 레이블과 칼럼 레이블의 프레임 영역을 채울 색상을 지정할 수도 있습니다.
막대 수	기본적으로 표시되는 구간 수는 데이터 범위에 따라 달라집니다. 그러나 표시할 최대 구간 수는 지정할 수 있습니다. 또한 표시되는 구간 수가 데이터 평균 이상 및 이하의 표준편차 값 범위를 기반으로 하도록 표준편차를 지정할 수 있습니다.

규격 한계 옵션 설정

주: 히스토그램, 확률 도표, Q-Q 도표 및 CDF 도표에서 이러한 옵션을 사용할 수 있습니다.

선택 영역의 **모양** 머리글 아래에서 **규격한계**를 클릭하여 규격하한, 목표 값 및 규격상한을 지정합니다.

주: **규격한계** 옵션은 적어도 하나의 분석변수에 대해 하나 이상의 한계를 지정할 때만 사용할 수 있습니다. 분석변수에 대한 규격한계를 지정하려면 선택 영역에서 **데이터**를 클릭하고 분석변수를 선택합니다.

공정능력분석: 분석 유형 지정

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

공정능력분석 리포트에는 기본적으로 기본 신뢰구간 테이블, 기본 측도 테이블, 위치모수에 대한 검정 테이블, 적률 테이블 등이 포함됩니다.

리포트에 테이블을 추가하려면 **테이블** 상자에서 테이블을 선택합니다.

분석 옵션을 지정하려면 **테이블** 상자에서 테이블을 선택합니다. 옵션을 사용할 수 있으면, 옵션을 지정하는 데 사용한 오른쪽 영역에 상자가 나타납니다.

리포트에서 모든 기술통계량 테이블을 숨기려면 **기술통계량 및 공정능력지수 테이블 생략** 체크 박스를 선택합니다. 이 옵션을 선택해도 다음 테이블은 숨겨지지 않습니다.

- 선택한 분포에 의해 생성된 테이블. 해당 분포에 대한 패널에서 **분포테이블 생략** 체크 박스를 선택하여 이러한 테이블을 숨길 수 있습니다.

- INTERVALS 문에 의해 생성된 테이블. NOPRINT 옵션을 지정하면 이러한 테이블을 생략할 수 있습니다. INTERVALS 문에 대한 자세한 내용은 CAPABILITY 프로시저에 대한 도움말을 참조하십시오.

결과를 출력 데이터셋에 저장하려면 **데이터셋에 출력 통계량 저장** 체크 박스를 선택합니다. 데이터셋에는 일반량 통계량 및 공정능력지수가 포함됩니다.

주: SAS Enterprise Guide는 **옵션 창**의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간

- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 속성 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

레이더 그래프

레이더 그래프 작업 정보	447
레이더 그래프: 분석 역할에 변수 할당	448
레이더 그래프: 모양 옵션 설정	449
레이더 그래프 옵션 설정	449
추가 모양 옵션	449
제목 및 각주 지정	449
속성 보기	450

레이더 그래프 작업 정보

레이더 그래프 작업은 품질 관리 또는 시장 조사 문제에서 나타나는 데이터 측정값의 상대적인 빈도를 보여 주는 레이더(또는 별) 그래프를 생성합니다. 그래프 통계량은 그래프 중심에서 뻗어나간 축을 따라 표시됩니다.

예를 들어 레이더 그래프는 세 곳의 제조 현장에 대한 총 매출을 보여 줄 수 있습니다. 각 현장에 대한 매출은 돌기의 길이로 나타납니다.

Java 애플릿 또는 Java 이미지 출력 출력형식을 선택한 경우에는 방사형 그래프를 생성할 수 없습니다.

자세한 내용은 [“그래픽 출력형식 선택” \(5페이지\)](#)에서 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GRADAR
필수 SAS 제품	Base SAS, SAS/GRAPH
추가 SAS 제품(권장)	없음

레이더 그래프: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
그래프 칼럼	이 역할에 할당하는 칼럼의 값에 따라 레이더 그래프의 스포크가 결정됩니다. 평면을 정의하기 위해서는 3개의 점이 필요하므로 이 칼럼에는 최소한 3개의 관측값이 있어야 합니다.
그룹 분석 칼럼	이 역할에 할당하는 칼럼의 값에 따라 생성되는 그래프의 수가 결정됩니다. 각 고유타값에 대해 별도의 그래프가 생성됩니다.
빈도 수	이 역할에 할당하는 칼럼의 값에 따라 그래프 칼럼 역할 값의 빈도 (발생 횟수)가 결정됩니다. 이 칼럼은 숫자 칼럼이어야 합니다. 이 역할을 할당하지 않으면 그래프 칼럼 의 각 값은 정확히 한 번씩 계산됩니다.
중첩 칼럼	이 역할에 할당하는 칼럼의 값은 지정된 칼럼의 레벨을 사용하여 비교 레이더 그래프를 생성합니다. 모든 그래프는 동일한 스포크 집합으로 표시됩니다. 수평 칼럼 역할 또는 수직 칼럼 역할에 칼럼을 할당하면 이 역할을 사용할 수 없습니다.
수평 칼럼	이 역할에 할당하는 칼럼은 지정된 칼럼의 각 값에 대한 레이더 그래프를 생성합니다. 그래프 영역의 왼쪽에서 오른쪽 순서로 그래프가 그려집니다. 수직 칼럼 역할에도 칼럼을 할당하면 왼쪽에서 오른쪽 및 위쪽에서 아래쪽 순서로 그래프가 그려집니다.
수직 칼럼	이 역할에 할당하는 칼럼은 지정된 칼럼의 각 값에 대한 레이더 그래프를 생성합니다. 그래프 영역의 위쪽에서 아래쪽으로 그래프가 그려집니다. 수평 칼럼 역할에도 칼럼을 할당하면 왼쪽에서 오른쪽 및 위쪽에서 아래쪽 순서로 그래프가 그려집니다.

레이더 그래프: 모양 옵션 설정

레이더 그래프 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **레이더**를 클릭합니다.
그래프 스타일을 선택합니다.

- **Corona** - 내부 원에서 발산하는 별 꼭짓점을 가진 다각형을 생성합니다.
- **Polygon** - 닫힌 다각형을 생성합니다.
- **Radial** - 중심에서 발산하는 광선을 생성합니다.
- **Spoke** - 내부 원에서 발산하는 광선을 생성합니다.
- **Wedge** - 중심에서 전체 스포크 길이(기본값)까지 방사되는 광선으로 닫힌 다각형을 생성합니다.

추가 모양 옵션

일부 옵션은 여러 그래프 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 [“그래프 영역 사용자 정의” \(597페이지\)](#)에서 참조하십시오.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

확률표본

확률표본 작업 정보	451
확률표본: 확률표본 생성	451

확률표본 작업 정보

확률표본 작업은 입력 데이터셋 행의 확률표본이 포함된 출력 데이터셋을 생성합니다.

데이터 서브셋이 필요하면 이 작업을 사용할 수 있습니다. 예를 들어 비용 리포트 절차를 향상시키고 가능하면 비용을 줄일 목적으로 직원의 여행 경비를 감사한다고 가정합니다. 비용 리포트를 검사할 모든 리소스를 가지고 있지 않으면 통계량 표본 작업을 사용하여 감사를 위한 비용 리포트를 객관적으로 선택할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	SURVEYSELECT
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	없음

확률표본: 확률표본 생성

- 1 작업 이름을 지정합니다.

주: 이 작업은 **SAS Enterprise Guide**에서만 사용 가능합니다.

2. 작업의 입력 데이터 소스를 선택합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

3. 출력 데이터셋에 포함할 변수를 지정합니다. 기본적으로 출력 데이터셋에는 모든 변수가 포함됩니다.

주: **계층변수** 역할에 변수를 할당하면 **출력변수** 역할에도 변수가 자동으로 할당됩니다. 출력에서 변수를 제거하면 계층변수에서도 변수가 제거됩니다.

4. 원하는 행 개수 또는 원하는 입력 행 백분율로 표본 크기를 지정합니다. 예를 들어 입력 행의 수가 400개인데 3%의 행을 지정하면 결과 표본은 12개의 행을 갖게 됩니다.

주: **계층변수** 역할에 변수를 할당하면 여기서 지정한 표본 크기 값은 전체 입력 데이터셋이 아닌 각 계층에 적용됩니다.

5. 표본추출 방법을 선택합니다. 다음 옵션 중에서 선택할 수 있습니다.

단순 확률표본(중복 없음)

한번 선택된 행은 이후 선택 대상에서 제외됩니다. 따라서 동일한 행이 중복 선택될 수는 없습니다.

비제약 확률표본(중복 허용)

한번 선택된 행을 이후에도 선택할 수 있습니다. 따라서 동일한 행을 중복하여 선택할 수 있습니다. 중복 선택한 행을 출력 데이터셋에 기록하는 방법을 지정할 수 있습니다. 다음 옵션 중에서 선택할 수 있습니다.

- **출력에 각 관측값을 한 번만 표시(중복 제외)** - n 번 선택된 행이 표본에 한 번만 나타납니다. 출력에서 NumberHits 변수(임의표본 작업에서 자동으로 계산됨)는 입력 데이터셋에서 관측값이 발생한 횟수를 나열합니다.

- **출력에 모든 관측값 표시(중복 포함)** - n 번 선택된 행이 표본에 n 번 나타납니다.

6. 계층이라고 하는, 함께 사용할 수 없는 비중복 서브셋으로 입력 데이터셋을 분할하는 데 사용할 변수를 지정합니다. 각 계층은 계층변수의 값 집합으로 정의되며, 각 계층은 개별적으로 표본이 추출됩니다. 완전표본은 모든 층에서 가져온 표본의 합집합입니다.

주: 계층 간 총 표본 크기를 계층의 크기에 비례하여 할당할 수 있습니다.

예를 들어 GENDER 변수가 M 및 F 값을 가질 수 있고, VOTED 변수는 Y 및 N 값을 가질 수 있다고 가정해 봅니다. GENDER와 VOTED를 모두 **계층변수** 역할에 할당하면 입력 데이터셋은 네 개의 계층, 즉 투표한 남성, 투표하지 않은 남성, 투표한 여성, 투표하지 않은 여성으로 분할됩니다.

입력 데이터셋에는 20,000개의 행이 포함되어 있습니다. 데이터 분포는 다음과 같습니다.

- 투표한 남성 7,000명

- 투표하지 않은 남성 4,000명
- 투표한 여성 5,000명
- 투표하지 않은 여성 4,000명

따라서 투표를 한 남성의 비율은 $7,000/20,000=0.35$ 즉, 35%입니다. 표본의 비율은 입력 데이터셋에 있는 계층의 비율을 반영해야 합니다. 예를 들어 표본 데이터셋에 100개의 관측값이 포함되었을 때 입력 데이터셋의 비율을 반영하려면 표본 값의 35%는 투표한 남성 계층에서 선택해야 합니다.

계층을 채울 관측값이 충분하지 않으면 모두 포함됩니다.를 선택하면 지정한 표본 크기보다 계층의 크기가 작으면 작업에서 계층에 있는 모든 관측값을 선택합니다.

- 7 출력 데이터셋의 이름과 위치를 지정합니다. 기본적으로 데이터셋은 Work 라이브러리에서 저장됩니다.
- 8 난수 생성을 위한 초기 시드를 지정합니다. 난수 초깃값을 지정하지 않으면 시스템 시계 기준의 난수 초깃값을 사용하여 표본이 산출됩니다.
- 9 표본 산출에 사용된 난수 초깃값을 포함하는 요약테이블을 생성할지를 지정합니다. 동일한 입력 데이터셋에 이와 동일한 초깃값을 지정하면 동일한 표본을 다시 얻을 수 있습니다.

데이터 순위화

데이터 순위화 작업 정보	455
데이터 순위화: 분석 역할에 변수 할당	456
데이터 순위화: 분석 옵션 설정	456
순위화: 결과 옵션 설정	457
속성 보기	458

데이터 순위화 작업 정보

순위 계산 작업은 SAS 데이터셋의 관측값 전반에 걸쳐 하나 이상의 숫자 변수에 대한 순위를 계산하고 새로운 SAS 데이터셋에 출력으로 보냅니다.

예를 들어 사용자의 회사가 판매한 각 제품에 대해 판매 순위를 계산하려고 합니다. 이때 순위 변수는 제품 판매의 순서를 표시합니다. 가장 높은 판매 숫자를 가진 제품이 처음에 나열됩니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	RANK
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	SAS/STAT

데이터 순위화: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
순위화 칼럼	이 역할에 지정된 각 칼럼에 대해 순위를 계산합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다. 기본적으로, 순위 칼럼의 이름은 <code>rank_칼럼 이름</code> 이며, 여기서 <code>칼럼 이름</code> 은 원본 칼럼의 이름입니다. 새로운 이름을 지정하려면 순위 칼럼 이름 텍스트 상자에 이름을 입력합니다.
순위화 그룹	이 역할에 칼럼을 지정하면 선택한 칼럼을 기준으로 입력 테이블이 정렬되고 각 그룹 내에서 순위가 계산됩니다.

데이터 순위화: 분석 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

옵션 이름	설명
순위화 방법	순위화 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 작은 값에서 큰 값으로 - 원 데이터를 작은 값에서 큰 값으로 순위화합니다. ■ 백분위 순위 - 원 데이터를 100개의 그룹으로 나눈 후, 가장 작은 값을 백분위 값 0, 가장 큰 값을 백분위 값 99로 할당합니다. ■ 십분위 순위 - 원 데이터를 10개의 그룹으로 나눈 후, 가장 작은 값을 십분위 값 0, 가장 큰 값을 십분위 값 9로 할당합니다. ■ 사분위 순위 - 원 데이터를 4개의 그룹으로 나눈 후, 가장 작은 값을 사분위 값 0, 가장 큰 값을 사분위 값 3으로 할당합니다.

옵션 이름	설명
	■ 그룹=n(NTILES) - 원 데이터를 n개의 그룹으로 나눕니다. 가장 작은 값을 0 값을 받고, 가장 큰 값은 $n-1$ 값을 받습니다. 그룹 수 텍스트 상자에 n 값을 지정합니다. ■ 분모가 n인 비율 순위 - 각 순위를 비결측 순위 변수값을 가지는 관측값 개수로 나누어서 순위를 계산합니다. ■ 분모가 $(n+1)$인 비율 순위 - 각 순위를 분모 $n+1$로 나누어 분수 순위를 계산합니다. 여기서 n은 순위 변수의 비결측값을 가진 관측값의 수입니다. ■ 백분율 - 각 순위를 변수의 비결측 값을 가진 관측값 수로 나누고 결과에 100을 곱하여 백분율을 구합니다. ■ 정규 점수(Blom 공식) - Blom 공식을 사용하여 순위에서 정규 점수를 계산합니다. ■ 정규 점수(Tukey 공식) - Tukey 공식을 사용하여 순위에서 정규 점수를 계산합니다. ■ 정규 점수(van der Waerden 공식) - van der Waerden 공식을 사용하여 순위에서 정규 점수를 계산합니다. ■ Savage 점수(지수) - 순위에서 Savage(또는 지수) 점수를 계산합니다.
동일 값 처리 방법	동일 값을 처리하는 방법을 지정합니다. 방법을 선택하지 않으면 평균 값(가운데 순위) 방법이 사용됩니다. ■ 평균값(가운데 순위) - 해당 순위의 평균값을 사용합니다. ■ 높은 순위값 - 해당 순위 중 큰 값을 사용합니다. ■ 낮은 순위값 - 해당 순위 중 작은 값을 사용합니다.
큰 값에서부터 작은 값으로 순위화	순위의 순서를 바꿉니다.

순위화: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

기본적으로, 출력 테이블에는 원본 칼럼과 함께 순위화된 칼럼도 포함됩니다. 원본 칼럼을 순위화된 칼럼으로 교체하려면 **순위화 값 포함** 체크 박스의 선택을 취소합니다.

원래 칼럼을 **순위화 칼럼** 역할에 할당하면 순위 칼럼의 이름이 지정됩니다. 기본적으로, 순위 칼럼의 이름은 `rank_칼럼 이름`이며, 여기서 **칼럼 이름**은 원본 칼럼의 이름입니다. 칼럼 이름을 변경하려면 선택 영역에서 **데이터**를 선택합니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

RFM(Recency, Frequency, Monetary) 분석

<i>RFM(Recency, Frequency, Monetary) 분석 작업 정보</i>	459
<i>RFM(Recency, Frequency, Monetary) 분석: 분석할 데이터 유형 선택</i>	460
<i>RFM(Recency, Frequency, Monetary) 분석: 분석 역할에 변수 할당</i>	461
<i>RFM(Recency, Frequency, Monetary) 분석: 데이터를 구간으로 범주화</i>	462
<i>RFM(Recency, Frequency, Monetary) 분석: 결과 저장</i>	463
<i>RFM(Recency, Frequency, Monetary) 분석: 도표 선택</i>	463

RFM(Recency, Frequency, Monetary) 분석 작업 정보

RFM(Recency, Frequency, Monetary) 분석은 새로운 캠페인이나 제품 제안에 가장 잘 반응할 것 같은 기존 고객을 식별하는 데 사용되는 기법입니다. 이 기법은 직접 마케팅에 공통적으로 사용됩니다.

이 작업의 입력 데이터는 거래 데이터 또는 고객 데이터입니다. RFM(Recency, Frequency, Monetary) 분석 작업은 거래 데이터를 고객 데이터에 집계합니다. 거래 데이터에서 각 레코드는 거래를 나타냅니다.

Recency, Frequency, Monetary 스코어를 결정하는 방법:

- Recency 스코어는 오름차순으로 가장 최근 거래 날짜의 값을 정렬하고 해당 값을 막대로 그룹화하여 결정됩니다. 가장 오래된 날짜를 가진 막대는 Recency 스코어로 1을 할당받습니다.
- Frequency 스코어는 오름차순으로 거래 수 값을 정렬하고 해당 값을 막대로 그룹화하여 결정됩니다. 가장 작은 거래 수를 가진 막대는 Frequency 스코어로 1을 할당받습니다.

- Monetary 스코어는 오름차순으로 총 거래 금액 값을 정렬하고 해당 값을 막대로 그룹화하여 결정됩니다. 가장 적은 금액의 막대는 Monetary 스코어로 1을 할당받습니다.

RFM 스코어는 다음 식을 사용하여 계산됩니다. $RFM \text{ 스코어} = 100 * \text{Recency 스코어} + 10 * \text{Frequency 스코어} + \text{Monetary 스코어}$ 가장 호의적이지 않은 고객 세그먼트는 최저 RFM 스코어를 가집니다. 가장 호의적인 고객 세그먼트는 최고 RFM 스코어를 가집니다. RFM(Recency, Frequency, Monetary) 분석 작업에는 SAS 9.4 이상이 필요합니다.

주: RFM(Recency, Frequency, Monetary Analysis) 작업에는 SAS 9.4 이상이 필요합니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	

RFM(Recency, Frequency, Monetary) 분석: 분석할 데이터 유형 선택

이 옵션에 액세스하려면 선택 영역에서 **RFM 유형**을 클릭합니다.

옵션 이름	설명
거래 레벨 데이터	데이터에서 각 레코드는 거래를 나타냅니다. RFM(Recency, Frequency, Monetary) 분석 작업은 고객 ID별로 개별 거래를 집계합니다.
고객 데이터	각 레코드는 고객을 나타냅니다. RFM(Recency, Frequency, Monetary) 분석 작업에서 데이터를 열기 전에 개별 거래를 결합합니다.

RFM(Recency, Frequency, Monetary) 분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

사용 가능한 분석 역할은 분석하려는 데이터 유형에 따라 달라집니다.

테이블 72.1 거래 레벨 데이터에 대한 옵션

역할 이름	설명
거래 날짜	거래 날짜를 지정하는 변수입니다. 이 정보는 가장 최근 거래 날짜를 결정합니다.
거래 금액	거래 금액을 지정하는 변수입니다. 이 데이터는 총 거래 금액을 계산하는 데 사용되며, 거래 금액의 합, 평균, 중위수 또는 최댓값이 될 수 있습니다.
고객 식별자	고객을 고유하게 식별할 수 있는 숫자 또는 문자열을 포함하는 변수입니다.

Recency, Frequency 및 Monetary 작업이 거래 데이터를 고객 데이터로 집계하기 때문에 거래에 대해 집계 방법을 선택해야 합니다. 기본적으로 집계 방법은 합입니다.

집계에서 불완전 거래 레코드(또는 결측값을 가진 레코드)를 가진 고객을 포함할지를 지정할 수 있습니다.

- 이 옵션을 선택하면, 출력 데이터셋은 불완전 거래 레코드를 가진 고객에 대한 스코어를 포함합니다. 각 고객에 대해서는 거래 날짜 및 거래 금액에 대해 비결측값을 가진 거래 레코드만 집계됩니다. 결측값을 가진 모든 레코드는 무시됩니다.
- 이 옵션을 선택하지 않으면, 출력 데이터셋은 불완전한 거래 레코드를 가진 고객의 스코어를 포함하지 않습니다. 거래 레코드는 완전 레코드를 가지는 고객에 대해서만 집계됩니다.

테이블 72.2 고객 데이터에 대한 옵션

역할 이름	설명
가장 최근 거래	각 고객의 가장 최근 거래 날짜를 지정하는 변수입니다. 이 데이터는 Recency 스코어를 계산하는 데 사용됩니다.
거래 수	각 고객의 거래 수를 지정하는 변수입니다. 이 데이터는 Frequency 스코어를 계산하는 데 사용됩니다.
총 거래 금액	각 고객이 지출한 총 금액을 지정하는 변수입니다. 이 데이터는 Monetary 스코어를 계산하는 데 사용됩니다.
고객 식별자	고객을 고유하게 식별할 수 있는 숫자 또는 문자열을 포함하는 변수입니다.

RFM(Recency, Frequency, Monetary) 분석: 데이터를 구간으로 범주화

이 옵션에 액세스하려면 선택 영역에서 **구간화**를 클릭합니다.

옵션 이름	설명
구간 모델	데이터를 범주화하는 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 네스트 방법은 각 Recency 값에 단순 순위를 할당합니다. 각 Recency 순위 내에서 고객은 빈도 순위로 할당되며, 각 Frequency 순위 내에서 고객은 Monetary 순위로 할당됩니다. 이 방법의 장점은 결합된 RFM 스코어의 균등 분포를 산출한다는 점과 구매 행동(예, Frequency 및 Monetary)이 시간에 따라 변화할 수 있다고 간주하는 점입니다. 이 방법의 단점은 모든 스코어가 다른 스코어에 대해 종속되어 있기 때문에 빈도 및 금융 스코어를 해석하는 데 더 많은 시간이 소요되는 데 있습니다. ■ 독립 방법은 Recency, Frequency 및 Monetary 값을 독립적으로 순위화합니다. 이 방법의 장점은 스코어 해석이 명확한 데 있습니다. 예를 들어 한 고객에 대한 Frequency 스코어가 5이면 다른 고객에 대한 Frequency 스코어도 5와 같음을 의미합니다. 작은 표본에서 이 방법의 단점은 결합된 RFM 스코어의 분포가 덜 균등하다는 것과 일부 세그먼트가 고객을 가지지 않을 확률이 있다는 것입니다.
막대 수	Recency, Frequency 및 Monetary 값에 대한 막대 수를 지정할 수 있습니다. 지정한 값을 곱해 총 막대 수를 생성합니다. 예를 들어 Recency에

옵션 이름	설명
	막대를 5개, Frequency에 막대를 5개, Monetary에 막대를 5개로 지정하면, 총 막대 개수는 125입니다.

RFM(Recency, Frequency, Monetary) 분석: 결과 저장

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
RFM 스코어 저장 위치	기본적으로 RFM(Recency, Frequency, Monetary) 스코어는 Work 라이브러리의 데이터셋에 저장됩니다. 이 위치를 변경하려면, 편집 을 클릭합니다.
출력 데이터셋 변수 수	출력 데이터셋의 변수 이름을 사용자 정의할 수 있습니다.

RFM(Recency, Frequency, Monetary) 분석: 도표 선택

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

옵션 이름	설명
금융 지도	각 Recency 및 Frequency 구간 값 조합에 대해 Monetary 값을 요약하는 열지도를 생성합니다. 이 도표는 기본적으로 표시됩니다.
RFM 분포 도표	막대 값 분포의 균등성을 확인하는 데 사용할 수 있는 막대 그래프를 생성합니다. 이 도표는 기본적으로 표시됩니다.
RFM 세그먼트 요약	막대 값 분포의 균등성을 확인하는 데 사용할 수 있는 테이블을 생성합니다.

자기회귀 오차가 있는 회귀분석

자기회귀 오차가 있는 회귀분석 작업 정보	465
자기회귀 오차가 있는 회귀분석: 분석 역할에 변수 할당	466
자기회귀 오차가 있는 회귀분석: 모델 옵션 설정	467
자기회귀 오차가 있는 회귀분석: 통계량 옵션 설정	468
자기회귀 오차가 있는 회귀분석: 도표 생성	470
자기회귀 오차가 있는 회귀분석: 결과 옵션 설정	471
제목 및 각주 지정	472
속성 보기	472

자기회귀 오차가 있는 회귀분석 작업 정보

자기회귀 오차가 있는 회귀분석 작업은 오차가 자기상관적이거나 이분산적일 때 시계열 데이터에 대한 선형 회귀모델을 추정하고 예측합니다. 이 작업은 독립변수의 미래값이 입력 데이터셋에 포함될 때 예측 및 예측 신뢰한계를 생성합니다. 예측값을 생성할 때 모델의 시계열 부분과 시스템 부분을 사용하기 때문에 유용한 예측 도구입니다. 이 모형은 예측값 생성 시 추세에서 최근 이탈을 고려합니다.

이 작업을 사용하여 다음과 같은 작업을 수행할 수 있습니다.

- 시계열에서 다른 추정 방법 및 모델로부터의 추정값을 비교. 예를 들어 일정 기간 동안의 투자 모형을 비교할 수 있습니다.
- 단기간 예측. 많은 시계열은 평활 모양의 확률보행을 가지므로 높은 양의 자기상관을 제시합니다. 이 작업은 모수를 추가하고 확률보행 성향을 캡처링하여 모델의 적합성을 향상시킵니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
----------	--------------

사용되는 SAS 프로시저	AUTOREG
---------------	---------

요구 사항 이름	프로시저 및 제품 이름
필수 SAS 제품	Base SAS, SAS/ETS
추가 SAS 제품(권장)	없음

자기회귀 오차가 있는 회귀분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
종속변수	회귀분석에 대한 종속변수로 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다.
설명변수	회귀모델에 대한 독립회귀변수 역할을 합니다. 이 역할에 변수를 할당하지 않으면 평균만 일치됩니다. (계열의 자기상관을 구하기 위한 방법입니다.)
시간 ID 변수	행에 날짜를 할당하는 데 사용될 변수를 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다. SAS 날짜, 시간 또는 날짜/시간 출력형식을 갖는 첫 번째 변수가 자동으로 이 역할에 할당됩니다. 시간 ID 변수를 생성할 수도 있습니다. 분석에 대한 그래픽 도표를 생성하려면 이 역할에 변수를 할당해야 합니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 종속변수나 독립변수로 이미 선택한 변수로는 분석을 그룹화할 수 없습니다.

자기회귀 오차가 있는 회귀분석: 모델 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

옵션 이름	설명
방법	모델 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 최대우도추정값 ■ 비조건부 최소제곱추정값 ■ Yule-Walker 추정값 ■ 반복 Yule-Walker 추정값
최대 반복 횟수	최대 반복 횟수를 지정합니다. 기본값은 50입니다.
비결측 데이터의 첫 번째 수열만 사용	결측값이 없는 첫 번째 연속 데이터 수열을 사용하도록 지정합니다. 이 옵션을 선택하지 않으면, 전체 관측값이 사용됩니다.
모델에 절편 모수 포함	결측값이 없는 첫 번째 연속 데이터를 사용하도록 지정합니다. 이 옵션을 선택하지 않으면 모든 전체 관측값이 사용됩니다.
모델에 절편 모수 포함	모델에 절편 모수를 포함하도록 지정합니다. 절편 모수는 기본적으로 포함됩니다. ■ 자기회귀 과정의 차수 체크 박스를 선택하고 상자에 값을 입력하거나 자기회귀 과정의 시차 지정을 선택하고 적합시킬 자기회귀 오차 시차 서브셋을 입력합니다. ■ 유의하지 않은 자기회귀 모수 제거 체크 박스를 선택하여 모델에서 유의도가 가장 낮은 순서대로 모수를 제거합니다(후진제거). 초기 보통최소제곱 추정 이후에 계산되는 Yule-Walker 추정에서 한 번만 제거가 수행됩니다. ■ 유의수준이 다음을 초과하면 제거 상자에 값을 입력하여 후진제거에 사용되는 유의수준 기준을 지정합니다.

자기회귀 오차가 있는 회귀분석: 통계량 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **통계량**을 클릭합니다.

이 통계량 중 하나 이상을 선택할 수 있습니다.

옵션 이름	설명
Durbin-Watson 통계량	최대 n 차수까지 Durbin-Watson 통계량을 표시합니다. 최대 순서까지 포함 상자를 사용하여 n 을 지정합니다. 올바른 값의 범위는 1에서 32,767까지입니다. 기본값은 1입니다.
Durbin-Watson 검정의 주변 확률	큰 표본 크기에 대해 일반화된 Durbin-Watson 검정의 주변 확률 통계량의 p -값을 표시합니다. 계획행렬의 선형화 근사 사용 옵션을 선택하여 자기회귀 오차 과정으로 인해 모델이 비선형일 때 계획행렬의 선형화된 근사를 사용하는 p -값을 계산합니다.
Durbin의 t 통계량	Durbin t 통계량을 표시합니다. 이 통계량은 시차 종속변수가 있을 때 잔차 자기상관을 탐지합니다. 주: 설명변수 역할에 변수를 할당하지 않은 경우 Durbin t 통계량만 계산할 수 있습니다.
Durbin의 h 및 t 통계량	Durbin의 h 또는 t 통계량을 표시할지를 선택할 수 있습니다. 유형 드롭다운 리스트를 사용하여 계산할 통계량을 선택합니다. Durbin h 통계량은 회귀변수에 시차 종속변수가 포함될 때 1차 자기상관이 있는지 결정합니다. 시차 종속변수 드롭다운 리스트에서 이 변수를 선택할 수 있습니다. (이 드롭다운 리스트의 변수는 설명 변수 역할에 할당한 변수입니다.) Durbin h 통계량을 계산할 수 없으면 점근적으로 동등한 t 통계량이 대신 표시됩니다. 주: 설명변수 역할에 변수를 할당하지 않은 경우 Durbin h 또는 t 통계량을 계산하도록 선택할 수 있습니다.
ARMA 오차에 대한 Godfrey의 일반 라그랑지 승수 검정	최대 r 차수까지 ARMA 오차에 대한 Godfrey의 일반 라그랑지 승수 검정을 표시합니다. 최대 차수까지 포함 상자를 사용하여 r 을 지정합니다. 유효한 값의 범위는 1 - 32,767입니다. 기본값은 4입니다.
Jarque-Bera 정규성 검정	회귀잔차에 대한 Jarque-Bera의 정규성 검정을 표시합니다.

옵션 이름	설명
Phillips-Perron 단위근 검정	모델에 회귀변수가 없을 때 Phillips-Perron 단위근 검정을 표시합니다. 가중 분산 추정량에 대한 절단점을 지정하려면 절단 시차 포함 체크 박스를 선택하고 절단 시차 상자에 시차 값을 입력합니다. Phillips-Perron 검정에는 영(0) 평균, 단일 평균 및 결정적 추세의 세 가지 유형이 있습니다. 주: 설명변수 역할에 변수를 할당하면 회귀변수가 모델에 포함되고 대신 'Phillips-Ouliaris 공적분 검정'이 표시됩니다.
ARCH 효과에 대한 Q 와 LM 통계량	이분산성을 검정하는 통계량 테이블을 표시합니다.
모수 추정값의 공분산 추정값	모수 추정값의 공분산 추정값을 표시합니다.
모수 추정값의 상관관계수 추정값	모수 추정값의 상관관계수 추정값을 표시합니다.
편자기상관계수	편자기상관계수를 표시합니다.
변환 계수	첫 번째 p 관측값의 변환 계수를 표시합니다. 이 계수는 스칼라에 자기공분산 Toeplitz 행렬의 Cholesky 근의 역함수를 곱하여 생성됩니다.
자기공분산의 Toeplitz 행렬의 역	Yule-Walker 해에 대한 자기공분산의 Toeplitz 행렬의 역을 표시합니다.
회귀모델의 로그 우도함수	오차가 정규분포를 따른다는 가정 하에서 회귀모델의 로그 우도함수 값을 표시합니다.
비중심화 회귀 R-제곱	비중심화 회귀 R^2 을 표시합니다. 비중심회귀 R^2 은 라그랑지 승수 검정 통계량을 계산하는 데 유용합니다. 대부분의 LM 검정 통계량이 $T \cdot URSQ$ 로 계산되기 때문입니다. 여기서 T 는 추정에 사용된 관측값의 수입니다.
모수 추정값과 목적 함수의 반복 히스토리	각 반복에서 모수 추정값과 목적 함수를 표시합니다. 목적함수는 최대우도법에 대한 완전로그우도함수(full log likelihood function)이지만, 오차제곱합이 비조건부 최소제곱의 목적함수로 생성됩니다. 최대우도추정법에 대해서는 집중우도가 아닌 완전로그우도함수의 값을 표시합니다.
Ramsey의 RESET 검정	Ramsey의 RESET 검정통계량을 표시합니다. 주: 이 옵션은 설명변수 역할에 변수를 할당할 때만 사용할 수 있습니다.

자기회귀 오차가 있는 회귀분석: 도표 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

기본적으로 현재 선택된 데이터에 적용 가능한 모든 도표가 출력에 포함됩니다. 그러나 **사용자 정의 도표 리스트** 옵션을 선택하여 출력에 포함할 도표를 선택할 수 있습니다. 다음 옵션 중에서 선택할 수 있습니다.

옵션 이름	설명
자기상관함수 도표	자기상관함수 도표를 생성합니다.
잔차에 대한 역 자기상관함수 도표	역자기상관함수 도표를 생성합니다.
잔차에 대한 편 자기상관함수 도표	편자기상관함수 도표를 생성합니다.
예측값 및 실제 데이터 값 도표	예측값 대 관측값 도표를 생성합니다.
Cook의 D 통계량 도표	Cook의 D 통계량 대 관측값 번호 도표를 생성합니다. 주: 옵션 패널에서 자기회귀 모델 적합 체크 박스를 선택했으면 이 옵션을 사용할 수 없습니다.
잔차에 대한 Q-Q 도표	잔차의 정규 분위수 도표를 생성합니다.
잔차 도표	잔차 대 예측값 도표를 생성합니다. 잔차는 관측값과 예측값의 차이입니다.
스튜던트화 잔차 도표	스튜던트화 잔차 대 예측값 도표를 생성합니다. 주: 옵션 패널에서 자기회귀 모델 적합 체크 박스를 선택했으면 이 옵션을 사용할 수 없습니다.
백색잡음 (White Noise) 확률 도표	백색잡음(White Noise) 확률 도표를 생성합니다.
잔차의 히스토그램 도표	적합 잔차 히스토그램을 생성합니다.

자기회귀 오차가 있는 회귀분석: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
통계량 포함	관측값 수준 통계량을 저장합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 비조건부 평균의 예측값, 잔차, 신뢰한계 - 모델의 구조적인 부분에서만 형성된 예측값, 잔차, 신뢰상한 및 신뢰하한을 저장합니다. 신뢰수준 드롭다운 리스트를 사용하여 모형의 구조적 부분 추정값에 대한 신뢰한계 크기를 설정할 수 있습니다. ■ 미래 실현값에 대한 예측값, 잔차, 신뢰한계 - 모델의 구조 및 자기회귀 부분 모두에서 형성된 예측값, 잔차, 신뢰상한 및 신뢰하한을 저장합니다. 신뢰수준 드롭다운 리스트를 사용하여 반응시계열의 미래값 추정치에 대한 신뢰구간 크기를 설정할 수 있습니다. ■ 누적합(CUSUM) 및 신뢰수준 - CUSUM 통계량, CUSUMSQ 통계량 및 각각의 신뢰상한과 하한을 저장합니다. 신뢰수준 드롭다운 리스트를 사용하여 상한 및 하한에 대한 유의수준을 설정할 수 있습니다. ■ RRecursive와 BLUS 잔차는 Theil의 BLUS 잔차 값, CUSUM 및 CUSUMSQ 통계량을 계산하는 데 사용되는 재귀 잔차, 재귀 잔차를 계산하는 데 사용되는 예측오차분산(vt)의 일부 값을 저장합니다.
모수 추정값 포함	모델에 사용된 모든 변수가 포함된 데이터셋을 생성합니다. 각 회귀변수에는 해당 회귀 모수의 추정값이 포함됩니다. 모수 추정값에 대한 공분산행렬을 표시하려면 공분산행렬 포함 체크 박스를 선택합니다.
모든 출력 숨기기	결과를 표시하지 않습니다. 주: SAS Enterprise Guide는 옵션 창의 작업> 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

패널 데이터의 회귀분석

패널 데이터의 회귀분석 작업 정보	473
패널 데이터의 회귀분석: 분석 역할에 변수 할당	474
패널 데이터의 회귀분석: 모델 옵션 설정	474
패널 데이터의 회귀분석: 결과 옵션 설정	475
제목 및 각주 지정	476
속성 보기	476

패널 데이터의 회귀분석 작업 정보

패널 데이터의 회귀분석 작업은 시계열 데이터 및 횡단면 데이터를 결합할 때 공통적으로 발생하는 선형 계량경제 모델의 종류를 분석합니다. 이 작업에서는 각 횡단면 단위에 대한 시계열 관측값으로 이루어진 패널 데이터셋을 다룹니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	TSCSREG
필수 SAS 제품	Base SAS, SAS/ETS
추가 SAS 제품(권장)	없음

패널 데이터의 회귀분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
종속변수	회귀분석할 변수를 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다.
설명변수	반응변수를 회귀분석할 독립변수를 지정합니다. 이 역할에는 변수를 하나 이상 지정해야 합니다.
횡단면 ID 변수	횡단면을 정의하는 변수를 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다. 데이터는 이 변수값을 기준으로 자동 정렬됩니다.
시간 ID 변수	각 횡단면의 관측값에 시간 기간을 할당하는 데 사용되는 변수를 지정합니다. 이 역할에는 변수를 하나만 할당해야 합니다. 데이터는 이 변수값을 기준으로 자동 정렬됩니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다.

패널 데이터의 회귀분석: 모델 옵션 설정

선택 영역에서 이 옵션은 **옵션** 머리글 아래에 표시됩니다.

도표 이름	설명
사용할 추정 방법	추정 방법을 지정하려면 원하는 방법의 옵션 버튼을 선택합니다. 모델 회귀모수에 대한 추정 방법의 성능은 모델에서 오차 요소의 통계적 특성에 따라 달라집니다. 기본값은 이원 임의효과 입니다.

도표 이름	설명
	■ 일원 고정 효과 모델은 규격이 관측값이 속한 횡단면에서만 영향을 받습니다. ■ 이원 고정 효과 모델은 규격이 관측값이 속한 횡단면과 시계열 모두에서 영향을 받습니다. ■ 일원 임의효과 모델은 규격이 관측값이 속한 횡단면에서만 영향을 받으며, 효과는 임의로 적용됩니다. 이 경우, 효과는 임의입니다. 밸런스 패널에서는 Fuller 및 Battese 방법이 기본값으로 사용됩니다. 언밸런스 패널에서는 Wansbeek 및 Kapteyn 방법이 기본적으로 사용됩니다. ■ 이원 임의효과(분산 요소 모형) 모델은 규격이 관측값이 속한 횡단면과 시간 기간에서 영향을 받습니다. 이 경우, 효과는 임의입니다. 밸런스 패널에서는 Fuller 및 Battese 방법이 기본값으로 사용됩니다. 언밸런스 패널에서는 Wansbeek 및 Kapteyn 방법이 기본적으로 사용됩니다. ■ 자기회귀 모형(Parks 방법)은 횡단면 간에 동시 상관관계가 있는 1차 자기회귀 오차 구조를 가정합니다. Parks 방법을 선택하면 3차 스플라인 곡선을 사용하여 계열의 결측값을 보간할 수 있습니다. 3차 스플라인은 서로 결합된 3차 다항 함수를 구성하는 분절된 함수로서, 전체 곡선과 1차 및 2차 미분계수는 연속됩니다. 3차 스플라인은 변수의 비결측값에 적합됩니다. 이 곡선은 입력 시계열에 대한 연속 시간 근사를 형성하는 데 사용되며, 출력 시계열은 스플라인 근사에서 생성됩니다. 결측값의에3차 스플라인 보간 체크 박스는 기본적으로 선택됩니다. 주: 결측값에 3차 스플라인 보간 옵션은 주의해서 사용합니다. 보간법은 종종 모델에 편향을 가져오며, 이로 인해 모수 추정값과 가설 검정이 제대로 수행되지 않을 수 있습니다. 이러한 방법에 대한 자세한 내용은 <i>SAS/ETS: User's Guide</i>에서 참조하십시오.
모델에 절편 모수 포함	이 절편 모수를 모델에서 제외하려면 체크 박스에 대한 선택을 취소하십시오. 이 체크 박스는 기본적으로 선택됩니다.

패널 데이터의 회귀분석: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

모수 추정값을 출력 테이블에 저장하려면 **모수 추정값 저장** 체크 박스를 선택합니다. **공분산 행렬**, **상관행렬** 또는 두 행렬을 모두 포함하려면 해당 체크 박스를 선택합니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

결과를 표시하려면 **분석 표시** 체크 박스를 선택합니다. 결과를 표시하지 않으려면 이 체크 박스의 선택을 취소합니다. 결과는 기본적으로 표시됩니다.

결과가 표시되도록 선택했으면 결과에 표시할 항목의 체크 박스를 선택합니다. 일부 항목은 Parks 추정 방법을 선택할 때만 사용할 수 있습니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간

- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 속성 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

SAS ODS 통계적 그래프

SAS ODS 통계적 그래프 정보	479
ODS Graphics Designer 사용	479
ODS 통계적 그래프 표시	480

SAS ODS 통계적 그래프 정보

SAS ODS Statistical Graphics는 일반적으로 SAS ODS Graphics로 불리며 **SAS Output Delivery System(ODS)**의 확장형입니다. ODS는 프로시저에서 생성된 모든 출력을 관리하며 HTML 및 PDF를 포함한 다양한 형식으로 출력을 표시할 수 있습니다.

많은 SAS 분석 프로시저는 ODS Graphics 기능을 사용하여 이러한 프로시저 테이블처럼 자동으로 그래프를 생성합니다. ODS Graphics는 그래픽 템플릿 언어(GTL) 구문을 사용하고, 많은 복합 그래프를 생성하는 데 강력하고 유연한 기능을 제공합니다. GTL은 통계 그래프를 정의하기 위한 범용 언어입니다.

SAS Enterprise Guide에서 ODS Graphics Designer를 사용하여 GTL을 몰라도 이러한 통계 그래프를 정의할 수 있습니다. 그래프 정의가 생성된 후, 해당 그래프 정의를 사용하여 **SAS Enterprise Guide** 또는 **SAS Add-In for Microsoft Office**에서 ODS Statistical Graph를 생성할 수 있습니다.

ODS Graphics Designer 사용

주: SAS ODS Graphics Designer는 **SAS Add-In for Microsoft Office**에서 사용 가능하지 않지만 SAS Add-In에서 SAS ODS Graphics Designer를 사용하여 생성된 SGD 파일을 열 수 있습니다.

SAS ODS Graphics Designer는 사용자 정의 그래프를 생성하고 디자인할 수 있는 대화식 그래픽 응용 프로그램입니다. Designer는 그래프 템플릿 언어(GTL)에 근거한 그래프를 생

성합니다. 이 언어는 SAS 분석 프로시저 및 SAS ODS Graphics 프로시저에 의해 사용된 언어와 같습니다. ODS Graphics Designer는 그래픽 사용자 인터페이스를 제공하여 템플릿의 상세 정보와 GTL을 몰라도 그래프를 손쉽게 디자인할 수 있습니다.

주: SAS 9.4의 첫 번째 유지 관리 릴리스부터 로컬 컴퓨터에 Base SAS를 설치하지 않아도 SAS Enterprise Guide의 SAS ODS Graphics Designer를 사용할 수 있습니다.

포인트앤클릭 동작을 사용해서 분석용 데이터의 단순 또는 복합 그래픽 뷰를 생성할 수 있습니다. ODS Graphics Designer로 여러 도표 유형을 사용하여 복잡한 그래프를 디자인할 수 있습니다. 다중 셀 그래프, 분류 패널 및 산점도 행렬을 디자인할 수 있습니다. 그래프는 제목, 주석, 범례 및 기타 그래프 요소를 가질 수 있습니다. 결과를 리포트에 포함할 이미지나 ODS Graphics Designer 파일(SGD)로 저장하여 나중에 편집할 수 있습니다. 자세한 내용은 “[ODS 통계적 그래프 표시](#)” (480페이지)에서 참조하십시오.

SAS Add-In for Microsoft Office 또는 SAS Enterprise Guide에서 ODS Graphics Designer를 열려면 **작업** 영역에서 **ODS 통계적 그래프 표시**를 더블 클릭합니다. 자세한 내용은 [SAS Output Delivery System: User's Guide](#)를 참조하십시오.

ODS 통계적 그래프 표시

ODS 통계적 그래프 표시 작업은 **SAS Enterprise Guide** 및 **SAS Add-In for Microsoft Office**에서 사용 가능 합니다. 이 작업에서는 선택한 데이터 소스에 사용할 그래프 정의(SGD 파일)를 선택합니다. **SAS Enterprise Guide**에서 그래프 정의를 생성할 수 있습니다. 자세한 내용은 “[ODS Graphics Designer 사용](#)” (479페이지)에서 참조하십시오.

그래프 정의가 생성된 후, 로컬 및 원격 SAS 세션에서 사용될 수 있습니다. SAS 스토어드 프로세스에서도 이러한 그래프를 렌더링할 수 있습니다.

ODS 통계적 그래프 생성을 위한 그래프 정의(SGD 파일) 사용 방법:

- 1 SAS Add-In for Microsoft Office 또는 SAS Enterprise Guide의 **작업** 영역에서 **ODS 통계적 그래프 표시**를 더블 클릭합니다.
ODS 통계적 그래프 표시 대화 상자가 나타납니다.
- 2 사용하려는 디자인을 포함하는 SGD 파일을 선택합니다.
- 3 그래프에 대한 데이터 소스를 선택합니다.
- 4 (선택 사항) 그래프에 대한 너비와 높이를 지정합니다.
- 5 (선택 사항) 제목과 각주를 지정합니다.
- 6 **실행**을 클릭합니다.

그래프 정의에 대한 자세한 내용은 [SAS Output Delivery System: User's Guide](#)를 참조하십시오.

SAS Rapid Predictive Modeler

SAS Rapid Predictive Modeler 정보	481
SAS Rapid Predictive Modeler를 위한 표본추출 전략	482
SAS Rapid Predictive Modeler: 데이터 정리	483
SAS Rapid Predictive Modeler: 분석 역할에 변수 할당	484
SAS Rapid Predictive Modeler: 모델 선택	485
모델 옵션 설정	485
SAS Rapid Predictive Modeler용 모델의 이해	486
SAS Rapid Predictive Modeler: 리포트 옵션 선택	489
리포트 옵션	489
SAS Rapid Predictive Modeler용 표준 리포트 정보	489
SAS Rapid Predictive Modeler: 출력 옵션 설정	490
SAS Rapid Predictive Modeler: 모델 등록	491
모델 등록 정보	491
SAS 메타데이터 레파지토리에 모델 등록	492

SAS Rapid Predictive Modeler 정보

SAS Rapid Predictive Modeler는 다음 유형의 데이터 마이닝 분류 및 회귀 문제에 적합한 모델을 생성하기 위해 설계되었습니다.

- 이산 변수 값을 예측하는 분류 모델입니다. 일부 예로는 True 또는 False와 같은 변수 값을 예측하는 분류 모델이 있습니다(예: 구매 또는 거부, 높음, 중간 또는 낮음, 또는 계속).
- 회귀 모델은 실수 변수의 값을 예측합니다. 몇 가지 예로는 실가를 사용하여 매출, 매출 또는 성공률과 같은 금액을 예측할 수 있는 회귀 모델이 있습니다.

SAS Rapid Predictive Modeler를 사용하여 모델을 생성하려면 데이터셋을 제공해야 합니다. 모든 행은 독립 예측변수(입력변수)와 최소 하나의 종속변수(타겟)를 포함합니다. 데이터에서 표본을 추출하는 방법에 대한 자세한 내용은 SAS Rapid Predictive Modeler를 위한 표본추출 전략을 참조하십시오. 작업의 데이터 표본추출 방법에 대한 자세한 내용은 SAS Rapid Predictive Modeler의 표본추출 전략을 참조하십시오.

생성한 모델을 SAS 코드로 저장한 다음 SAS 환경에서 배포할 수 있습니다. SAS 모델 코드를 사용하여 새 데이터를 스코어링한 다음 그 결과를 바탕으로 보다 합리적인 비즈니스 결정을 내릴 수 있습니다. 이러한 프로세스를 모델 스코어링이라고 합니다. 스코어 데이터를 사용하여 특정 제안에 대해 선택할 고객을 결정하거나, 제조, 창고 보관 또는 물류 수요를 예측하거나, 고객 이탈을 관리하거나, 사기성 거래를 발견할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	
필수 SAS 제품	SAS Enterprise Miner
추가 SAS 제품(권장)	

SAS Rapid Predictive Modeler를 위한 표본추출 전략

SAS Rapid Predictive Modeler는 복합 표본추출 접근법을 사용합니다. 데이터 표본에 포함된 관측값 수는 다음 요인에 따라 달라집니다.

- 입력 변수 수
- 데이터 소스 내
- 데이터가 희귀 이벤트 타겟을 포함하는지 여부
- 데이터의 이벤트 수

다음은 SAS Rapid Predictive Model 작업에서 처리되는 관측값 수를 결정하는 데 사용하는 지침입니다.

입력 변수 수	처리되는 관측값 수
< 100	80,000
100 - 200	40,000
> 200	20,000

조건	희귀 이벤트: 예	희귀 이벤트: 아니요
< 처리되는 관측값 수 또는 총 이벤트 수 < (0.10*처리되는 관측값 수)	비이벤트 대 이벤트의 비율이 10:1이 되도록 데이터를 표본추출합니다.	표본추출 안 함
총 이벤트 수 > (0.10*처리되는 관측값 수)	희귀 이벤트를 다음 비율로 표본추출합니다 $10 * [(0.10 * (\text{처리되는 관측값 수})) / \text{이벤트 수}]$	층화추출

SAS Rapid Predictive Modeler: 데이터 정리

모델을 구성하기 전에 예측에 사용할 수 있는 히스토리 이벤트와 특성을 나타내는 입력 데이터가 필요합니다. 예측할 이벤트 또는 값을 나타내는 타겟 데이터도 필요합니다. 대부분 입력 데이터는 이전 기간에서 파생되며 타겟 데이터는 이후 기간에서 파생됩니다. 모델을 개발하는 데 사용하는 통합된 입력 및 타겟 데이터를 분석용 데이터라고 합니다.

예를 들어 작년 판매 영수증을 마이닝하여 내년 예상 수익을 예측하거나 특별 제안에 반응할 고객을 예측할 수 있습니다. 과거 이벤트의 히스토리 데이터를 사용하여 미래 이벤트의 성과를 예측하는 것을 모델 분석이라고 합니다.

최적의 모델 결과를 위해 모델 분석 데이터는 데이터의 행으로 저장된 많은 수의 관측값을 포함하고 있어야 합니다. 예를 들어 많은 소매 고객 모델은 수십만 개의 관측값을 가진 입력 데이터를 사용합니다.

타겟 변수에 고객 중 1%만 제안에 응답할 때와 같이 희귀 이벤트가 포함될 때 분석용 데이터에 상당 수의 해당 고객이 데이터셋으로 포함되어야 합니다. 해당 제안을 수락한 모든 고객을 선택하고, 제안을 받아들이지 않은 고객 수와 동일하도록 분석용 데이터를 과표본추출할 수도 있습니다. 과표본추출은 희귀 이벤트가 있는 모델에서 보다 쉽게 안정적인 솔루션을 찾는 데 유용합니다.

분석용 데이터에서 희귀 이벤트가 더 많이 나타나도록 과표본추출을 수행할 때 분석용 데이터에서 자연 모집단에 비례하여 타겟 이벤트의 수를 인위적으로 부풀려야 합니다. 분석용 데이터와 모집단 데이터 간의 차이를 보완하기 위해 Rapid Predictive Modeler에서는 사전 확률 설정을 제공합니다. 사전확률 설정은 모집단 데이터에 나타나는 타겟 이벤트의 실제 비례 빈도를 지정합니다.

SAS Rapid Predictive Modeler를 사용하여 마이닝한 데이터는 행(관측값)과 칼럼(변수)으로 구성되어야 합니다. 칼럼 중 하나가 타겟변수를 나타내야 합니다. 다음 예를 참조하십시오.

Name	Age	Gender	Income	Treatment	Purchase
Ricardo	29	M	33000	Y	Y

Name	Age	Gender	Income	Treatment	Purchase
Susan	35	F	51000	Y	N
Jeremy	49	M	110000	N	Y

Name

각 관측값에 대한 ID 값을 포함하는 칼럼입니다. SAS Rapid Predictive Modeler는 분석 콘텐츠에 대해 ID 변수 칼럼을 처리하지 않습니다.

Age, Gender, Income 및 Treatment

SAS Rapid Predictive Modeler가 사용하는 입력 칼럼입니다.

Purchase

타겟 칼럼입니다.

입력 데이터 테이블을 구성할 때 빈도 칼럼도 지정할 수 있습니다. 빈도 변수는 행(또는 관측값)에 할당해야 할 상대 가중값을 나타낼 수 있습니다. 예를 들어 일부 데이터셋에서 한 행이 두 개 이상의 관측값을 나타낼 수도 있습니다. 분석 중 SAS Rapid Predictive Modeler에서 무시할 칼럼을 선택할 수도 있습니다.

분석용 데이터에는 항상 입력 및 타겟 변수 값이 필요합니다. 스코어링에 사용되는 데이터에는 입력 변수 값만 필요하며 타겟 칼럼은 선택 사항입니다. 모델을 사용하여 새 데이터를 기반으로 예측할 때에는 타겟 칼럼이 필요하지 않습니다. 하지만 모델을 사용하여 효과를 모니터링할 때에는 타겟 칼럼이 있어야 합니다. 스코어링에 사용되는 데이터에는 일반적으로 ID 칼럼도 포함됩니다.

SAS Rapid Predictive Modeler: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

사용 가능한 분석 역할은 분석하려는 데이터 유형에 따라 달라집니다.

역할 이름	설명
종속변수	예측 또는 분류할 값을 지정합니다. 종속 변수는 타겟 변수로도 알려져 있습니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
빈도 수	데이터 행에 적용할 가중값을 나타내는 데 사용되는 변수를 지정합니다. 이 역할에는 변수를 하나만 할당할 수 있습니다.

역할 이름	설명
ID	선택 함수를 리포팅하고 스코어링하는 데 유용한 변수를 지정합니다. 이러한 변수는 분석에 포함되지 않습니다.
제외됨	분석에 포함시키지 않을 변수를 지정합니다.

SAS Rapid Predictive Modeler: 모델 선택

모델 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **모델**을 클릭합니다.

이 옵션을 사용하여 생성할 모델의 복잡성 수준을 지정할 수 있습니다. 중급 방법은 기본 및 중급 모델을 포함하고, 고급 방법은 기본, 중급, 고급 모델을 모두 포함합니다.

기본 방법을 사용하여 생성한 모델은 중급 방법을 사용하여 생성한 모델보다 더 빠르게 실행될 수 있지만 정확성이 떨어질 수도 있습니다. 중급 방법을 사용하여 생성한 모델과 고급 방법을 사용하여 생성한 모델을 비교할 때도 마찬가지입니다. 각 모델에 대해 사용된 데이터 마이닝 작업에 대한 자세한 내용은 [“SAS Rapid Predictive Modeler용 모델의 이해” \(486 페이지\)](#)를 참조하십시오.

SAS Enterprise Miner 모델링 함수는 SAS Rapid Predictive Modeler를 실행할 때 실행됩니다. 소프트웨어에서 실행하는 모델링 함수는 선택한 모델링 방법에 따라 달라집니다.

종속변수에서는 **의사결정 및 사전확률**을 클릭하여 다음 정보를 지정합니다.

테이블 76.1 종속변수에 대한 의사결정 및 사전확률 옵션

옵션 이름	설명
이벤트 레벨	모델링하려는 클래스 타겟 값을 지정합니다. SAS Rapid Predictive Modeler가 각 타겟 이벤트에 대한 확률을 제공하는 모델을 생성하기는 하지만 원하는 타겟 레벨이 인식되면 리포트 성능이 향상됩니다.
사전 확률	모델 분석용 데이터에 발생하는 타겟 변수 레벨의 개수와 비율을 표시합니다. 타겟이 범주형 변수이고 분석용 데이터와 모집단 데이터에서 서로 다른 타겟 분포가 있으면 이러한 값을 조정할 수 있습니다. 예를 들어 관측값의 50%가 응답자이고 나머지 50%는 비응답자인 과표본 추출된 데이터에 대해 분석한 모델이 있다고 가정해 봅시다. 하지만 모델에서 스코어링할 모집단 데이터는 과거에 10%의 응답자만 포함하고 90%의 비응답자를 포함합니다. 사전확률 설정을 사용하여 과거에 예측된 비응답자 대비 응답자 비율을 모델에 알려 줄 수 있습니다.

옵션 이름	설명
	■ 사전확률을 지정하지 않으려면 표시 안 함(기본값)을 선택합니다. ■ 타겟 변수의 모든 레벨에 대해 동일한 확률을 지정하려면 같은 값으로 지정을 선택합니다. ■ 스코어드 데이터의 타겟변수 레벨에 대한 사용자 정의 사전확률을 지정하려면 사용자 정의를 선택하고 확률을 지정하십시오. 지정한 사전확률의 합은 1이어야 합니다.
의사결정 함수	예측된 결과와 연결할 비용, 이익 또는 가중값을 지정합니다. 값 테이블은 의사결정 행렬이라고 합니다. 의사결정 행렬을 통해 가능한 각각의 의사결정 결과와 값을 연결할 수 있습니다. ■ 모델에 의사결정 행렬이 필요하지 않으면 표시 안 함(기본값)을 선택합니다. ■ 모델을 사용하여 이익을 최대화하려면 필요하면 최댓값을 선택하고 행렬의 True Positive 셀에 더 높은 가중값을 입력합니다. ■ 모델을 사용하여 비용을 최소화하려면 필요하면 최솟값을 선택하고 행렬의 True Negative 셀에 더 높은 가중값을 입력합니다. ■ 모델을 사용하여 희귀 이벤트를 예측하려면 역을 선택하여 True Positive 및 True Negative 예측을 식별합니다. 이때 False Positive 및 False Negative 예측을 잘못 평가할 위험이 있습니다.

테이블 76.2 모델화 옵션

옵션 이름	설명
기본	예측 또는 분류할 값을 지정합니다. 종속 변수는 타겟 변수로도 알려져 있습니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
중급	데이터 행에 적용할 가중값을 나타내는 데 사용되는 변수를 지정합니다. 이 역할에는 변수를 하나만 할당할 수 있습니다.
고급	선택 함수를 리포팅하고 스코어링하는 데 유용한 변수를 지정합니다. 이러한 변수는 분석에 포함되지 않습니다.

SAS Rapid Predictive Modeler용 모델의 이해

SAS Rapid Predictive Modeler는 기본, 중급 및 고급 모델을 제공합니다. 정교함과 복잡성에 따라 모델이 달라집니다.

- 기본 모델은 단순 회귀분석입니다.
- 중급 모델은 기본 모델의 분석과 함께 더 정교한 분석을 포함하고 있고 좀 더 나은 모델을 선택합니다.

- 고급 모델은 기본 및 중급 모델의 분석과 함께 훨씬 더 정교한 분석을 포함하고 있고 최적의 모델을 선택합니다.

기본 모델

일련의 3가지 데이터 마이닝 작업을 수행합니다.

- 변수 선택: 기본 모델은 모델링을 위해 상위 100개 변수를 선택합니다.
- 변환: 기본 모델은 모델링을 위해 선택된 상위 100개 변수에 대해 최적 묶음변환을 수행합니다. 최적 묶음변환은 결측 변수 값을 교정하므로 결측값 처리가 수행되지 않습니다.
- 모델링: 기본 모델은 전진 회귀모델을 사용합니다. 전진 회귀모델은 단계별 프로세스에서 변수를 한 번에 하나씩 선택합니다. 단계별 프로세스에서는 변수 기여도가 무의미해 때까지 한 번에 하나의 변수를 선형 방정식에 추가합니다. 전진 회귀모델은 분석에서 예측 능력이 없는 변수 또는 다른 예측변수와 상관성이 높은 변수를 제외하려고 합니다.

중급 모델

중급 모델은 일련의 7가지 데이터 마이닝 작업을 수행합니다.

- 변수 선택: 중급 모델링을 위해 상위 200개 변수를 선택합니다.
- 변환: 중급 모델은 모델링을 위해 선택된 200개의 변수에 대해 최적 거듭곱 변환을 수행합니다. 최적 거듭곱 변환은 Box-Cox 변환으로 알려진 일반적인 종류의 변환 중 하나입니다. 최적 거듭곱 변환에서는 지수 거듭곱 변환의 부분 집합을 평가하고 지정된 기준에 대해 최적의 결과를 생성하는 변환을 선택합니다.
- 결측값 처리: 중급 모델은 결측값 처리를 수행하여 결측 변수를 평균 변수 값으로 대체합니다. 결측값 처리 작업에서는 사용자가 결측값 처리된 변수 값을 포함하는 관측값을 식별할 수 있도록 하는 표시자 변수도 생성합니다.
- 변수 선택: 중급 모델은 카이제곱 및 R-제곱 기준 검증을 사용하여 타겟 변수와 관련이 없는 변수를 제거합니다.
- 변수 선택 기법 통합: 중급 모델은 메타데이터 카이제곱 및 R-제곱 기준 검증에서 선택된 변수 집합을 병합합니다.
- 모델링: 중급 모델은 분석용 데이터를 3가지 경쟁 모델 알고리즘으로 실행합니다. 이러한 모델은 의사결정트리, 로지스틱 회귀 및 단계별 회귀입니다. 로지스틱 회귀모델에서는 먼저 분석용 데이터가 의사결정트리로 실행되고, 여기에서 입력으로 회귀모델에 전달되는 NODE_ID 변수가 생성됩니다. NODE_ID 변수는 변수 교호작용 모델을 활성화하기 위해 생성됩니다.
- 챔피언 모델 선택: 중급 모델은 경쟁 모델의 예측 또는 분류 성능에 대한 분석 평가를 수행합니다. 평가 후 최적의 예측 또는 분류 성능을 보여 주는 모델이 모델링 분석을 수행하기 위해 선택됩니다. 최적 모델 선택과 관련한 중급 모델에서는 중급 모델의 성능뿐 아니라 기본 모델의 성능도 평가합니다.

SAS Rapid Predictive Modeler가 중급 최적 모델을 선택한 후 중급 최적 모델의 예측 성능을 기본 모델과 비교한 다음 그 결과로 최적 모델을 선택합니다.

고급 모델

고급 모델은 일련의 7가지 데이터 마이닝 작업을 수행합니다.

- 변수 선택: 고급 모델링을 위해 상위 400개 변수를 선택합니다.
- 변환: 고급 모델은 모델링을 위해 선택된 400개의 변수에 대해 다중 변환 알고리즘을 수행합니다. 다중 변환 작업에서는 후행 변수 선택 노드에서 사용할 몇 가지 변수 변환을 생성합니다. 다중 변환의 결과로 입력 변수의 수가 증가합니다. 입력 변수의 수가 증가하기 때문에 SAS Rapid Predictive Modeler는 다중 변환 알고리즘에 의해 생성된 출력에서 400개의 최적 입력 변수를 선택합니다.
- 결측값 처리: 고급 모델은 결측값 처리를 수행하여 결측 변수를 평균 변수 값으로 대체합니다. 결측값 처리 작업에서는 사용자가 결측값 처리된 변수 값을 포함하는 관측값을 식별할 수 있도록 하는 표시자 변수도 생성합니다.
- 변수 선택: 고급 모델은 카이제곱 및 R-제곱 기준 검증을 사용하여 타겟 변수와 관련이 없는 변수를 제거합니다. R-제곱 분석 중에 AOV16 변수가 생성됩니다.
- 변수 선택 기법 통합: 고급 모델은 메타데이터 카이제곱 및 R-제곱 기준 검증에서 선택된 변수 집합을 병합합니다.
- 모델링: 고급 모델은 분석용 데이터를 4가지 경쟁 모델 알고리즘으로 실행합니다. 이러한 모델은 의사결정트리 모델, 신경망 모델, 후진 회귀모델 및 앙상블 모델입니다. 신경망 모델은 최적의 피드 전진망을 찾기 위해 제한된 검색을 수행합니다. 후진 회귀는 R-제곱 스코어가 유의하게 떨어질 때까지 한 번에 하나씩 변수를 제거하여 변수를 소거하는 선형회귀모델입니다. 앙상블 모델은 여러 선행 입력 모델의 사후확률(Class 타겟일 때) 또는 예측값(Interval 타겟일 때)을 결합하여 새로운 모델을 생성합니다. 그런 다음 새로운 앙상블 모델이 새 데이터를 스코어링하는 데 사용됩니다. 고급 모델에서 사용하는 앙상블 모델은 기본 모델의 출력 모델, 중급 모델의 최적 모델, 고급 모델의 최적 모델을 기반으로 생성됩니다.
- 챔피언 모델 선택: 고급 모델은 경쟁 의사결정 트리, 신경망, 회귀분석 모델의 예측 또는 분류 성능에 대한 분석 평가를 수행합니다. 평가 후 최적의 예측 또는 분류 성능을 보여주는 모델이 기본 및 중급 모델의 최적 모델과 함께 입력으로 사용되어 앙상블 모델을 생성합니다. 그런 다음 새로 생성된 고급 앙상블 모델, 의사결정트리 모델, 신경망 모델 및 후진 회귀모델을 모두 분석 비교하여 모든 기본, 중급 및 고급 최적 모델의 표본 공간을 기반으로 최적의 모델을 선택합니다.

SAS Rapid Predictive Modeler가 고급 모델의 예측 성능을 중급 및 기본 모델의 최적 모델과 비교한 다음 그 결과로 최적 성능의 최적 모델을 선택합니다.

SAS Rapid Predictive Modeler: 리포트 옵션 선택

리포트 옵션

이 옵션에 액세스하려면 선택 영역에서 **리포트**를 클릭합니다.

리포트에서는 모델에서 유의한 항목을 식별하여 향상도 도표와 같은 일반적인 비즈니스 그래픽을 생성합니다. 결과는 분석용 및 평가용 데이터에 대한 통계량을 포함합니다. SAS Rapid Predictive Modeler 프로세스에서는 입력 데이터를 분석용 데이터와 평가용 데이터로 구분합니다. 분석용 데이터는 각 모델의 모수를 계산하는 데 사용되어 분석용 적합통계량을 생성합니다. 그런 다음 검증용 데이터는 각 모델로 스코어링하여 검증용 적합통계량을 산출합니다. 검증용 적합통계량은 모델을 비교하고 과적합을 감지하는 데 사용됩니다. 분석용 통계량이 평가용 통계량보다 훨씬 의미가 있을 때 과적합(Overfitting)을 의심할 수 있으며, 이 때 모델을 분석하여 데이터에서 임의 신호를 감지할 수 있습니다. 최량의 검증용 통계량을 가진 모델이 일반적으로 우선 설정됩니다.

SAS Rapid Predictive Modeler는 모델링에 사용된 데이터 소스와 변수에 대한 요약, 중요한 예측변수의 순위, 모델의 정확도를 평가하는 여러 가지 적합통계량 및 모델 스코어카드를 제공하는 간결한 핵심 리포트 집합을 자동으로 생성합니다. 자세한 내용은 SAS Rapid Predictive Modeler용 표준 리포트 정보를 참조하십시오.

결과에 다음 정보를 포함하도록 선택할 수도 있습니다.

- 모델 요약화
- 변수 순위
- 교차 테이블
- 분류 행렬
- 적합통계량
- 향상도 도표
- 모델 비교

주: 한 가지 모델 알고리즘만 사용하는 기본 방법에는 이 옵션을 사용할 수 없습니다.

SAS Rapid Predictive Modeler용 표준 리포트 정보

다음은 SAS Rapid Predictive Modeler에서 자동으로 생성되는 표준 리포트입니다.

이득 도표

이득 도표는 Class 타겟 변수가 있는 모델에 대해서만 사용할 수 있습니다. 이 그래프는 예측값별로 순위화된 데이터의 백분위수를 보여줍니다. 향상도는 임의 선택으로 발견된 타겟 이벤트 수에 비해 모델이 식별한 타겟 이벤트 수의 비율을 나타내는 척도입니다.

ROC(Receiver Operating Characteristic Plot)

Receiver Operating Characteristic 도표는 단일 십분위수보다 전체 표본에 대한 모델의 최대 예측검정력을 보여줍니다. 데이터는 민감도 대 (1 - 특이도)로 도표화됩니다. 모델 곡선과 대각선(임의 선택 모델을 나타냄) 간의 차이를 Kolmogorov-Smirnov(KS) 값이라고 합니다. KS 값이 높을수록 더욱 강력한 모델을 나타냅니다.

스코어카드

모델의 특성을 비즈니스 용도에 맞게 해석할 수 있도록 결과에 스코어카드를 포함합니다. 스코어카드가 구성되면 각 Interval 변수는 개별 값 범위로 그룹화됩니다. 그런 다음, 각 변수는 모델 중요도에 따라 순위화되고 최대 1,000개 포인트로 척도화됩니다. 그러면 각 변수의 개별 값에 척도화된 총 포인트 중 일부가 부여됩니다.

프로젝트 정보

프로젝트 정보는 모델을 생성한 사용자, 모델 생성 시기 및 모델의 구성 요소 파일 저장 위치를 보여줍니다.

SAS Rapid Predictive Modeler: 출력 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

옵션 이름	설명
Enterprise Miner 프로젝트 데이터 저장	이 작업에서 SAS Enterprise Miner 데이터를 저장할지를 지정합니다. SAS Rapid Predictive Modeler의 모델은 SAS Enterprise Miner 프로젝트의 예제입니다. SAS Enterprise Miner 데이터를 저장하면, SAS Enterprise Miner 인터페이스를 사용하여 SAS Rapid Predictive Modeler가 생성한 모델을 열고 편집할 수 있습니다. SAS Enterprise Miner에서는 SAS Enterprise Miner 외부에서 사용하기 위해 분석을 저장하고 내보내며, SAS Metadata Repository에 모델을 등록할 수 있습니다. Enterprise Miner 프로젝트 데이터 저장 체크 박스는 기본적으로 선택되어 있습니다. ■ SAS Enterprise Miner의 워크스테이션 설치가 있는 경우 프로젝트가 SAS App 서버 폴더 필드에 지정된 워크스테이션 위치에 로컬로 저장됩니다. 찾아보기를 클릭하여 이 위치를 선택합니다. ■ SAS Enterprise Miner의 클라이언트/서버 설치가 있는 경우 프로젝트가 서버의 User Root 폴더(예 : C:\users\username\Documents)에 저장됩니다. SAS Management Console을 사용한 경우 서버에서 SYSTEM 루트 위치(예 : C:\projects)를 정의하려면 SAS App 서버 폴더 필드가 비어 있습니다. 이 경우 찾아보기를

옵션 이름	설명
	클릭하여 프로젝트 저장소 위치를 지정합니다. SAS 응용 프로그램 서버 폴더 필드가 비어 있고 프로젝트 저장 전 위치를 지정하지 않으면, 올바른 위치를 지정하라는 오류 메시지가 나타납니다.
데이터셋 스코어	SAS Rapid Predictive Modeler가 생성한 모델로 스코어링할 데이터셋의 이름을 지정합니다.

SAS Rapid Predictive Modeler: 모델 등록

이 옵션에 액세스하려면 선택 영역에서 **등록**을 클릭합니다.

모델 등록 정보

SAS Rapid Predictive Modeler를 사용하여 생성한 모델은 쉽게 배포할 수 있습니다. 이 코드는 SAS 9.2 이상 설치에서 실행할 수 있는 SAS 코드로 저장됩니다.

주: SAS Metadata Repository에 모델을 등록하려면 먼저 SAS Rapid Predictive Modeler를 한 번 이상 실행하거나 저장해야 합니다. 모델을 생성할 때 등록 옵션을 사용하려면 **Enterprise Miner 프로젝트 데이터 저장** 옵션을 선택해야 합니다.

모델을 등록하면 해당 사이트의 다른 SAS 사용자도 등록된 모델을 사용할 수 있습니다.

- **SAS Enterprise Guide** 및 **SAS Add-In for Microsoft Office** 사용자는 모델을 사용하여 자신만의 스코어링 프로세스를 수행할 수 있습니다. 예를 들어 등록된 모델은 모델 스코어링 작업에 새로운 데이터를 스코어링하는 데 사용될 수 있습니다.
- **SAS Enterprise Miner** 사용자는 이러한 모델을 **SAS Enterprise Miner**에서 프로젝트 및 다이어그램으로 열 수 있으며, 추가 진단을 수행하거나 동작 및 대안 계획을 조사하도록 해당 모델을 수정할 수 있습니다.
- **SAS Model Manager** 사용자는 다른 모델링 자산과 함께 관리될 모델을 가져올 뿐 아니라 모델 성능을 모니터링하여 시간에 따른 성능 저하를 확인할 수도 있습니다.
- **SAS Data Integration Studio**의 사용자는 모델을 가져와서 관리 및 스케줄된 스코어링 프로세스를 생성할 수 있습니다.

SAS Rapid Predictive Modeler를 사용하여 생성된 모델은 SAS Scoring Accelerator for Teradata, DB2, Netezza를 사용하여 게시될 수도 있습니다.

SAS 메타데이터 레파지토리에 모델 등록

- 1 모델 이름, 경로 및 설명을 지정합니다.
- 2 **등록**을 클릭합니다.

.....
주: **등록** 버튼은 SAS Workspace 서버에 연결되어 있지 않으면 사용할 수 없습니다.
.....

산점도

산점도 작업 정보	493
산점도: 도표 유형 선택	494
산점도: 분석 역할에 변수 할당	494
산점도: 모양 옵션 설정	495
3D 산점도 옵션 설정	495
기타 도표 옵션 설정	496
추가 모양 옵션	496
제목 및 각주 지정	497
속성 보기	497

산점도 작업 정보

산점도 작업은 데이터 포인트의 패턴이나 집중도를 통해 2개 또는 3개의 변수 간 관계를 보여 주는 2D 산점도, 3D 산점도 또는 3D 바늘 도표를 생성합니다.

예를 들어 2D 산점도는 임상 연구에 포함된 모든 환자의 몸무게와 나이를 표시할 수 있습니다.

자세한 내용은 [“그래픽 출력형식 선택” \(5페이지\)](#)에서 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GPLOT, G3D
필수 SAS 제품	Base SAS, SAS/GRAPH
추가 SAS 제품(권장)	없음

산점도: 도표 유형 선택

이 옵션에 액세스하려면 선택 영역에서 **산점도**를 클릭합니다.

역할 이름	설명
2D 산점도	2개의 변수값을 가로축(X) 및 세로축(Y) 상에 데이터 포인트로 표시하는 2차원 산점도를 생성합니다. X, Y 값으로 이루어진 각 쌍은 데이터 포인트를 형성합니다.
3D 산점도	3개의 숫자 변수값을 데이터 포인트로 표시하는 3차원 도표를 생성합니다. X, Y, Z 값으로 이루어진 각 그룹은 데이터 포인트를 형성합니다.
바늘을 포함한 3D 산점도	3개의 숫자 변수값을 데이터 포인트로 표시하는 3차원 도표를 생성합니다. X, Y, Z 값으로 이루어진 각 그룹은 데이터 포인트를 형성합니다. 각 데이터 포인트는 바늘이라고 하는 선으로 수평면에 연결됩니다.

산점도: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
수평	이 역할에 할당하는 칼럼은 그래프의 가로축 또는 X축 변수입니다.
수직	이 역할에 할당하는 칼럼은 그래프의 세로축 또는 Y축 변수입니다. 가로축 변수의 고윳값에 대한 요약통계를 생성할 수 있습니다. 세로 역할에서 변수를 선택합니다. 가로축의 개별값 요약 을 선택하고 요약 함수를 선택합니다. 예를 들어 입력 데이터는 각 국가의 판매 및 비용을 지역별로 나열합니다. 이때 가로 역할에 국가 변수를, 세로 역할에 판매 변수를 할당합

역할 이름	설명
	니다. 각 국가에 대한 판매 합계를 도표로 나타내려면 체크 박스를 선택한 다음 드롭다운 리스트에서 합을 선택합니다. 이 작업은 국가의 지역별 판매를 추가하고 최종 출력에 각 국가의 단일 판매 값을 표시합니다.
세로(오른쪽)	이 역할에 할당하는 칼럼은 그래프의 오른쪽 면에 대한 세로축 또는 Y축 변수입니다. 이 변수는 X축 변수에 대응하여 그려지므로 그래프에 별도의 도표가 생성됩니다. 가로축 변수의 고윳값에 대한 요약통계를 생성할 수 있습니다. 세로 역할에서 변수를 선택합니다. 가로축의 개별값 요약을 선택하고 요약 함수를 선택합니다. 예를 들어 입력 데이터는 각 국가의 판매 및 비용을 지역별로 나열합니다. 이때 가로 역할에 국가 변수를, 세로(오른쪽) 역할에는 비용 변수를 할당합니다. 각 국가에 대한 비용 합계를 도표로 나타내려면 체크 박스를 선택한 다음 드롭다운 리스트에서 합을 선택합니다. 이 작업은 국가의 지역별 비용을 추가하고 최종 출력에 각 국가에 대해 단일 비용 값을 표시합니다. 주: 이 역할은 2D 산점도에만 사용할 수 있습니다.
깊이	이 역할에 할당하는 칼럼은 그래프의 세 번째 차원 또는 Z축 변수입니다. 주: 이 옵션은 3D 도표에만 사용할 수 있습니다.
그룹 분석 칼럼	이 역할에 할당하는 칼럼의 값에 따라 생성되는 그래프의 수가 결정됩니다. 각 고윳값에 대해 별도의 그래프가 생성됩니다. 생성된 모든 그래프에 대해 각 축에 대한 배율을 동일하게 하려면 모든 그래프에 대해 축 길이 같음 체크 박스를 선택합니다. 기본적으로, 각 축에 대한 값의 범위는 도표화된 데이터에서 최소 및 최댓값에 기반합니다. 이 옵션을 선택하면 각 축의 범위는 모든 그래프에 대해 동일합니다.

산점도: 모양 옵션 설정

3D 산점도 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **3D 산점도**를 클릭합니다.

기호 종류, 색상 및 크기를 선택하고 각 데이터 포인트에 대해 바늘을 표시할지를 선택합니다.

슬라이더 또는 **좌우 회전** 및 **상하 회전** 컨트롤을 사용하여 보기 각도를 설정합니다. 좌우 회전 각도와 상하 회전 각도는 70입니다. 보기 각도를 기본 설정으로 되돌리려면 **재설정**을 클릭합니다.

주: 상하 및 좌우 회전 값 지정 체크 박스를 선택 취소하면 그래프의 상하 및 좌우 회전이 그래프 출력형식에 따라 결정됩니다. 이 옵션은 기본적으로 선택되며 보기 각도를 지정할 수 있습니다.

기타 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **도표**를 클릭합니다.

주: 이 옵션은 2D 산점도에만 사용할 수 있습니다.

특정 변수에 대한 도표 모양을 제어하는 방법

- 1 리스트에서 변수를 선택합니다. 리스트의 변수는 **세로** 및 **세로(오른쪽)** 역할에 할당된 변수입니다. I
- 2 **선** 영역에서 선 유형, 선 두께 및 선 색상을 지정합니다.
- 3 외곽선 색상을 선택합니다. 이 옵션은 다음 부분의 색상을 지정합니다.
 - 둘러싸인 영역의 외곽선
 - 회귀분석의 신뢰한계
 - 표준 편차 도표의 선 또는 외곽선
- 4 **데이터 포인트 표식** 영역에서 표식 유형, 사용할 기호, 각 표식의 높이 및 데이터 포인트 표식의 색상을 지정합니다.

주: 일부 기호는 모든 그래픽 출력형식에서 지원되지 않습니다. 결과 출력형식에서 지원되지 않는 기호를 선택하면 지원되는 다른 기호가 대신 표시됩니다.

- 5 **결측값**에서 도표 선을 끊을지를 선택합니다.

주: 결측값에서 두 번째 도표 선 분리 체크 박스는 **세로(오른쪽)** 역할에 변수를 할당할 때만 활성화됩니다.

추가 모양 옵션

일부 옵션은 여러 그래프 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- “보간법 선택” (601페이지)
- “축 옵션 설정” (593페이지)
- “참조선 추가” (595페이지)
- “그래프 영역 사용자 정의” (597페이지)

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

산점도 행렬

산점도 행렬 작업 정보	499
산점도 행렬: 산점도 행렬 생성	499

산점도 행렬 작업 정보

산점도 행렬 작업에서는 다양한 변수의 조합에 대해 산점도 패널 그래프를 생성합니다. 산점도에서 적합 3도표 및 타원을 중첩하는 옵션을 사용할 수 있습니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	SGSCATTER
필수 SAS 제품	Base SAS, SAS/GRAPH
추가 SAS 제품(권장)	없음

산점도 행렬: 산점도 행렬 생성

옵션 이름	설명
레이블	프로젝트 트리 및 프로세스 플로우에서 작업을 식별합니다. 레이블 상자에 작업 이름을 입력합니다.
데이터 소스	작업의 입력 데이터 소스를 수정할 수 있습니다. 기본적으로 작업을 열기 전에 선택한 데이터 소스가 입력 데이터 소스입니다. 입력 데이

옵션 이름	설명
	터 소스의 이름은 데이터 소스 필드에 표시됩니다. 데이터 소스에 지정된 필터는 작업 필터 필드에 표시됩니다.
행렬 변수	산점도를 생성하도록 그래프화할 변수를 지정합니다. 이 역할에는 적어도 두 개의 변수를 지정해야 합니다.
그룹 변수	값을 그룹으로 나눌 분류변수를 지정합니다. 이 역할에는 변수를 할당하지 않아도 됩니다.
그래프 제목	결과에 표시되는 제목을 지정합니다. 기본적으로 결과의 제목은 산점도 행렬입니다. 그러나 이 제목은 사용자 정의할 수 있습니다. 제목의 텍스트를 편집하려면 아래쪽 화살표를 클릭합니다. 제목에 매크로 변수를 사용할 수 있습니다. 매크로 변수의 값은 작업 실행 시 결정됩니다.
대각선	도표 행렬의 대각선 셀에 그래프를 추가합니다. 대각선 셀에 히스토그램을 표시하고 이러한 히스토그램을 정규 밀도함수 곡선 또는 커널 밀도함수 추정값과 중첩시킬 수 있습니다. 또한 대각선이 왼쪽 상단 모서리 또는 오른쪽 상단 모서리에서 시작할지를 지정할 수 있습니다. 대각선에 그래프가 포함되도록 선택하지 않으면 대각선 셀에 변수 이름이 포함됩니다.
타원	산점도가 포함된 각 셀에 신뢰 또는 예측 타원을 추가합니다. 두 가지 타원 유형은 모두 이변량 정규분포로 간주합니다. 다음 타원 유형 중에서 선택할 수 있습니다. ■ 없음 - 타원을 추가하지 않습니다. ■ 평균 - 모평균에 신뢰 타원을 추가합니다. ■ 예측값 - 새로운 관측값에 예측 타원을 추가합니다. 타원에 대해 신뢰수준을 지정할 수도 있습니다.
범례	산점도 행렬에 범례를 표시할지를 지정합니다. 범례는 그룹 변수를 할당할 때에만 사용할 수 있습니다. 그룹 변수를 할당하면 범례가 기본적으로 표시됩니다. 범례를 제거하려면 범례 표시 체크 박스의 선택을 취소합니다. 범례 제목을 지정하고 위치 드롭다운 리스트에서 범례의 위치를 선택합니다. 범례 주위에 테두리를 표시할지를 지정할 수도 있습니다.
그래프 영역	그래프의 높이와 너비를 지정합니다. 또한 산점도에 있는 데이터 포인트 위로 포인터를 이동할 때 나타나는 데이터 팁을 표시할지를 지정할 수도 있습니다. 기본적으로 그래프 영역은 800 x 800이며 데이터 팁은 표시되지 않습니다.
그래프 주석	결과에 표시되는 주석을 지정합니다. 기본적으로 작업을 실행한 날짜 및 시간은 그래프의 주석에 표시됩니다. (그래프 주석 텍스트 상자에는 현재 날짜 및 시간에 대한 SAS 매크로 변수가 표시됩니다.) 그러나 결과의 주석은 사용자 정의할 수 있습니다. 각주의 텍스트를 편집하

옵션 이름	설명
	려면 아래쪽 화살표를 클릭합니다. 주석에 매크로 변수를 사용할 수 있습니다.
	주석을 삭제하려면 텍스트 상자의 텍스트를 모두 삭제하고 확인 을 클릭합니다. 그러면 결과에 주석이 포함되지 않습니다.

데이터 정렬

데이터 정렬 작업 정보	503
데이터 정렬: 분석 역할에 변수 할당	504
데이터 정렬: 데이터 정렬 옵션 설정	504
데이터 정렬: 결과 옵션 설정	506
속성 보기	506

데이터 정렬 작업 정보

데이터 정렬 작업을 통해 프로젝트 내에서 데이터셋을 칼럼별로 정렬할 수 있습니다. 하나 이상의 칼럼을 기준으로 데이터를 정렬할 수 있으며, 각 칼럼에 대해 오름차순 또는 내림차순을 선택할 수 있습니다. 데이터 정렬 작업을 수행하면 정렬된 데이터셋이 프로젝트에서 새로운 데이터 항목으로 표시됩니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	SORT
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	없음

데이터 정렬: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
정렬 기준	이 역할에 하나 이상의 변수를 지정하면 테이블이 선택한 변수별로 정렬됩니다. 이 역할 내에서 변수가 나타나는 순서에 따라 1차 정렬키, 2차 정렬키 등이 결정됩니다. 1차 정렬키는 항상 정렬 기준 역할에서 첫 번째 표시되는 변수가 됩니다.
제외할 데이터	이 역할에 하나 이상의 변수를 지정하면 출력에 해당 변수가 포함되지 않습니다. 이 역할에는 최대 $(n - 1)$ 개의 변수를 할당할 수 있습니다. 여기서 n 은 정렬 중인 데이터셋의 총 변수 수입니다.

데이터 정렬: 데이터 정렬 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

옵션 이름	설명
출력 순서	다음 옵션을 수정하여 데이터 출력 순서를 결정할 수 있습니다. ■ 순서 맞추기 - 문자 값 변수를 정렬하는 데 사용할 순서 맞추기 방법을 나타냅니다. 서버에 정의된 데이터 정렬 순서를 사용하거나, ASCII 또는 EBCDIC 데이터 정렬 순서를 선택하거나, 덴마크어, 핀란드어, 노르웨이어 또는 스웨덴 국가 표준을 사용할 수 있습니다. 해당 설치 사이트에서 정의한 사용자 정의 순서 맞추기를 사용하여 문자 변수를 정렬하도록 선택할 수도 있습니다. 주: 사용자 정의를 선택하면 해당 사이트에 국가별 정렬 표준이 정의되어 있어야 합니다. 사용자 정의된 국가별 정렬 순서의 사용 여부는 관리자가 결정할 수 있습니다. 사용 가능한 정

옵션 이름	설명
	렬 순서가 없으면 서버에 정의된 기본 순서 맞추기 방법이 사용됩니다. ■ '정렬 기준' 그룹 내에서 원래 데이터 순서 유지 - 분석 역할에 설정한 순서에 따라 데이터를 그룹화합니다. 이 옵션을 선택하지 않으면 정렬 기준 그룹에서 정의되지 않은 순서로 출력 데이터셋이 그룹화됩니다.
중복 레코드	다음 옵션을 사용하여 출력 시 중복 레코드의 처리 방법을 지정할 수 있습니다. ■ 모든 레코드 유지 - 중복된 레코드를 포함하여 모든 레코드를 출력 데이터셋에 그대로 유지합니다. ■ 각 '정렬 기준' 그룹의 첫 레코드만 유지 - 정렬 기준 그룹에 대해 동일한 값을 갖는 중복 관측값을 제거합니다. '정렬 기준' 그룹 내에서 원래 데이터 순서 유지 옵션을 선택하면 각 정렬 기준 그룹에 유지되는 관측값은 원래 데이터셋에서 첫 번째로 읽은 데이터입니다. 그러나 '정렬 기준' 그룹 내에서 원래 데이터 순서 유지 옵션을 선택하지 않으면 각 정렬 기준 그룹에 유지되는 관측값을 미리 알 수 없습니다. ■ 이웃한 중복 레코드 유지 안 함 - 각 레코드를 출력 데이터셋의 이전 레코드에 비교합니다. 정확히 일치하는 항목이 발견되면 중복 레코드는 출력 데이터셋에 기록되지 않습니다. 주: 정렬 기준 역할에 모든 변수를 할당하지 않으면 일부 중복 레코드가 제거되지 않을 수 있습니다.
고급 정렬 옵션	출력 데이터의 정렬 방법에 대해 고급 옵션을 설정할 수 있습니다. 다음과 같은 옵션을 수정할 수 있습니다. ■ 정렬에 사용할 메모리 - 데이터 정렬 작업에 사용할 수 있는 최대 메모리 크기를 지정합니다. 메모리 크기는 바이트(B), 킬로바이트(KB), 메가바이트(MB), 기가바이트(GB) 단위로 지정할 수 있습니다. 사용 가능한 메모리를 모두 사용하거나, 서버에 할당된 기본 메모리 크기를 사용하도록 지정할 수도 있습니다. ■ 필요한 임시 디스크 공간 줄이기 - 데이터 정렬 시 정렬 기준 변수 및 관측값 번호만 임시 파일에 저장하여, 정렬 작업에 필요한 스토리지 양을 줄일 수 있습니다. 정렬 마지막 단계에서 임시 파일을 인덱스로 사용하여 원래 데이터에 액세스한 후 데이터를 정렬된 순서로 결과 데이터셋에 전송합니다. ■ 인덱싱된 데이터 강제 정렬 - 원하는 순서로 이미 정렬된 데이터셋 또는 사용자 정의 인덱스와 정렬 기준 역할에 지정한 순서가 반영된 키가 들어 있는 데이터셋을 포함하여, 모든 데이터셋에 대해 정렬을 수행합니다. 이 옵션을 지정하면 현재의 데이터셋 순서 또는 인덱스 포함 여부에 관계 없이 모든 데이터셋이 정렬됩니다.

데이터 정렬: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

출력 데이터셋의 이름 또는 저장 위치를 다르게 지정할 수 있습니다. **정렬 기준** 그룹에 대해 동일한 값을 갖는 중복 관측값을 제거하거나, 값이 정확히 일치하는 레코드 중 하나만 유지하도록 하면 중복된 레코드를 데이터셋에 저장할 수도 있습니다. 자세한 내용은 **중복 레코드** 옵션을 참조하십시오.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

칼럼 분할

칼럼 분할 작업 정보.....	507
칼럼 분할: 분석 역할에 변수 할당.....	509
칼럼 분할: 결과 옵션 설정.....	510
속성 보기.....	510

칼럼 분할 작업 정보

칼럼 분할 작업에서는 입력 데이터셋에서 선택한 칼럼 값의 고유한 조합을 다중 칼럼으로 분할하여 출력 데이터셋을 생성합니다. 출력 데이터셋을 사용하면 여러 입력 데이터셋 행을 포함하는 칼럼들을 개별적으로 분석할 수 있습니다.

이 작업은 한 칼럼에 여러 하위그룹에 대한 다중 관측값이 포함되어 있으며, 하위그룹 측도를 개별 칼럼으로 분할할 때 유용합니다.

예를 들어 노스캐롤라이나주 각 지역의 월별 기온값 칼럼이 포함된 다음 데이터셋을 살펴보겠습니다.

	📍 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	Charlotte	January	51
7	Charlotte	February	52,8
8	Charlotte	March	61,7
9	Charlotte	April	70,5
10	Charlotte	May	79,6
11	Raleigh	January	51
12	Raleigh	February	54,1
13	Raleigh	March	62,1
14	Raleigh	April	71,6
15	Raleigh	May	79
16	Wilmington	January	56
17	Wilmington	February	59
18	Wilmington	March	65,8
19	Wilmington	April	74,4
20	Wilmington	May	81,1

칼럼 분할 작업을 통해 각 지역의 월별 기온값을 각 월의 칼럼으로 전치하여 다음과 같은 데이터셋을 생성할 수 있습니다.

	📍 Location	⊕ January	⊕ February	⊕ March	⊕ April	⊕ May
1	Asheville	46,9	50,6	59	68	74,9
2	Charlotte	51	52,8	61,7	70,5	79,6
3	Raleigh	51	54,1	62,1	71,6	79
4	Wilmington	56	59	65,8	74,4	81,1

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
----------	--------------

사용되는 SAS 프로시저	TRANSPOSE
---------------	-----------

요구 사항 이름	프로시저 및 제품 이름
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	SAS/STAT

칼럼 분할: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분할할 칼럼	이 역할에 지정한 칼럼에는 새로운 칼럼으로 전치되는 값이 포함되어 있습니다. 이 역할에는 하나의 변수만 지정해야 합니다.
데이터 값 식별 칼럼	이 역할에 지정한 칼럼은 분할할 칼럼 의 전치값을 포함하는 새로운 칼럼의 이름을 제공합니다. 이 칼럼의 개별값 순서는 어느 행이 새로운 칼럼으로 전치될지를 나타냅니다. 그룹 변수를 지정하지 않으면 이 칼럼에는 한 집합의 개별값만 포함되며, 분할할 칼럼 의 값은 출력 데이터셋의 한 행으로 전치됩니다. 데이터 값 식별 칼럼 역할에는 반드시 칼럼을 지정해야 합니다. 그룹 변수를 선택하려면 그룹 분석변수 역할에 칼럼을 지정합니다.
데이터 값 레이블 칼럼	이 역할에 지정한 칼럼은 분할할 칼럼 의 전치값을 포함하는 새로운 칼럼의 레이블을 제공합니다. 주: 이 칼럼의 효과는 출력 데이터셋에 있는 새로운 칼럼의 속성을 보면 알 수 있습니다.
그룹 분석변수	이 역할에 지정한 각 변수는 분할할 칼럼 의 행을 새 칼럼으로 전치되는 하위그룹으로 분할할 때 사용됩니다. 그룹 변수값의 개별 조합을 사용하여 정의되는 각 하위그룹은 출력 데이터셋을 구성하는 행이 됩니다. 하위그룹을 구성하는 행의 개수는 데이터 값 식별 칼럼 에 있는 개별값 개수에 의해 결정됩니다.

칼럼 분할: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

출력 데이터셋에 다른 이름을 지정할 수 있습니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

출력 데이터셋 칼럼 이름에 접두어를 추가할 수 있습니다. 기본적으로 칼럼 이름은 **데이터 값 식별 칼럼**의 고유타입에서 생성됩니다. 접두어를 지정하려면 **접두어 사용** 체크 박스를 선택하고 텍스트 상자에 접두어를 입력합니다. 기본 접두어는 **Column**입니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

칼럼 세로 배열

칼럼 세로 배열 작업 정보	511
칼럼 세로 배열: 분석 역할에 변수 할당	513
칼럼 세로 배열: 칼럼 명명 옵션 설정	513
칼럼 세로 배열: 결과 옵션 설정	514
속성 보기	515

칼럼 세로 배열 작업 정보

칼럼 세로 배열 작업은 해당 칼럼이 관측값으로 전치될 수 있도록 입력 데이터셋에서 선택한 칼럼을 재구성하여 출력 데이터셋을 생성합니다. 출력 데이터셋을 사용하면 여러 입력 데이터셋 칼럼의 값을 분석할 수 있습니다. 관측값을 그룹화하면 선택된 칼럼은 그룹 값의 개별 조합에 기반한 하위그룹으로 나누어집니다. 각 하위그룹은 출력 데이터셋의 행을 구성합니다.

이 작업은 데이터셋의 각 관측값이 여러 칼럼에 동일한 유형의 데이터를 포함하고 있으며, 여러 칼럼의 데이터를 분석할 때 유용합니다. 예를 들어 노스캐롤라이나주 각 지역의 월별 기온값이 포함된 다음 데이터셋을 예로 들어 보겠습니다.

	 Location	 January	 February	 March	 April	 May
1	Asheville	46,9	50,6	59	68	74,9
2	Charlotte	51	52,8	61,7	70,5	79,6
3	Raleigh	51	54,1	62,1	71,6	79
4	Wilmington	56	59	65,8	74,4	81,1

이 데이터 구조를 통해 노스캐롤라이나주 전역의 월별 기온 요약통계량을 계산할 수 있습니다. 그러나 관측값 전체에 대한 통계량만 계산되고 특정 지역에 대한 요약통계량은 계산할 수 없습니다.

모든 기온에 대한 통계량 또는 각 지역에 대한 통계량을 계산하려면 월별 기온값 칼럼을 지역별로 세로 배열하여 다음과 같은 출력 데이터셋을 생성합니다.

	 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	Charlotte	January	51
7	Charlotte	February	52,8
8	Charlotte	March	61,7
9	Charlotte	April	70,5
10	Charlotte	May	79,6
11	Raleigh	January	51
12	Raleigh	February	54,1
13	Raleigh	March	62,1
14	Raleigh	April	71,6
15	Raleigh	May	79
16	Wilmington	January	56
17	Wilmington	February	59
18	Wilmington	March	65,8
19	Wilmington	April	74,4
20	Wilmington	May	81,1

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	TRANSPOSE
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	SAS/STAT

칼럼 세로 배열: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
세로로 배열할 칼럼	이 역할에 지정되는 각 칼럼은 출력 데이터셋의 행이 됩니다. 선택된 그룹 변수는 각 칼럼을 하위칼럼으로 분할하는 데 사용되고 각 하위칼럼은 행으로 바뀝니다. 이때 칼럼은 그룹 변수로 정의된 그룹 수와 동일한 수의 행으로 전치됩니다. 세로로 배열할 칼럼 역할에 변수를 하나 이상 지정해야 합니다. 그룹 변수를 선택하려면 그룹 분석변수 역할에 칼럼을 지정합니다. 주: 그룹 변수를 지정하지 않으면 입력 데이터셋에는 한 개의 관측값만 포함되며 각 칼럼은 출력 데이터셋의 한 행으로 전치됩니다.
그룹 분석변수	이 역할에 지정되는 각 변수는 칼럼을 분할하여 개별적으로 전치되는 하위 그룹으로 누적하는 데 사용됩니다. 그룹 변수값의 개별 조합을 사용하여 정의되는 각 하위그룹은 출력 데이터셋의 행이 됩니다. 그룹 변수값의 개별 조합은 유일해야 합니다.

칼럼 세로 배열: 칼럼 명명 옵션 설정

출력 데이터셋의 새로운 칼럼에 이름 및 레이블을 지정하려면 선택 영역에서 **칼럼 이름**을 클릭합니다.

역할 이름	설명
새로운 값을 가지는 칼럼의 이름 및 레이블 지정	새로운 칼럼은 입력 데이터셋의 관측값을 출력 데이터셋 변수값으로 전치하여 생성됩니다. 전치값이 포함된 출력 칼럼의 머리글을 지정하려면 새로운 값을 가지는 칼럼 영역의 이름 상자에 머리글을 입력합니다. 처음에는 이름 상자에 기본 머리글 StackedValues가 포함됩니다.

역할 이름	설명
	출력 데이터셋과 함께 저장된 칼럼의 설명을 지정하려면 새로운 값을 가지는 칼럼 영역의 레이블 상자에 레이블을 입력합니다. 설명 문자열의 길이는 최대 256자입니다. 레이블을 생략하려면 이 상자를 비워 둡니다.
파생 칼럼의 이름 및 레이블 지정	출력 데이터셋의 각 행에는 현재 관측값을 생성하기 위해 전치된 입력 데이터셋의 변수 이름이 포함됩니다. 이 변수 이름이 포함된 출력 칼럼의 머리글을 지정하려면 파생 칼럼 영역의 이름 상자에 머리글을 입력합니다. 처음에는 이름 상자에 기본 머리글 ValueSource가 포함됩니다. 출력 데이터셋과 함께 저장된 칼럼의 설명을 지정하려면 파생 칼럼 영역의 레이블 상자에 레이블을 입력합니다. 설명 문자열의 길이는 최대 256자입니다. 레이블을 생략하려면 이 상자를 비워 둡니다.
설명 칼럼의 이름 및 레이블 지정	출력 데이터셋의 각 행에는 해당 출력 행의 값이 속하는 입력 데이터셋의 변수 레이블이 포함됩니다. 이 변수 이름이 포함된 출력 칼럼의 머리글을 지정하려면 설명 칼럼 영역의 이름 상자에 머리글을 입력합니다. 처음에는 이름 상자에 기본 머리글 ValueDescription이 포함됩니다. 출력 데이터셋과 함께 저장된 칼럼의 설명을 지정하려면 설명 칼럼 영역의 레이블 상자에 레이블을 입력합니다. 설명 문자열의 길이는 최대 256자입니다. 레이블을 생략하려면 이 상자를 비워 둡니다. 주: 칼럼 레이블 지정 효과는 출력 데이터셋에 있는 새로운 칼럼의 속성을 보면 알 수 있습니다.

칼럼 세로 배열: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

출력 데이터셋에 다른 이름을 지정할 수 있습니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

데이터 표준화

데이터 표준화 작업 정보	517
데이터 표준화: 분석 역할에 변수 할당	518
데이터 표준화: 표준화 옵션 설정	518
데이터 표준화: 결과 옵션 설정	519
속성 보기	519

데이터 표준화 작업 정보

데이터 표준화 작업은 SAS 데이터셋의 변수를 주어진 평균 및 표준편차로 표준화한 후 이 표준화된 값을 포함하는 새로운 SAS 데이터셋을 생성합니다.

예를 들어 이 작업으로 지정된 평균 및 표준편차를 사용하여 테스트 스코어를 표준화할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	STANDARD
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	없음

데이터 표준화: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분석변수	이 역할에 할당하는 각 변수에 대해 표준화가 수행됩니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
상대 가중값 변수	분석변수에 대한 가중 평균 및 가중 분산을 계산하는 데 사용할 값을 지정합니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 다른 역할에 이미 지정한 변수로는 테이블을 그룹화할 수 없습니다.

데이터 표준화: 표준화 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **표준화**를 클릭합니다.

옵션 이름	설명
표준화 기준 설정	기본적으로 이 작업에서는 데이터를 평균 0, 표준편차 1로 표준화합니다. 분석할 변수에 대해 새로운 평균 및 표준편차를 지정하거나, 현재의 평균 또는 표준편차를 그대로 사용할 수도 있습니다. 주: 모든 분석변수에 동일한 평균 및 표준편차가 적용됩니다. 새로운 평균 또는 새로운 표준편차 체크 박스의 선택을 취소하여 각 변수의 평균 또는 표준편차를 그대로 유지할 수 있습니다.

옵션 이름	설명
결측값을 칼럼의 평균값으로 바꾸기	기본적으로 입력 분석변수의 결측값은 표준화 프로세스에서 제외됩니다. 결측값을 칼럼의 평균값으로 바꾸기 체크 박스를 선택하여 입력 결측값을 해당 변수에 대한 비결측값 평균으로 바꿀 수 있습니다.

데이터 표준화: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

새로운 표준화된 변수 추가 옵션은 기본적으로 선택되므로, 출력 데이터에 표준화된 변수가 원래 변수와 함께 포함됩니다. 원래 변수를 표준화된 변수로 바꾸려면 **기존 변수 바꾸기**를 선택합니다.

주: SAS Enterprise Guide는 **옵션 창**의 **작업 > 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식

■ 결과의 그래픽 출력에 대한 출력형식

.....
주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.
.....

요약통계량

요약통계량 작업 정보	521
요약통계량: 분석 역할에 변수 할당	522
요약통계량: 요약통계량 선택	523
기본 요약통계량 선택	523
백분위수 요약통계량 선택	523
추가 요약통계량 선택	524
요약통계량: 도표 옵션 설정	524
요약통계량: 결과 옵션 설정	525
제목 및 각주 지정	527
속성 보기	527

요약통계량 작업 정보

요약통계량 작업은 모든 관측값 및 관측값 그룹 내의 변수에 대한 기술통계량을 계산할 수 있는 데이터 요약 도구를 제공합니다. 히스토그램과 같은 그래픽 화면에서 데이터를 요약할 수도 있습니다.

예를 들어 이 작업을 사용하여 제품 유형과 국가에 따라 정렬된 새로운 판매 수에 대한 리포트를 생성할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	MEANS, UNIVARIATE
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	SAS/GRAPH

요약통계량: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분석변수	통계량 생성에 사용할 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
분류변수	입력 데이터를 범주 또는 하위그룹으로 분류하는 데 사용할 변수를 지정합니다. 변수는 문자 또는 숫자가 될 수 있습니다. 분류변수들의 유일한 조합에 대해 선택된 모든 분석변수에 대한 통계량이 계산됩니다. 다중 레이블 출력형식 허용 옵션을 설정하여 이 변수에 다중 레이블 출력형식이 지정될 때, 주어진 범위 또는 겹치는 범위에 대한 첫 번째 및 두 번째 출력형식 레이블을 사용하여 하위그룹 조합이 생성되도록 설정할 수 있습니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
상대 가중값 변수	가중 평균, 가중 분산 및 가중합을 계산하는 데 사용할 변수를 지정합니다.
COPY 변수	출력 데이터셋에 포함할 변수를 지정합니다. COPY 변수 역할에 대하여 가져올 값 옵션을 사용하여 입력 데이터셋의 해당 관측값에 있는 변수의 최댓값 또는 최솟값을 출력 데이터셋으로 복사할 수 있습니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다.

요약통계량: 요약통계량 선택

기본 요약통계량 선택

이 옵션에 액세스하려면 선택 영역의 통계량 머리글 아래에서 **기본**을 클릭합니다.

요약통계량 리포트에는 기본적으로 평균, 표준편차, 최솟값, 최댓값, 관측값 개수 등의 통계량이 포함됩니다. 계산할 통계량, 리포트에 표시될 소수점 이하 자릿수, 분위수 통계량 계산 방법 등을 제어할 수 있습니다.

옵션 이름	설명
기본 통계량	분석에 포함할 기본 통계량을 지정합니다.
최대 소수점 이하 자릿수	계산된 통계량에 대한 최대 소수점 이하 자릿수를 지정합니다. 기본적으로 통계량은 최적 적합을 사용하여 표시되는데, 일반적으로 소수점 이하 7자리입니다.
표준편차 및 분산에 사용할 분모	표준편차 및 분산 계산에 사용할 분모를 지정합니다. 기본 분모는 자유도입니다.

백분위수 요약통계량 선택

이 옵션에 액세스하려면 선택 영역의 통계량 머리글 아래에서 **백분위수**를 클릭합니다.

옵션 이름	설명
백분위수 통계량	계산할 백분위수를 지정합니다. 제1 사분위수(25번째 백분위수), 중위수(50번째 백분위수) 및 제3 사분위수(75번째 백분위수)를 계산하도록 선택할 수도 있습니다. 기본적으로 백분위수는 선택되지 않습니다.
분위수 계산 방법	분위수, 중위수, 백분위수 계산에 사용할 방법을 지정합니다. 다음 두 가지 방법 중에서 선택할 수 있습니다. ■ 순서 통계량 - 모든 데이터를 메모리로 읽어와서 유일한 값을 기준으로 정렬하는 방법입니다. ■ 불연속 포물선 알고리즘 - 분위수의 근사값을 구하는 메모리 점유도가 낮은 방법입니다.

추가 요약통계량 선택

이 옵션에 액세스하려면 선택 영역의 통계량 머리글 아래에서 **추가**를 클릭합니다.

다음 통계량을 계산할 수도 있습니다.

- 평균의 신뢰한계
- t 통계량과 p-값(Prob > |t|)
- 변동계수
- 수정제곱합
- (비수정)제곱합

평균의 신뢰한계에 대한 신뢰수준을 계산할 수도 있습니다.

요약통계량: 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

주: **분류변수** 역할에 변수를 할당했으면 도표는 분류변수의 n -원 조합에 대해서만 생성됩니다.

도표 이름	설명
히스토그램	히스토그램은 데이터셋의 분포를 결정하는 데 사용되는 그래프입니다. 히스토그램에 대한 중간점을 지정할 수 있습니다. 드롭다운 리스트에서 사용할 수 있는 옵션은 히스토그램을 생성하는지 또는 비교 히스토그램을 생성하는지에 따라 달라집니다. 분류변수 역할에 변수를 할당할 때에는 비교 히스토그램을 생성합니다. 비교 히스토그램에서는 다음 옵션을 사용할 수 있습니다. ■ 중간점 지정 - 시작 중간점과 마지막 중간점을 지정할 수 있습니다. 다른 중간점을 계산하는 데 사용할 단계 간격도 지정할 수 있습니다. 중간점은 각 막대로 표시되는 값 범위의 중간 값입니다. ■ 키 - 키 셀의 데이터에 대해 중간점을 결정합니다. 기본적으로 분류변수의 레벨은 첫 번째 분류변수의 내부(서식화되지 않은) 값의 증가 순으로 위에서 아래로(왼쪽에서 오른쪽으로) 표시됩니다. 분류변수를 하나만 지정하면 키 셀은 이러한 순서로 첫 번째에 나타나는 레벨이 됩니다. 분류변수를 두 개 지정하면 키 셀은 이러한 순서로 첫 번째에 나타나는 변수 1과 변수 2의 레벨 결합이 됩니다. 따라서 키 셀의 선택에 따라 모든 셀에 사용되는 균등 수평 축이 결정됩니다. 그런 다음

도표 이름	설명
	키 셀에 대한 중간점 리스트에 나머지 셀의 데이터가 포함될 때까지 리스트가 어느 방향이든지 필요한 만큼 확장됩니다. ■ 균등 - 총 표본 크기를 기반으로 중간점 개수를 결정합니다. 분류변수 역할에 변수를 할당하지 않을 때에는 히스토그램을 생성합니다. 히스토그램에서는 다음 옵션을 사용할 수 있습니다. ■ 중간점 지정 - 시작 중간점과 마지막 중간점을 지정할 수 있습니다. 다른 중간점을 계산하는 데 사용할 단계 간격도 지정할 수 있습니다. 중간점은 각 막대로 표시되는 값 범위의 중간 값입니다. ■ 그룹 수 지정 - 데이터에 대한 막대 수(히스토그램 구간이라고도 함)를 지정할 수 있습니다. 기본적으로 표시되는 구간 수는 데이터 범위에 따라 달라집니다. 그러나 표시할 최대 구간 수는 지정할 수 있습니다.
상자수염도	상자수염도는 중앙 위치의 측도(중위수), 산포의 두 가지 측도(범위 및 사분위 범위), 왜도(사분위에 대해 중위수의 방향으로부터) 및 가능한 이상치를 표시합니다. 상자 도표는 둘 이상의 데이터셋을 비교하는 데 특히 유용합니다.

요약통계량: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

옵션 이름	설명
데이터셋에 통계량 저장	통계량을 데이터 테이블에 영구히 저장할지를 지정합니다. 주: SAS Enterprise Guide는 옵션 창의 작업 > 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
결과에 포함할 통계량을 선택합니다.	결과에 포함할 통계량을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ COPY 변수 역할에 대하여 가져올 값 - COPY 변수 역할에 변수를 할당하면 입력 데이터셋의 해당 관측값에 있는 변수의 최댓값 또는 최솟값을 출력 데이터셋으로 복사할 수 있습니다. ■ 분류변수 조합 - 분류변수 역할에서 분류변수에 대한 요약통계량을 테이블로 그룹화할 수 있습니다. n-원 테이블을 생성할 수 있는데, 여기에서 n은 분류변수의 개수입니다. 그룹은 분류변수의 유일한 조합

옵션 이름	설명
	에 의해 결정됩니다. 테이블에서 변수의 순서는 분류변수 역할의 변수 순서에 따라 결정됩니다. 다음 옵션을 사용하여 출력 데이터를 하위그룹으로 분류할 수 있습니다. □ N-원 조합만 - 분류변수의 n-원 조합에 대한 요약통계량을 보여주는 단일 테이블을 생성합니다. 예를 들어 두 가지 마케팅 캠페인의 매출 결과를 비교한다고 가정합니다. 분석변수 역할에 Profit 변수를 할당하고 분류변수 역할에 Month, Marketing Campaign 및 Region 변수를 할당합니다. 그런 다음, N-원 조합만 옵션을 선택하면 Month*Campaign*Region의 테이블이 생성됩니다. □ 모든 가능한 조합 - 분류변수의 각 고유 조합의 요약통계량에 대한 별도의 테이블을 생성합니다. 예를 들어 변수로 Month, Campaign, Region을 사용하고 모든 가능한 조합 옵션을 사용하면 다음과 같은 테이블이 생성됩니다. ■ 총 매출 테이블. 이를 무원테이블이라고 하며, 분류변수는 포함되지 않습니다. ■ 각 월(예: 5, 6, 7)에 대한 요약통계량을 표시하는 일원테이블 ■ 각 지역(예: 북쪽, 남쪽, 동쪽, 서쪽)에 대한 요약통계량을 표시하는 일원테이블 ■ 각 마케팅 캠페인 유형(예: TV 광고, 잡지 광고)에 대한 요약통계량을 표시하는 일원테이블 ■ Month*Region에 대한 요약통계량을 표시하는 이원테이블 ■ Month*Campaign에 대한 요약통계량을 표시하는 이원테이블 ■ Region*Campaign에 대한 요약통계량을 표시하는 이원테이블 ■ Month*Campaign*Region에 대한 요약통계량을 표시하는 삼원테이블 □ 조합 지정 - 분류변수의 각 고유 조합에 대한 요약통계량에 대해 별도의 테이블을 생성합니다. 예를 들어 방법 지정 상자에 1, 2 또는 3을 입력하면 결과에는 가능한 모든 일원, 이원 또는 삼원 테이블이 포함됩니다. 이 상자에 0을 지정하면 분류변수가 포함되지 않는 무원테이블이 결과에 표시됩니다.
분석 레이블 표시	분석변수에 대한 레이블을 표시합니다.
모든 출력 숨기기	결과를 표시하지 않습니다. 출력이 표시되지 않도록 선택하면 출력 데이터셋에 대한 결과를 저장해야 합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

요약통계량 마법사

요약통계량 마법사 정보	529
요약통계량 마법사: 데이터 선택	529
요약통계량 마법사: 역할에 변수 할당	529
요약통계량 마법사: 리포트 통계량 및 출력 결과 지정	530
요약통계량 마법사: 제목 및 각주 지정	531

요약통계량 마법사 정보

요약통계량 마법사에서는 선택 사항인 그룹 변수로 기술통계량 테이블, 히스토그램, 상자 도표 등을 생성할 수 있습니다.

요약통계량 마법사: 데이터 선택

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

다음을 클릭하여 역할에 변수를 할당합니다.

요약통계량 마법사: 역할에 변수 할당

1 데이터를 선택해야 역할에 변수를 할당할 수 있습니다.

- 한 개 이상의 변수를 **분석변수(요약통계량을 생성할 변수)** 역할에 할당합니다. 이 역할에 지정하는 변수는 통계량 생성에 사용할 숫자 변수입니다.

주: 이 역할에는 변수를 하나 이상 할당해야 합니다. 다른 모든 역할은 선택 사항입니다.

- 한 개 이상의 변수를 **분류변수(이 변수의 값에 따라 통계량 생성됨)** 역할에 할당합니다. 이 역할에 지정하는 변수는 입력 데이터를 범주 또는 하위그룹으로 분류하는 데 사용될 문자 또는 이산형 숫자 변수입니다. 분류변수들의 유일한 조합에 대해 선택된 모든 분석 변수에 대한 통계량이 계산됩니다.
- 한 개 이상의 변수를 **그룹변수(이 변수의 값에 따라 별도의 테이블 생성됨)** 역할에 할당합니다. 이 역할에 할당된 변수는 각 개별값 또는 **그룹 분석변수**의 값 조합에 대해 별도의 통계량을 계산하는 데 사용됩니다. 이 역할의 변수를 기준으로 데이터가 자동 정렬된 후 통계량이 계산됩니다. 각 변수에 대한 정렬 순서를 선택하려면 위쪽/아래쪽 화살표를 사용합니다.
- (선택 사항) **고급**을 클릭하여 다음과 같은 역할에 변수를 할당합니다.
 - **빈도 수** - 이 역할에 변수를 할당하면 테이블의 각 관측값이 n 관측값을 나타내게 되며, 여기서 n은 해당 행의 빈도변수값입니다. 통계량은 자동으로 계산됩니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.
 - **상대 가중값** - 이 역할에 변수를 할당하면 각 관측값의 변수값은 가중 평균, 가중 분산 및 가중합을 계산하는 데 사용됩니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.
 - **COPY 변수** - 이 역할에 할당되는 변수는 출력 데이터셋에 추가 변수로 포함됩니다.
- 다음**을 클릭하여 통계량과 결과를 지정합니다.

요약통계량 마법사: 리포트 통계량 및 출력 결과 지정

- 리포트 통계량과 출력 결과를 지정하기 전에 역할에 변수를 할당해야 합니다.
- 결과에 포함할 통계량을 선택합니다. 기본적으로 평균, 표준편차, 최소, 최대 및 관측값 개수가 포함됩니다. 다음 탭에 있는 옵션을 사용하여 추가적인 통계량을 포함할 수 있습니다.

기본

설명 옵션을 포함합니다. 계산된 통계량에 대한 최대 소수점 이하 자릿수를 선택할 수도 있습니다. 기본적으로 통계량은 최적 적합을 사용하여 표시되는데, 일반적으로 소수점 이하 7자리입니다.

백분위수

계산할 백분위수를 지정합니다. 제1 사분위수(25번째 백분위수), 중위수(50번째 백분위수) 및 제3 사분위수(75번째 백분위수)를 계산하도록 선택할 수도 있습니다. 기본적으로 백분위수는 선택되지 않습니다.

분위수 계산 방법 옵션을 사용하여 분위수, 중위수, 백분위수 계산에 사용할 방법을 선택합니다. 다음 두 가지 방법 중에서 선택할 수 있습니다.

- **순서 통계량** - 모든 데이터를 메모리로 읽어와서 유일한 값을 기준으로 정렬하는 방법입니다.
- **불연속 포물선 알고리즘** - 분위수의 근사값을 구하는 메모리 점유도가 낮은 방법입니다.

추가

다음 통계량을 포함합니다.

- **평균의 신뢰한계**
- **t 통계량과 p-값(Prob > |t|)**
- **변동계수**
- **수정제곱합**
- **(비수정)제곱합**

평균의 신뢰한계에 대한 신뢰수준을 계산할 수도 있습니다.

- 3 **통계량 표시** 옵션을 선택하여 결과를 표시합니다. 통계량을 결과에 표시할 것인지 아니면 데이터 테이블에 영구 저장할 것인지를 선택하거나 두 가지 모두를 선택해야 합니다.
- 4 **히스토그램**이나 **상자수염도** 체크 박스를 선택하여 생성하려는 도표 유형을 선택합니다. 한 가지 또는 두 가지를 선택할 수 있습니다. 기본적으로 어떠한 도표도 선택되지 않습니다.

주: **분류변수** 역할에 변수를 할당했으면 도표는 분류변수의 n -원 조합에 대해서만 생성됩니다.

- 5 **데이터셋에 통계량 저장** 체크 박스를 선택하여 출력을 저장합니다. 통계량을 결과에 표시할 것인지 아니면 데이터 테이블에 영구 저장할 것인지를 선택하거나 두 가지 모두를 선택해야 합니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

- 6 **다음**을 클릭하여 제목과 각주를 지정합니다.

요약통계량 마법사: 제목 및 각주 지정

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

- 1 제목과 주석을 지정하기 전에 리포트 통계량과 출력 결과를 지정해야 합니다.
- 2 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다. 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

.....
주: 결과에 포함되도록 선택한 도표에 따라 일부 제목을 사용할 수 없습니다.
.....

- 3 **마침**을 클릭하여 작업을 실행합니다.

요약테이블

요약테이블 작업 정보	533
요약테이블: 분석 역할에 변수 할당	534
요약테이블: 요약 테이블 설정	534
요약 테이블 생성	534
테이블 모양 사용자 정의	535
요약테이블: 결과 옵션 설정	537
제목 및 각주 지정	537
속성 보기	537

요약테이블 작업 정보

요약테이블 작업에서는 데이터셋의 전체 또는 일부 변수를 사용하여 기술통계량이 테이블 형식으로 표시됩니다. 단순한 형태에서 고급 사용자 정의된 형태에 이르는 다양한 테이블을 생성할 수 있습니다.

예를 들어 SAS 데이터셋에 두 가지 유형, 즉 거주지와 직업별 고객에 대한 제품 판매 데이터가 미국의 북동부와 서부 지역의 각 주별로 포함되어 있다고 가정합니다. 요약테이블 작업을 사용하여 지리적 지역의 각 주에 대해 각 고객 유형에 대한 판매를 합할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	TABULATE
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	없음

요약테이블: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분석변수	분석에 대한 변수를 지정합니다. 이 역할 또는 분류변수 역할에는 한 개 이상의 변수를 지정해야 합니다.
분류변수	입력 데이터를 범주 또는 하위그룹으로 분류하는 데 사용할 문자 변수 또는 이산형 숫자 변수를 지정합니다. 분류변수들의 유일한 조합에 대해 선택된 모든 분석변수에 대한 통계량이 계산됩니다. 다중 레이블 출력형식 옵션을 설정하여 이 변수에 다중 레이블 출력형식이 지정될 때, 주어진 범위 또는 겹치는 범위에 대한 첫 번째 및 두 번째 출력형식 레이블을 사용하여 하위그룹 조합이 생성되도록 설정할 수 있습니다.
페이지	이 변수의 각 개별값에 대해 별도의 요약 테이블이 생성됩니다.
상대 가중값 변수	가중 평균, 가중 분산 및 가중합을 계산하는 데 사용할 변수를 지정합니다.
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.

요약테이블: 요약 테이블 설정

요약 테이블 생성

이 옵션에 액세스하려면 선택 영역에서 **요약테이블**을 클릭합니다.

요약테이블을 설정하려면 원하는 정보와 정보의 표시 방식을 결정해야 합니다. 요약테이블 작업에서 사용자는 자신이 원하는 방식으로 분류변수 및 분석변수를 정렬할 수 있습니다.

변수를 사용하는 방법:

- 1 **사용 가능한 변수** 상자에서 변수를 선택합니다.
- 2 **미리 보기** 상자의 위쪽 또는 왼쪽 영역에 있는 셀로 변수를 끌어 놓습니다. 새 변수를 기존 변수 위에 끌어 놓으면 기존 변수가 새로운 변수로 바뀝니다. 새 변수를 기존 변수에 추가하려면 셀 경계선에 새 변수를 끌어 놓습니다. 마우스 포인터를 정확한 위치에 놓으면 경계선이 파란색으로 바뀝니다.

주: 사용하는 분석변수에 적용되는 통계량은 기본적으로 관측값의 개수(N)입니다. 위에서 설명한 것과 동일한 방법으로 통계량을 추가하거나 바꿀 수 있습니다.

테이블 모양 사용자 정의

요약테이블 모양을 사용자 정의하는 옵션을 설정하려면 미리 보기 테이블을 마우스 오른쪽 버튼으로 클릭합니다. **셀 제거**를 제외한 팝업 메뉴의 각 항목에 대해 다양한 옵션을 설정할 수 있는 창이 열립니다.

도표 이름	설명
셀 제거	미리 보기 테이블의 분석변수, 분류변수 또는 통계량을 마우스 오른쪽 버튼으로 클릭하면 이 항목이 활성화됩니다. 이 항목을 사용하여 요약테이블에서 변수 또는 통계량을 제거할 수 있습니다.
테이블 속성	이 항목은 항상 활성화됩니다. 테이블 속성 대화 상자에 설정된 옵션은 전체 요약테이블에 적용됩니다. ■ 일반 탭: 결측값에 대한 레이블을 지정합니다. 기본값은 마침표('.')입니다. 체크 박스를 사용하여 결측값을 분류변수의 정당 값으로 포함하고 행 분류변수의 머리글(레이블)을 숨길 수 있습니다. ■ 출력형식 탭: 표준 SAS 출력형식 또는 사용자 작성 출력형식을 요약 테이블의 모든 데이터 값에 할당합니다. 필드 너비, 소수점 이하 자릿수 및 출력형식을 선택합니다. ■ 글꼴 탭: 테이블의 모든 텍스트에 대한 글꼴, 크기, 스타일, 전경색, 배경색, 정렬 및 수직 정렬을 지정합니다.
상자 영역 속성	이 항목은 항상 활성화됩니다. 이 항목을 사용하여 테이블 왼쪽 상단 모서리에 대한 옵션을 설정합니다. 이러한 옵션은 테이블 속성 대화 상자에 선택된 옵션을 무시합니다. ■ 일반 탭: 상자 영역에 사용할 텍스트를 입력합니다. 페이지 역할에 하나 이상의 변수를 할당하면 상자 영역에 페이지 머리글이 표시되도록 선택할 수 있습니다. 상자 영역의 텍스트가 공간에 비해 길면 공간에 맞춰 줄 바꿈을 하거나, 절단하도록 선택할 수도 있습니다. ■ 출력형식 탭: 이 탭은 상자 영역에는 적용되지 않습니다.

도표 이름	설명
	■ 글꼴 탭: 상자 영역의 텍스트에 대한 글꼴, 크기, 스타일, 전경색, 배경색, 정렬 및 수직 정렬을 지정합니다.
머리글 속성	이 항목을 활성화하려면 칼럼 또는 행 머리글을 마우스 오른쪽 버튼으로 클릭해야 합니다. 이 항목을 활성화하려면 칼럼 또는 행 머리글을 마우스 오른쪽 버튼으로 클릭해야 합니다. 이러한 옵션은 테이블 속성 대화 상자에 선택된 옵션을 무시합니다. ■ 일반 탭: 행 또는 칼럼 머리글에 레이블을 입력합니다. 머리글의 텍스트가 머리글 공간에 비해 길면 공간에 맞춰 줄 바꿈하거나 자를 수도 있습니다. ■ 출력형식 탭: 이 탭은 분류변수 값에 적용되지 않습니다. 분류변수에 머리글 출력형식을 할당하려면 선택 영역에서 데이터를 클릭하고 분류변수를 선택합니다. 출력형식을 지정하려면 머리글 출력형식 옵션을 선택합니다. ■ 글꼴 탭: 머리글 텍스트의 글꼴, 크기, 스타일, 전경색, 배경색, 정렬 및 수직 정렬을 지정합니다.
클래스 레벨 속성	이 항목은 미리 보기 테이블에서 분류변수를 마우스 오른쪽 버튼으로 클릭했을 때만 활성화됩니다. 이 항목을 사용하여 분류변수값에 대한 옵션을 설정할 수 있습니다. ■ 일반 탭: 분류변수 값 텍스트를 줄 바꿈할지를 선택합니다. ■ 출력형식 탭: 이 탭은 분류변수 값에 적용되지 않습니다. 분류변수에 머리글 출력형식을 할당하려면 선택 영역에서 데이터를 클릭하고 분류변수를 선택합니다. 출력형식을 지정하려면 머리글 출력형식 옵션을 선택합니다. ■ 글꼴 탭: 분류 변수 값 텍스트의 글꼴, 크기, 스타일, 전경색, 배경색, 정렬 및 수직 정렬을 지정합니다.
데이터 값 속성	미리 보기 테이블의 변수 또는 통계량을 마우스 오른쪽 버튼으로 클릭하면 이 항목이 활성화됩니다. 이 항목을 사용하여 요약 테이블의 선택한 행 또는 칼럼에서 데이터 값에 대한 옵션을 설정합니다. 이러한 옵션은 테이블 속성 대화 상자에 선택된 옵션을 무시합니다. ■ 일반 탭: 데이터 값 텍스트를 줄 바꿈할지를 선택합니다. ■ 출력형식 탭: 표준 SAS 출력형식 또는 사용자 작성 출력형식을 요약 테이블에서 선택한 행 또는 칼럼의 모든 데이터 값에 할당합니다. 필드 너비, 소수점 이하 자릿수 및 출력형식을 선택합니다. ■ 글꼴 탭: 데이터 값 텍스트의 글꼴, 크기, 스타일, 전경색, 배경색, 정렬 및 수직 정렬을 지정합니다.

요약테이블: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

결과를 출력 데이터셋에 저장하려면 **데이터셋에 결과 저장** 체크 박스를 선택합니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

요약테이블 마법사

요약테이블 마법사 정보	539
요약테이블 마법사: 데이터 선택	539
요약테이블 마법사: 역할에 변수 할당	540
요약테이블 마법사: 분류변수 선택	540
요약테이블 마법사: 합계 지정	541
요약테이블 마법사: 추가 출력 선택	541
요약테이블 마법사: 제목 및 각주 지정	541

요약테이블 마법사 정보

요약테이블 마법사에서는 요약통계량에 대한 옵션을 사용하여 교차 테이블과 같은 다양한 구성의 테이블을 생성할 수 있습니다.

요약테이블 마법사: 데이터 선택

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

다음을 클릭하여 변수 및 통계량을 할당합니다.

요약테이블 마법사: 역할에 변수 할당

- 1 데이터를 선택해야 역할에 변수를 할당할 수 있습니다.
- 2 변수를 **분석변수** 상자에 추가하려면 **추가**를 클릭합니다. 팝업 메뉴에서 이 역할에 추가할 변수를 선택합니다. 이 역할에 할당하는 변수는 계산에 사용될 숫자 변수입니다.
- 3 **분석변수** 상자에서 각 분석변수에 대해 계산하려는 통계량을 선택합니다. 기본적으로 합 통계량이 계산됩니다. 다른 통계량을 계산하려면 변경하려는 통계량에 대한 **통계** 칼럼의 셀을 선택합니다. 드롭다운 리스트를 사용하여 계산할 통계량을 선택합니다.

주: 데이터에 숫자변수가 포함되어 있지 않으면 분석변수를 할당할 수 없습니다. 이때 빈도 통계량이 분석변수 대신 사용됩니다.

- 4 **분석변수 레이블** 드롭다운 리스트에서 분석변수에 대한 레이블의 위치를 선택합니다. 사용 가능한 옵션에는 **칼럼에 위치**, **행에 위치**, **페이지에 위치** 또는 **숨김**이 있습니다.
- 5 **통계량 레이블** 드롭다운 리스트에서 통계량에 대한 레이블의 위치를 선택합니다. 사용 가능한 옵션에는 **칼럼에 위치**, **행에 위치**, **페이지에 위치** 또는 **숨김**이 있습니다.
- 6 **다음**을 클릭하여 통계량과 결과를 지정합니다.

요약테이블 마법사: 분류변수 선택

- 1 분석변수 및 통계량을 할당할 수 있으면(선택 사항) 분류변수도 선택할 수 있습니다.
- 2 칼럼, 행 또는 페이지에 분류변수를 추가하려면 해당 상자에 대해 **추가**를 클릭합니다. 팝업 메뉴에서 추가할 변수를 선택합니다. 분류변수로 할당하는 변수는 입력 데이터를 범주로 나누는 데 사용되는 문자 또는 이산 숫자 변수입니다. 분류변수로 지정되는 변수는 입력 데이터를 범주로 나누는 데 사용되는 문자 변수 또는 이산형 숫자 변수입니다. **페이지** 상자에 할당한 각 변수에 대해 별도의 요약 테이블이 생성됩니다.
- 3 **기타 옵션**을 클릭하여 테이블 구조를 사용자 정의합니다. **테이블 구조** 대화 상자에서 칼럼, 행 및 페이지에 대한 머리글을 숨길 수 있습니다. 칼럼, 행 및 테이블을 중첩하거나 연결할 수 있습니다.
- 4 **다음**을 클릭하여 제목과 각주를 지정합니다.

요약테이블 마법사: 합계 지정

- 1 분류변수를 지정해야 합계를 지정할 수 있습니다.
- 2 **칼럼, 행 및 페이지** 드롭다운 리스트에서 각 칼럼, 행 또는 페이지에 대한 전체 합계를 표시할지를 선택합니다. 합계를 표시하려면 **표시**를 선택합니다. 합계를 보지 않으려면 **없음**을 선택합니다. 페이지에서 페이지 합계를 페이지의 처음 또는 마지막에 표시할지 여부도 지정할 수 있습니다.

주: 칼럼, 행 또는 페이지에 대한 분류변수를 지정하지 않으면, 해당 옵션을 사용할 수 없습니다.

- 3 **합계에 대한 레이블** 상자에 합계에 대한 레이블을 입력합니다.
- 4 **다음**을 클릭하여 제목과 각주를 지정합니다.

요약테이블 마법사: 추가 출력 선택

- 1 추가 출력을 선택하기 전에 합계를 지정할 수 있습니다.
- 2 **데이터셋에 결과 저장** 체크 박스를 선택합니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

- 3 **다음**을 클릭하여 제목과 각주를 지정합니다.

요약테이블 마법사: 제목 및 각주 지정

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

- 1 제목과 각주를 지정하기 전에 출력 데이터셋에 결과를 저장할지를 지정할 수 있습니다.

- 2 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다. 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

.....
주: 결과에 포함되도록 선택한 도표에 따라 일부 제목을 사용할 수 없습니다.
.....

- 3 **마침**을 클릭하여 작업을 실행합니다.

표면도

표면도 작업 정보	543
표면도: 도표 유형 선택	544
표면도: 분석 역할에 변수 할당	544
표면도: 모양 옵션 설정	545
표면도 옵션 설정	545
추가 옵션	545
데이터 평활화	546
제목 및 각주 지정	546
속성 보기	547

표면도 작업 정보

표면도 작업은 3개의 숫자 변수 간의 수학적 관계를 보여 주는 3차원 와이어프레임 도표, 3차원 평활 도표 또는 3차원 색상 도표를 생성합니다. 3D 도표에서 두 개의 수평 변수(x와 y)로 지정된 평면 상의 위치를 수직 변수(z)로 나타냅니다. 각 점의 좌표는 입력 데이터셋의 관측값에 있는 3개의 숫자 변수값에 대응됩니다. 관측값에는 $z = f(x, y)$ 형태의 값 또는 주어진 경도 및 위도에서의 고도와 같은 독립적인 값이 포함될 수 있습니다.

자세한 내용은 [“그래픽 출력형식 선택” \(5페이지\)](#)에서 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	G3D
필수 SAS 제품	Base SAS, SAS/GRAPH
추가 SAS 제품(권장)	없음

표면도: 도표 유형 선택

이 옵션에 액세스하려면 선택 영역에서 **표면도**를 클릭합니다.

역할 이름	설명
3D 와이어프레임 표면도	와이어프레임 테이블에 3가지 변수 간의 수학적 관계를 보여 주는 3차원 도표를 생성합니다.
3D 평할 표면도	표면의 색상이 모두 동일한 3차원 도표를 생성합니다.
3D 색상 표면도	색상을 사용하여 값 변경을 나타내는 3차원 도표를 생성합니다. 영역이 분명하게 정의됩니다.
3D 평할 색상 표면도	색상을 사용하여 값 변경을 나타내는 3차원 도표를 생성합니다. 색상 표면도처럼 영역이 분명히 정의되지는 않습니다.

표면도의 표시 방법은 선택한 그래픽 출력형식에 따라 달라집니다.

ActiveX 및 ActiveX 이미지

모든 표면도 유형을 렌더링합니다.

Java 및 Java 이미지

3D 와이어프레임 표면도를 렌더링합니다. 3D 평할 표면도, 3D 색상 표면도 및 3D 평할 색상 표면도는 3D 평할 색상 표면도로 렌더링됩니다.

GIF, JPEG, PNG 및 SAS EMF

모든 표면도 유형을 3D 와이어프레임 표면도로 렌더링합니다.

표면도: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
수평	이 역할에 할당하는 칼럼은 그래프의 가로축 또는 X축 변수입니다.
수직	이 역할에 할당하는 칼럼은 그래프의 세로축 또는 Y축 변수입니다.
깊이	이 역할에 할당하는 칼럼은 그래프의 세 번째 차원 또는 Z축 변수입니다. 주: 이 옵션은 3D 도표에만 사용할 수 있습니다.
그룹 분석 칼럼	이 역할에 할당하는 칼럼의 값에 따라 생성되는 그래프의 수가 결정됩니다. 각 고윳값에 대해 별도의 그래프가 생성됩니다.

표면도: 모양 옵션 설정

표면도 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **표면**을 클릭합니다.

그래프의 위쪽(양의 Z 방향에 접한 쪽)과 아래쪽(음의 Z 방향에 접한 쪽) 색상을 선택합니다.

그래프의 옆면을 실선으로 그릴지를 선택합니다.

슬라이더 또는 **좌우 회전** 및 **상하 회전** 컨트롤을 사용하여 보기 각도를 설정합니다. 좌우 회전 각도와 상하 회전 각도는 70입니다. 보기 각도를 기본 설정으로 되돌리려면 **재설정**을 클릭합니다. 노트

주: **상하 및 좌우 회전 값 지정** 체크 박스를 선택 취소하면 그래프의 상하 및 좌우 회전이 그래프 출력형식에 따라 결정됩니다. 이 옵션은 기본적으로 선택되며 보기 각도를 지정할 수 있습니다.

3D 색상 표면도와 3D 평활 색상 표면도에서는 표면에 데이터 등고선을 그릴지를 선택할 수 있습니다. 표면 위쪽 또는 아래쪽에 등고선을 표시할 수 있습니다.

추가 옵션

일부 옵션은 여러 그래프 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- ["축 옵션 설정" \(593페이지\)](#)

- “그래프 영역 사용자 정의” (597페이지)

데이터 평활화

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **데이터 그리드**를 클릭합니다.

다음과 같을 때 보간 또는 평활된 값의 사각 그리드를 생성할 수 있습니다.

- 입력 테이블의 데이터 포인트가 불규칙하게 분포할 때
- 입력 테이블의 행 수가 부족하여 3차원 표면도를 생성할 수 없을 때
- 입력 테이블의 데이터에 잡음이 많을 때

기본적으로 보간법은 사용되지 않습니다. 하지만 다음과 같은 보간법을 선택할 수 있습니다.

Join

입력 데이터셋에서 생성된 삼각 영역에 선형보간을 적용합니다. 이 보간법은 수직 변수의 초기 값 범위 내의 값을 생성하지만 생성되는 보간 표면은 평활하지 않습니다.

No scale

보간 전에 X 및 Y 변수를 동일한 범위로 확대/축소하지 않습니다. 보간 시에는 X와 Y가 동등한 것으로 가정하기 때문에, 기본적으로 두 변수가 모두 비슷하게 확대/축소된 후 보간이 수행됩니다.

Partial

5 차 다항 보간을 위한 미분계수를 추정하는 데 사용할 스플라인을 지정합니다. 이변량 스플라인은 가장 인접한 데이터에 적합하여 필요한 미분계수를 추정하는 데 사용됩니다.

Spline

보간할 이변량 스플라인 사용을 지정합니다. 이 옵션을 사용하면 차수 n3 알고리즘이 사용됩니다. 여기서 n은 입력 데이터 점의 수입니다. 이 방법은 시간이 많이 소요될 수 있습니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

t 검정

<i>t</i> 검정 작업 정보.....	549
<i>t</i> 검정: <i>t</i> 검정 유형 선택.....	550
<i>t</i> 검정: 분석 역할에 변수 할당.....	550
<i>t</i> 검정: 분석 옵션 설정.....	551
<i>t</i> 검정: 도표 생성.....	552
제목 및 각주 지정.....	552
속성 보기.....	553

t 검정 작업 정보

t 검정은 일표본, 이표본, 쌍체 관측값에 대하여 *t* 검정을 수행합니다. 일표본 *t* 검정은 표본 평균을 주어진 수치와 비교합니다. 이표본 *t* 검정은 (첫 번째 표본평균 - 두 번째 표본평균) 값을 주어진 수치와 비교합니다. 쌍체 *t* 검정은 관측값 차의 평균을 주어진 수치와 비교합니다.

때때로 두 그룹의 데이터가 독립적이고 데이터의 고유쌍이 존재한다고 추정하는 것이 불가능할 때가 있습니다. 이때 쌍체 *t* 검정을 사용하는 것이 유용합니다. 예를 들어 자극의 수축기 혈압에 대한 영향을 실험한다고 가정합니다. 12명의 사람이 연구에 참여하고 수축기 혈압이 자극 전후로 모두 측정됩니다. *t* 검정 작업을 사용하여 수축기 혈압의 평균 변화가 0에서 크게 달라지는지를 결정합니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	TTEST
필수 SAS 제품	Base SAS, SAS/STAT
추가 SAS 제품(권장)	SAS/GRAPH

t 검정: t 검정 유형 선택

이 옵션에 액세스하려면 선택 영역에서 **t 검정 유형**을 클릭합니다.

요구 사항 이름	프로시저 및 제품 이름
이표본	(첫 번째 표본의 평균 - 두 번째 표본의 평균) 값을 주어진 수치와 비교합니다.
쌍체	관측값 차의 평균을 주어진 수치와 비교합니다.
일표본	표본평균을 주어진 수치와 비교합니다.

t 검정: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
분류변수	이표본 t 검정에 사용할 수 있습니다. 이 역할에 지정되는 분류변수는 독립 이표본 분석에 사용됩니다. 변수에는 각 표본에 대해 하나씩 두 개의 레벨만 있어야 합니다. 이 역할에 할당할 수 있는 변수는 최대한 개입니다.
분석변수	일표본 및 이표본 t 검정에 사용할 수 있습니다. 이 리스트의 각 변수에 대해 t 검정이 수행됩니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
쌍체 변수	쌍체 t 검정에 사용할 수 있습니다. 쌍체 t 검정을 수행하려면 이 역할에 분석할 두 개의 변수를 지정합니다. 쌍체 t 검정은 첫 번째 변수값에서 두 번째 변수값을 빼는 방식으로 수행되므로, 변수의 순서가 중요합니다. 주: 이 역할에는 두 개의 변수만 지정해야 합니다.

역할 이름	설명
그룹 분석변수	이 역할에 변수를 지정하면 선택한 변수를 기준으로 테이블이 정렬되고, 각 그룹에 대해 별도의 t 검정이 수행됩니다. 주: 분석변수 또는 쌍체 변수로 이미 선택한 변수로는 t 검정을 그룹화할 수 없습니다.
빈도 수	이 역할에 변수를 지정하면 테이블의 각 행은 n개의 관측값을 나타내게 되며, 여기서 n은 해당 행의 빈도변수값입니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다.
상대 가중값 변수	이 역할에 변수를 할당하면 각 통계량의 변수값이 가중 통계량 계산에 사용됩니다. 이 역할에 할당할 수 있는 변수는 최대 한 개입니다. 주: 이 작업의 입력 데이터셋에 요약통계량만 포함되어 있으면 상대 가중값 변수 역할을 사용할 수 없습니다. 입력 데이터셋에 요약통계량이 포함되어 있으면 가중값 옵션이 적용되지 않습니다.

t 검정: 분석 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **분석**을 클릭합니다.

옵션 이름	설명
귀무가설 지정	기본적으로 귀무가설의 값은 0입니다. 새로운 귀무가설 값을 지정하려면 $H_0 =$ 상자에 새로운 값을 입력합니다.
표준편차 신뢰구간 옵션 설정	다음과 같이 신뢰구간 옵션을 설정합니다. ■ 양측 체크 박스는 기본적으로 선택됩니다. 양측 신뢰구간을 선택하면 카이제곱 분포의 각 끝에 동일한 영역이 놓이게 됩니다. 이 옵션은 양측 신뢰구간을 표시합니다. ■ UMPU 체크 박스를 선택하면 α에 대한 균등최강력 불편검정에서 유도된 신뢰구간이 생성됩니다. 주: 신뢰구간이 표시되지 않게 하려면 양측 및 UMPU 체크 박스의 선택을 모두 취소합니다.

t 검정: 도표 생성

이 옵션에 액세스하려면 선택 영역에서 **도표**를 클릭합니다.

도표 이름	설명
요약 도표	요약 도표를 생성합니다. 요약 도표의 내용은 수행하는 t 검정 유형에 따라 다릅니다. ■ 일표본 t 검정에서는 요약 도표에 정규 및 커널밀도함수가 표시된 히스토그램, 상자 도표 및 신뢰구간 띠가 포함됩니다. ■ 이표본 t 검정에서는 요약 도표에 밀도함수가 표시된 비교 히스토그램과 상자 도표가 포함됩니다. ■ 쌍체 t 검정에서는 요약 도표에 히스토그램, 밀도함수, 상자 도표, 차이 또는 비의 신뢰구간이 포함됩니다.
히스토그램	정규 및 커널밀도함수가 표시된 히스토그램 또는 비교 히스토그램을 생성합니다. 일표본 t 검정 및 쌍체 t 검정의 경우 히스토그램과 밀도는 검정 기준(쌍체 설계에 대한 평균 차 또는 비율)을 기반으로 합니다. 이표본 t 검정의 경우 비교 히스토그램(클래스당 하나씩)이 표시됩니다.
상자 도표	상자 도표 또는 비교 상자 도표를 생성합니다. 상자는 25번째 백분위수(제1 사분위수)에서 75번째 백분위수(제3 사분위수)까지 표시됩니다. 상자 내 수직선은 중위수 위치를 나타냅니다. 일표본 및 쌍체 t 검정에서 평균에 대한 신뢰구간은 배경에 구간으로 표시됩니다. 이표본 t 검정의 경우 비교 상자 도표(클래스당 하나씩)가 표시됩니다.
신뢰구간 도표	평균에 대한 신뢰구간 도표를 생성합니다.
정규 분위수-분위수(Q-Q) 그림	정규 분위수-분위수(Q-Q) 도표를 생성합니다. 이표본 t 검정에서는 하나의 패널에 각 클래스에 대한 별도의 도표가 표시됩니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

.....

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

.....

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

.....

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

.....

테이블 분석

테이블 분석 작업 정보	555
테이블 분석: 분석 역할에 변수 할당	556
테이블 분석: 테이블 정의	556
테이블 분석: 셀 통계량 선택	557
테이블 분석: 테이블 통계량 지정	558
연관성 검정 수행	558
일치성 검정 수행	559
순서화 차이 검정 수행	560
추세 검정 수행	560
계산 옵션 설정	561
테이블 분석: 결과 옵션 설정	562
셀 통계량 저장	562
테이블 통계량 저장	562
제목 및 각주 지정	563
속성 보기	563

테이블 분석 작업 정보

테이블 분석 작업을 통해 데이터로부터 교차 테이블(분할표라고도 함)을 생성할 수 있습니다.

예를 들어 두 개의 다른 식이요법에 대해 관상심장질환의 가능성을 비교한다고 가정합니다. SAS 데이터에 고지방 식이요법과 관상심장질환의 위험 간의 사례-대조 연구에 대한 데이터가 있습니다. 데이터는 셀 수로 기록되고 변수 수에는 각 식이요법과 위험 조합에 대한 빈도가 포함됩니다. 테이블 분석 작업을 사용하여, 각 식이요법(고지방 및 저지방)과 위험(고위험 및 저위험)에 대한 빈도값을 나타내는 2x2 분할표를 생성할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	FREQ
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	없음

테이블 분석: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
빈도 수	테이블의 각 행에 대한 빈도가 포함된 변수를 지정합니다. 각 역할은 n 개의 관측값을 나타내게 되며, 여기서 n 은 해당 행의 빈도변수값입니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 다른 역할에 할당한 변수로는 분석을 그룹화할 수 없습니다.
테이블 변수	이러한 변수는 이원 - n 원 빈도 및 교차 테이블을 생성하는 데 사용됩니다. 테이블 변수에 대한 통계량이 계산됩니다. 이 역할에는 적어도 두 개의 변수를 지정해야 합니다.

테이블 분석: 테이블 정의

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

새로운 테이블을 생성하는 방법:

- 1 **테이블에 사용 가능한 변수** 상자에서 테이블에 추가할 변수를 선택합니다. 이 리스트의 변수는 **테이블 변수** 역할에 지정한 변수입니다.

2. 선택한 변수를 **미리 보기** 상자로 끌어 놓습니다. 첫 번째로 끌어 놓은 변수값은 교차 테이블의 변수가 되고, 두 번째로 끌어 놓은 변수값은 교차 테이블의 행이 됩니다.

셋 이상의 변수를 **미리 보기** 상자로 끌면, 끌어온 세 번째(또는 그 이후) 변수의 각 값 조합에 대해 하나씩 여러 개의 교차 테이블이 생성됩니다.

주: 변수를 테이블에 추가한 후에는 **테이블에 사용 가능한 변수** 상자에서 변수가 제거됩니다.

테이블에서 변수를 제거하려면 **미리 보기** 상자에서 변수를 선택하여 **테이블에 허용된 변수** 상자로 끌어 놓습니다.

새로운 테이블을 시작하려면 정의된 테이블 리스트에서 < **새로운 테이블을 정의하려면 여기를 클릭합니다** >를 선택합니다.

테이블 분석: 셀 통계량 선택

이 옵션에 액세스하려면 선택 영역에서 **셀 통계량**을 클릭합니다.

선택된 통계량은 교차 테이블의 각 셀에 표시됩니다.

옵션 이름	설명
누적칼럼 백분율	각 셀에 누적칼럼 백분율을 표시합니다.
행 백분율	각 셀에 대한 행 백분율을 표시합니다.
칼럼 백분율	각 셀에 대한 칼럼 백분율을 표시합니다. 기본적으로 이러한 통계량은 출력에 포함됩니다.
셀 빈도	각 셀에 대한 빈도변수를 표시합니다. 기본적으로 이러한 통계량은 출력에 포함됩니다.
셀 백분율	교차 테이블의 백분율, 행 백분율, 칼럼 백분율 또는 일원빈도 테이블의 백분율과 누적백분율을 표시합니다.
결측값 빈도	결측값의 빈도를 표시합니다.
각 셀의 카이제곱 통계량	전체 피어슨 카이제곱 통계량에 대한 각 셀의 기여도를 표시합니다.
기대빈도로부터의 편차	각 셀에 대해 기대빈도로부터의 셀 빈도의 편차를 표시합니다.
각 셀의 기대빈도	각 셀에 대하여 기대빈도를 표시합니다.
전체 빈도의 백분율	n 원 테이블($n > 2$)에서의 전체 빈도 백분율을 표시합니다.

옵션 이름	설명
데이터셋에 백분율 포함	출력 데이터셋의 칼럼 및 행에 대한 백분율을 표시합니다.

테이블 분석: 테이블 통계량 지정

연관성 검정 수행

이 옵션에 액세스하려면 선택 영역의 테이블 통계량 머리글 아래에서 **연관성**을 클릭합니다.

주: **정확 p-값**을 선택하면 정확 p -값을 계산하는 데 메모리가 많이 소요되고 처리 시간이 길어질 수 있으므로 **계산 시간 제한** 옵션을 사용하여 정확 p -값 계산의 제한 시간을 지정하십시오.

옵션 이름	설명
연관성 검정	다음 옵션을 지정할 수 있습니다. ■ 카이제곱 검정 - 테스트 동질성 또는 독립성에 대한 카이제곱 검정, 카이제곱을 기반으로 한 연관성 측도를 계산합니다. Pearson 카이제곱, 우도비 카이제곱 및 Mantel-Haenszel 카이제곱 검정이 포함됩니다. 2×2 테이블에서 이 검정에는 Fisher의 정확 검정 및 연속성-수정된 카이제곱이 포함되며, ■ 정확 p-값 - 일원테이블에서는 카이제곱 적합도 검정, 이원테이블에서는 Pearson 카이제곱, 우도비 카이제곱, Mantel-Haenszel 카이제곱 검정통계량에 대해 정확 p-값을 계산합니다. ■ $r \times c$ 테이블에 대한 Fisher의 정확 검정 - 2×2를 넘는 테이블에 대해 Fisher의 정확 검정을 계산합니다. 이 검정은 Freeman-Halton 검정이라고도 합니다.
연관성 측도	다음 옵션을 지정할 수 있습니다. ■ 측도 - 여러 가지 연관성 측도 및 근사표준오차(ASE)를 계산합니다. 감마, Kendall의 tau-b, Stuart의 tau-c, Somers의 $D(C R)$, Somers의 $D(R C)$, Pearson 및 Spearman 상관계수, 람다(대칭과 비대칭), 불확실성계수(대칭과 비대칭) 측도가 포함됩니다. 2×2 테이블에서 이 측도에서는 오즈비, 칼럼 1 상대위험도, 칼럼 2 상대위험도 및 해당 신뢰한계도 제공됩니다. ■ 오즈비에 대한 정확 p-값과 신뢰경계 - Pearson 상관계수 및 Spearman 상관계수에 대한 정확 p-값, 2×2 테이블에 대한 오즈비 신뢰한계를 계산합니다.

옵션 이름	설명
	■ 측도=0에 대한 검정은 감마, Kendall tau-b, 스튜어트 tau-c, Somers' $D(C R)$, Somers' $D(R C)$, Pearson 상관계수 및 Spearman 상관계수에 대한 점근 검정을 계산합니다. ■ 2 x 2 테이블에 대한 위험도 차 - 2x2 테이블에 대한 위험 추정값(또는 이항비) 및 위험도 차를 제공합니다. 이 분석은 두 그룹의 일부 특성에 대한 비율을 비교할 때(1행과 2행이 두 그룹에 대응되고, 칼럼은 두 개의 가능한 특성 또는 결과에 대응) 적합합니다. ■ 2 x 2 테이블에 대한 상대위험도 - 2x2 테이블에 대한 상대위험도 측도 및 신뢰한계를 계산합니다. 오즈비 및 칼럼 1과 2 상대위험도 측도가 포함됩니다.
Cochran-Mantel-Haenszel 통계량	Cochran-Mantel-Haenszel 통계량을 모두 계산하려면 CMH 통계량 체크 박스를 선택합니다. 이러한 통계량에는 CMH 상관관계 통계량, 행 평균 스코어(ANOVA), 수정 상대위험/오즈비가 포함됩니다. 기본 스코어 유형은 테이블 입니다. 스코어 유형을 변경하려면 스코어 유형 옵션을 사용하십시오.

일치성 검정 수행

이 옵션에 액세스하려면 선택 영역의 테이블 통계량 머리글 아래에서 **일치성**을 클릭합니다. $n \times n$ 테이블에 대해 다음과 같은 일치성 검정 및 측도를 선택할 수 있습니다.

옵션 이름	설명
측도	정방 테이블에 대한 분류 일치성의 검정 및 측도를 계산합니다. 이 옵션은 2x2 테이블에 대한 McNemar 검정 및 반응 범주가 둘 이상인 테이블에 대한 Bowker의 대칭성 검정을 제공합니다. 단순 카파 계수, 가중 카파 계수, 단순/가중 카파에 대한 근사 표준오차 및 해당 신뢰한계도 제공됩니다. 여러 계층 및 두 개의 반응 범주가 있으면 이 옵션은 Cochran의 Q 검정도 계산합니다.
정확 p-값	2x2 테이블에 대한 McNemar 검정의 정확 p-값, 단순 카파 계수 및 가중 카파 계수를 계산합니다. 주: 정확 p-값을 선택하면 정확 p-값을 계산하는 데 메모리가 많이 소요되고 처리 시간이 길어질 수 있으므로 계산 시간 제한 옵션을 사용하여 정확 p-값 계산의 제한 시간을 지정하십시오.
측도=0에 대한 검정	단순 카파 계수 및 가중 카파 계수에 대한 근사 검정을 계산합니다.

순서화 차이 검정 수행

이 옵션에 액세스하려면 선택 영역의 테이블 통계량 머리글 아래에서 **순서화 차이**를 클릭합니다.

옵션 이름	설명
Jonckheere-Terpstra 검정	분류 간의 순서화 차이를 비모수적으로 검정하는 Jonckheere-Terpstra 검정을 계산합니다. 반응변수 분포가 분류 간에서 다르지 않다는 귀무가설을 검정합니다.
정확 p-값	Jonckheere-Terpstra 검정에 대해 정확 p -값을 계산합니다. 주: 정확 p-값 을 선택하면 정확 p -값을 계산하는 데 메모리가 많이 소요되고 처리 시간이 길어질 수 있으므로 계산 시간 제한 옵션을 사용하여 정확 p -값 계산의 제한 시간을 지정하십시오.

추세 검정 수행

이 옵션에 액세스하려면 선택 영역의 테이블 통계량 머리글 아래에서 **추세 검정**을 클릭합니다.

옵션 이름	설명
Cochran-Armitage 검정	단일 인수 또는 공변량으로 이루어진 레벨에서 이항비 추세를 검정하는 Cochran-Armitage 추세 검정을 계산합니다. 이 검정은 한 변수에 두 개의 레벨이 있고 다른 변수는 순서를 나타내는 분할표에 적합한 방법입니다. 두 레벨 변수는 반응을 나타내며, 다른 변수는 순서화된 레벨을 가진 설명변수를 나타냅니다.
정확 p-값	Cochran-Armitage 추세 검정에 대한 정확 p -값을 계산합니다. 주: 정확 p-값 을 선택하면 정확 p -값을 계산하는 데 메모리가 많이 소요되고 처리 시간이 길어질 수 있으므로 계산 시간 제한 옵션을 사용하여 정확 p -값 계산의 제한 시간을 지정하십시오.

계산 옵션 설정

이 옵션에 액세스하려면 선택 영역의 테이블 통계량 머리글 아래에서 **계산 옵션**을 클릭합니다.

옵션 이름	설명
스코어 유형	연관성, 일치성 또는 추세 검정을 수행하려면 스코어 유형을 선택하십시오. 스코어 유형은 Mantel-Haenszel 카이제곱 통계량, Cochran-Mantel-Haenszel 통계량, Pearson 상관계수, Cochran-Armitage 추세 검정, 가중 카파 계수에서 사용할 행 스코어 및 칼럼 스코어 유형을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 테이블 - 숫자 변수에서는 테이블 스코어는 행 및 칼럼 레벨 값입니다. 행 또는 칼럼 변수의 출력형식을 지정하면 테이블 스코어는 레벨에 해당하는 내부 숫자값입니다. 둘 이상의 숫자값을 동일한 출력형식 레벨로 분류하면 해당 레벨의 내부 숫자값이 최솟값이 됩니다. - 문자 변수에서는 테이블 스코어를 행 번호와 칼럼 번호로 정의합니다 (즉, 첫 번째 행은 1, 두 번째 행은 2, ...). ■ 순위 - 순위 스코어는 행 스코어 및 칼럼 스코어로 정의됩니다. ■ RIDIT - RIDIT 스코어는 순위 스코어에서 파생됩니다. ■ 수정 RIDIT - 수정 Ridit(MODRIDIT) 스코어는 (0,1)의 균등분포에 대한 차수통계량의 예상값을 나타냅니다. 수정된 RIDIT 스코어는 순위 스코어에서 파생됩니다.
정확 계산	검정에서 정확 p-값을 계산하면 정확 p-값을 계산하는 데 메모리가 많이 소요되고 처리 시간이 길어질 수 있습니다. 정확 계산 영역에서 계산 시간 제한 체크 박스를 선택하고 각 테이블의 p-값 계산에 대한 시간 제한을 초 단위로 지정합니다. 기본값은 900초입니다. 정확 p-값을 직접 계산하지 않고 정확 p-값의 Monte Carlo 추정값을 계산하려면 Monte Carlo 추정 사용 체크 박스를 선택합니다. Monte Carlo 추정은 정확한 계산을 위해 상당한 시간과 메모리를 요구하는 복잡한 문제에 유용합니다.
계산에 결측값 포함	백분율 및 기타 통계량 계산에 결측값을 포함할 수 있습니다.
신뢰수준	계산에 대한 신뢰수준을 지정할 수 있습니다. 기본 신뢰수준은 95% 신뢰한계입니다.
값 순서	교차 테이블에서 변수값이 출력되는 순서를 지정할 수 있습니다. 값 순서에서는 다음 옵션을 설정할 수 있습니다. ■ 데이터셋 순서 - 데이터셋에 나오는 순서에 따라 값을 정렬합니다. ■ 서식화된 값 - 출력형식 값의 오름차순으로 값을 정렬합니다.

옵션 이름	설명
	■ 빈도 내림차순 - 빈도변수가 가장 높은 레벨이 맨 앞에 오도록 값을 정렬합니다. ■ 서식화되지 않은 값 - 변수의 입력값에 따라 값을 정렬합니다. 기본값입니다.

테이블 분석: 결과 옵션 설정

셀 통계량 저장

이 옵션에 액세스하려면 선택 영역의 결과 머리글 아래에서 **셀 통계량 결과**를 클릭합니다.

옵션 이름	설명
출력 데이터 저장	원하는 교차 테이블의 체크 박스를 선택합니다. 표시된 테이블은 테이블 변수 역할에 할당된 변수입니다. 셀 통계량 데이터셋 상자에서 선택한 테이블에 대한 셀 통계량을 저장할 데이터셋을 지정합니다. 데이터셋 이름은 선택한 각 테이블에 대해 변경됩니다. 주: SAS Enterprise Guide는 옵션 창의 작업 > 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
모든 출력 숨기기	결과를 표시하지 않습니다.

테이블 통계량 저장

이 옵션에 액세스하려면 선택 영역의 결과 머리글 아래에서 **테이블 통계량 결과**를 클릭합니다.

옵션 이름	설명
다음 테이블의 통계량 저장	하나의 교차 테이블에 대한 테이블 통계량을 저장합니다. 테이블 통계량 데이터셋 상자에서 선택한 테이블에 대한 테이블 통계량을 저장할 데이터셋을 지정합니다. 데이터셋 이름은 선택한 각 테이블에 대해 변경됩니다. 주: SAS Enterprise Guide는 옵션 창의 작업> 출력 라이브러리 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다. SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 찾아보기를 클릭합니다.
리스트 형식으로 테이블 표시	각 계층에 대한 개별 이원테이블을 표시하지 않고 하나의 테이블에 전체 다원 테이블을 표시합니다. 주: 테이블 통계량을 생성하는 옵션을 선택하면 리스트 형식으로 테이블 표시 옵션이 무시됩니다.
모든 출력 숨기기	결과를 표시하지 않습니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

타일 그래프

타일 그래프 작업 정보	565
타일 그래프: 분석 역할에 변수 할당	566
타일 그래프: 모양 옵션 설정	567
타일 그래프의 모양 지정	567
범례 옵션 설정	568
타일 그래프: 고급 옵션 설정	568
제목 및 각주 지정	570

타일 그래프 작업 정보

타일 그래프 작업으로 타일 그래프를 생성하여 대용량의 계층 데이터를 제한된 공간에서 볼 수 있습니다. 고유한 각 범주 조합이 사각 타일로 표시되며 이 타일의 크기 및 색상은 반응 변수에 따라 결정됩니다. 이러한 타일은 계층적 배열로 배치됩니다.

타일 그래프는 사각 트리 지도입니다. 타일 그래프는 트리로 표시되므로 웹과 같은 대화식 환경에서 특히 유용합니다. 대화식 환경에서 타일 그래프가 제공하는 이점은 다음과 같습니다.

- 원본 그래프의 작은 부분을 자세히 보려면 드릴다운(또는 확대/축소) 기능을 사용할 수 있습니다.
- 타일 위에 마우스 포인터를 놓으면 특정 타일에 대한 메트릭이 데이터 팁에 표시됩니다.
- 특정 레벨에서 타일을 리포트 또는 웹 페이지에 연결할 수 있습니다. 해당 레벨에서 타일을 클릭하면 타일과 연결된 리포트 또는 웹 페이지가 열립니다.

타일 그래프 작업에서는 ActiveX, Java, ActiveX 이미지 및 Java 이미지 그래프 출력형식만 허용됩니다. 현재 그래프 출력형식이 이러한 출력형식이 아니면 ActiveX 그래프 출력형식이 사용됩니다.

HTML 결과의 경우 특정 그래프 출력형식에 대해서만 일부 기능을 사용할 수 있습니다.

- 드릴다운 기능은 그래픽 출력형식으로 ActiveX 또는 Java를 선택할 때 사용할 수 있습니다.

- 타일 그래프에서 웹 페이지에 연결하는 기능은 그래픽 출력형식으로 Java를 선택할 때 사용할 수 있습니다.

옵션 대화 상자에서 그래프 출력형식 및 결과의 출력형식을 지정할 수 있습니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	GTILE
필수 SAS 제품	Base SAS, SAS/GRAPH
추가 SAS 제품(권장)	없음

타일 그래프: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **작업 역할**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

주: 선택한 그래프 유형에 따라 사용할 수 있는 역할이 다릅니다.

역할 이름	설명
분류변수	계층 범주(또는 ID) 변수를 지정합니다. 분류변수는 반드시 지정해야 하며 이 역할에 최대 6개의 변수를 할당합니다. 이 역할에 여러 변수를 할당하면 변수 순서에 따라 계층이 정의됩니다. 예를 들어 분류변수 역할의 변수 순서가 국가, 지역 및 부서이면 결과 계층은 국가 > 지역 > 부서입니다.
색상 분석 변수	타일 색상을 결정하는 변수값을 지정합니다. 최소 색상 리스트 값에는 가장 작은 색상 반응이 할당되고 최대 색상 리스트 값에는 가장 큰 색상 반응이 할당됩니다. 리스트의 색상 간 그라데이션에는 다른 색상 반응이 선형으로 할당됩니다. 이 역할에는 변수를 하나만 할당할 수 있습니다. 주: 색상 분석변수로 할당한 변수가 없으면 크기 분석 변수 역할에 할당하는 변수가 색상 분석변수로도 사용됩니다.
크기 분석 변수	타일의 상대적 크기를 결정하는 변수값을 지정합니다. 모든 타일의 총 크기에는 전체 타일 영역이 포함됩니다. 결측값, 0 및 음수 값을 지정하면 해당 타일이 무시됩니다. 크기 분석변수는 반드시 지정해야 하며 이 역할에는 변수를 하나만 할당할 수 있습니다.

역할 이름	설명
그룹 분석 칼럼	이 역할에 할당하는 칼럼의 값에 따라 생성되는 그래프의 수가 결정됩니다. 각 고유티값에 대해 별도의 그래프가 생성됩니다. 이 역할에는 둘 이상의 변수를 할당할 수 있습니다.

타일 그래프: 모양 옵션 설정

타일 그래프의 모양 지정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **타일 그래프**를 클릭합니다.

역할 이름	설명
타일 레이아웃	그래프의 타일 레이아웃을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 플로우 - 타일 레이아웃을 정렬된 여러 행으로 분할합니다. 그래프에서 정렬된 타일이 왼쪽에서 오른쪽, 위에서 아래의 순서로 표시됩니다. ■ 스퀘어 - 기본 직사각형 트리 맵입니다. 일반적으로 타일이 큰 순서대로, 그리고 그래프의 왼쪽 맨 아래에서 오른쪽 맨 위 방향으로 배열됩니다. 데이터 변수에 대한 정렬은 무시되며 타일의 모양은 정사각형에 가깝습니다. ■ 토글 - 분류변수 역할에 할당한 변수를 통해 생성된 각 계층 레벨에 대한 방향을 분할합니다. 계층의 최상위 레벨은 사용 가능한 공간이 수직 조각으로 분할되어 단일 축을 따라 정렬된 순서로 표시됩니다. 하위 레벨에서는 분할 방향이 변경(또는 토글)되어 두 번째 레벨이 수평으로 분할됩니다.
분류 레벨 지정	결과에 표시할 레벨의 개수를 지정합니다. 예를 들어 분류변수 역할의 변수 순서가 국가, 지역 및 부서이면 결과 계층은 국가 > 지역 > 부서입니다. 타일 그래프에 표시되는 레벨 수는 분류 레벨 지정 옵션의 값에 따라 달라집니다. ■ 이 옵션을 1로 설정하면 계층의 레벨 하나가 결과에 표시됩니다. 즉, 캐나다, 독일, 미국 같은 개별 국가에 대한 타일이 타일 그래프에 표시됩니다. 캐나다 타일로 드릴다운하면 캐나다의 서부 및 동부 지역에 대한 타일이 타일 그래프에 표시됩니다. ■ 이 옵션을 2로 설정하면 계층의 두 레벨이 결과에 표시됩니다. 즉, 국가 및 지역(캐나다 서부, 캐나다 동부, 독일 서부, 독일 동부, 미국 서부, 미국 동부 등)에 대한 타일이 타일 그래프에 표시됩니다. 캐나다

역할 이름	설명
	타일로 드릴다운하면 캐나다 서부와 캐나다 동부 지역에 있는 모든 부서에 대한 타일이 표시됩니다. 기본적으로 계층의 모든 레벨이 표시됩니다.
색상 지정	타일 그래프에 대한 색상 램프를 사용자 정의합니다. 색상 램프는 전체 데이터 값 범위에서의 색상 분포를 나타냅니다. 기본적으로 두 개 이상의 색상을 지정해야 합니다. 이러한 색상은 색상 램프의 시작 색상과 마지막 색상이 됩니다. 기본적으로 색상 램프의 색상은 그라데이션 범례 전체에 고루 배포됩니다. 색상 포인트를 사용하여 데이터 전체에 색상을 배포하는 방법을 지정할 수 있습니다. 색상 포인트를 지정하려면 편집을 클릭합니다. 색상 포인트 편집 대화 상자에 있는 값의 범위는 0~1이며 오름차순이어야 합니다. 예를 들어 색상 램프에 세 가지 색상을 지정하면 0, .75 및 1을 색상 포인트로 지정할 수 있습니다. 여러 가지 다른 색상 포인트에 대해 같은 값을 지정할 수 있습니다. 예를 들어 색상 램프에 다섯 가지 색상을 지정하면 0, .2, .4, .6을 색상 포인트로 지정할 수 있습니다. 색상 포인트의 개수는 색상 램프의 색상 개수와 동일하게 지정해야 합니다.

범례 옵션 설정

범례의 시작 및 마지막 포인트를 지정하려면 **범례 범위 지정** 체크 박스를 선택합니다. 데이터의 값 범위가 0에서 100,000까지라고 가정합니다. 그러나 타일 그래프에서 색상 매핑으로 22,000~26,000의 값을 강조 표시하려고 합니다. 이 예에서 22,000을 최소(시작)값으로 지정하고 26,000을 최대(마지막)값으로 지정할 수 있습니다. 타일 값이 22,000 미만이거나 26,000을 초과하면 회색으로 표시됩니다. 값이 22,000~26,000에 해당하는 타일은 타일 그래프에 대해 지정한 색상 램프를 사용합니다.

기본적으로 범례의 색상 매핑에는 반응 데이터의 색상이 사용됩니다. 그러나 **범례 색상 매핑이 고정 값에 대해 대칭이 되도록 설정** 체크 박스를 선택하여 색상 매핑이 고정 값의 양쪽에 대해 대칭이 되는지를 지정할 수 있습니다. 이 옵션을 사용하여 색상 척도의 중간점을 지정할 수 있습니다. 기본적으로 대칭 점 또는 중간점은 0입니다. 새로운 대칭 점 또는 중간점을 지정할 수 있습니다.

타일 그래프: 고급 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **고급**을 클릭합니다.

주: 그래픽 출력에 대한 출력형식으로 Java를 선택하고 **SAS Enterprise Guide**에서 결과에 대한 출력형식으로 HTML을 선택할 때만 타일 그래프에서 웹 페이지에 연결할 수 있습니다. 그래프 출력형식과 결과의 출력형식을 지정하려면 **옵션** 대화 상자의 **결과** ⇒ **그래프** 패

널을 참조하십시오. Java 요구 사항으로 인해 이 링크는 지원되지 않습니다.**SAS Add-In for Microsoft Office .**

타일 그래프에서 웹 페이지에 대한 연결을 추가할 수 있습니다. 선택한 연결 변수의 레벨에서 타일을 마우스 오른쪽 버튼으로 클릭할 때 표시되는 팝업 메뉴에서 연결을 사용할 수 있습니다. **연결 변수**는 분류변수 역할에 할당된 변수입니다. 이러한 연결이 표시되는 레벨은 분류 레벨을 지정했는지 여부에 따라 달라집니다.

- 분류 레벨을 지정하지 않으면 타일 그래프의 모든 레벨에서 링크를 사용할 수 있습니다. 예를 들어 타일 그래프의 계층이 국가 > 지역 > 부서이면 국가, 지역 및 부서 레벨에서 타일을 마우스 오른쪽 버튼으로 클릭하여 연결을 사용할 수 있습니다.
- 분류 레벨을 지정하면 타일 그래프의 가장 낮은 레벨에서만 링크를 사용할 수 있습니다. 예를 들어 타일 그래프의 계층에는 국가, 지역 및 부서의 세 가지 레벨이 있습니다. 분류 레벨이 2이면 링크를 가장 낮은 두 레벨인 지역 및 부서에서 사용할 수 있습니다. 분류 레벨이 1이면 가장 낮은 레벨인 부서에서만 링크를 사용할 수 있습니다.

타일 그래프 작업에서 연결 변수는 **사용 가능한 변수** 상자에 나열되어 있습니다. 각 연결 변수에는 치환 번호가 자동으로 할당됩니다. 이 번호는 **사용 가능한 변수** 상자에서 관련 변수 옆에 표시됩니다. 이 치환 번호를 웹 연결 URL 또는 레이블에 포함하여 연결 변수의 각 값에 대한 고유한 연결을 생성할 수 있습니다.

예를 들어 데이터에서 국가 값은 fr(프랑스), jp(일본) 및 uk(영국)입니다. 대체 번호가 1이고 국가 변수에 자동으로 할당되고 사용자가 `http://www.sales.{&1}.com`을 웹 링크에 대한 URL로 지정하고 타일 그래프에서 다음 링크를 사용할 수 있습니다.

- `http://www.sales.fr.com`
- `http://www.sales.jp.com`
- `http://www.sales.uk.com`

그래프에 연결을 포함하는 방법:

- 1 연결에 대한 레이블을 지정합니다. 연결 변수의 각 값에 대해 고유한 레이블을 생성하려면 레이블에 연결 변수의 치환 번호를 포함할 수 있습니다. 레이블에서 연결 변수에 대한 치환 번호를 추가하려면 **변수 삽입**을 클릭합니다.

다음은 연결 레이블을 지정하는 방법의 예입니다.

- **{&1}에 대한 상태 웹 페이지**
- **{&2}에 대한 Wikipedia 웹 페이지**

- 2 웹 페이지에 대한 URL을 지정합니다. URL에 링크 변수의 대체 번호를 포함하여 각 링크 변수 값에 대해 고유한 웹 링크를 생성할 수 있습니다. 링크 변수를 지정하는 구문은 `{&replacement-number}`입니다. 이 치환 번호는 웹 링크에 대한 URL의 아무 곳에서나 나타날 수 있습니다. URL에서 연결 변수 지정에 대한 구문을 추가하려면 **변수 삽입**을 클릭합니다.

이 치환 값은 웹 연결 URL에서 원하는 위치에 표시할 수 있습니다.

- `http://www.state.{&1}.us`
- `http://en.wikipedia.org/wiki/{&2}`

- 3 **추가**를 클릭하여 현재 타일 그래프에 대한 웹 링크 리스트에 이 웹 링크를 추가합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

전치

전치 작업 정보	571
전치: 분석 역할에 변수 할당	572
전치: 분석 옵션 설정	573
전치: 결과 옵션 설정	573
속성 보기	574

전치 작업 정보

데이터 전치 작업은 입력 데이터셋에서 선택한 칼럼을 출력 데이터셋의 행으로 변환합니다. 그룹 변수를 사용하지 않으면 선택된 각 칼럼은 한 행이 됩니다. 그룹 변수를 사용하면 선택된 칼럼은 그룹 변수의 값을 기반으로 하위칼럼으로 나뉘며, 각 하위칼럼은 출력 데이터셋의 한 행이 됩니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	TRANSPOSE
필수 SAS 제품	Base SAS
추가 SAS 제품(권장)	SAS/STAT

전치: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
전치 변수	이 역할에 지정되는 각 칼럼은 출력 데이터셋의 행이 됩니다. 그룹 변수를 선택하지 않으면 전체 칼럼이 한 행이 됩니다. 그룹 변수를 선택하면 그룹 변수에 의해 각 칼럼이 하위칼럼으로 분할되고, 각 하위칼럼이 한 행이 됩니다. 이때 칼럼은 그룹 변수로 정의된 그룹 수와 동일한 수의 행으로 전치됩니다. 전치 변수 역할에는 하나 이상의 칼럼을 지정해야 합니다. 그룹 변수를 선택하려면 그룹 분석변수 역할에 칼럼을 지정합니다.
COPY 변수	이 역할에 지정되는 각 칼럼은 입력 데이터셋에서 출력 데이터셋으로 전치 없이 직접 복사됩니다. 칼럼이 출력 데이터셋으로 직접 복사되기 때문에 출력 데이터셋의 행 수와 입력 데이터셋의 행 수가 같습니다. 입력 데이터셋의 행 수가 전치되는 변수의 수와 같지 않으면 출력 데이터셋은 결측값으로 채워집니다.
새로운 칼럼 이름	이 역할에 입력 칼럼을 할당하면 해당 칼럼 값이 출력 데이터셋의 데이터 칼럼 머리글 접미어가 됩니다. 이 역할에 입력 칼럼을 할당하지 않으면 출력 데이터셋의 데이터 칼럼 머리글 접미어로 정수가 사용됩니다. 데이터 칼럼 머리글의 접두어를 지정하려면 선택 영역에서 옵션을 클릭하고 접두어 사용 체크 박스를 선택합니다. 그런 다음 칼럼 이름 접두어 상자에 접두어를 입력합니다. 일반적으로 출력 데이터셋에 대한 칼럼 이름을 제공하는 데 사용되는 칼럼은 중복된 값을 포함하면 안 됩니다. 중복된 칼럼 이름이 발생하기 때문입니다. 중복 값이 포함된 칼럼을 사용하려면 해당 칼럼을 지정하고 선택한 다음 오른쪽에 나타나는 중복 허용 체크 박스를 선택합니다. 이때 주어진 값이 포함된 마지막 행만 전치됩니다. 중복 값을 포함하는 칼럼을 사용할 때 중복 허용 체크 박스를 선택하지 않으면 전치가 되지 않습니다.
그룹 분석변수	이 역할에 지정된 각 변수는 전치될 칼럼을 개별적으로 전치될 하위칼럼으로 분할하는 데 사용됩니다. 그룹 변수의 값으로 정의된 각 하위칼럼은 출력 데이터셋의 행이 됩니다.
새로운 칼럼 레이블	이 역할은 새로운 칼럼 이름 역할에 입력 칼럼을 할당할 때 사용할 수 있습니다.

역할 이름	설명
	이 역할에 입력 칼럼을 할당하면 해당 칼럼의 값이 출력 데이터셋의 데이터 변수 레이블로 사용됩니다. 이 역할에 입력 칼럼을 할당하지 않으면 출력 데이터셋의 데이터 변수 레이블이 변수의 이름이 됩니다.

전치: 분석 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **옵션**을 클릭합니다.

출력 데이터셋의 각 행에는 입력 데이터셋의 변수 이름 및 레이블이 포함됩니다. 이 변수 이름 또는 변수 레이블이 포함된 출력 칼럼의 머리글을 지정할 수 있습니다. 레이블에는 특수 문자, 선두 번호 및 공백을 포함할 수 있지만 32자를 넘을 수 없습니다. 변수 이름 칼럼에 대한 기본 머리글은 **소스**입니다. 변수 레이블 칼럼에 대한 기본 머리글은 **레이블**입니다.

출력 데이터셋의 모든 데이터 칼럼 머리글은 동일한 접두어로 시작됩니다. 접두어에는 특수 문자, 선두 번호 및 공백을 포함할 수 있지만 32자를 넘을 수 없습니다. 기본 접두어는 **Column**입니다.

주: **새로운 칼럼 이름** 역할에 변수를 할당하는 경우 접두어 지정은 선택 사항입니다. 이 역할에 변수를 할당하지 않으면 접두어가 필요하며, **접두어 사용** 체크 박스를 사용할 수 없습니다.

전치: 결과 옵션 설정

이 옵션에 액세스하려면 선택 영역에서 **결과**를 클릭합니다.

출력 데이터셋에 다른 이름을 지정할 수 있습니다.

주: SAS Enterprise Guide는 **옵션** 창의 **작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

단위당 결점수관리도

단위당 결점수관리도 정보.....	575
단위당 결점수관리도: 분석 역할에 변수 할당.....	576
관리도: 관리한계 지정.....	577
관리도: 수행할 검정 선택.....	578
관리도: 도표 모양 옵션 설정.....	579
관리도: 분석 결과 저장.....	579
제목 및 각주 지정.....	580
속성 보기.....	581

단위당 결점수관리도 정보

Shewhart 관리도는 공정이 통계적 관리 상태에 있는지 여부를 결정하기 위한 그래픽 및 분석 도구입니다. 단위당 결점수관리도 작업은 임의의 단위수를 포함하는 하위그룹 표본에서 검사 단위당 불량(결함)수에 대한 단위당 결점수관리도를 생성합니다.

이 작업을 통해 제조 공정에서 발생한 검사 단위당 결점수를 확인할 수 있습니다. 예를 들어 한 섬유 회사에서 단위당 결점수관리도를 사용하여 단위 직물당 결점수를 모니터링한다고 가정합니다. 직물은 두루마리에 감겨 있고 각 직물 조각은 가로 1미터, 세로 30미터 길이입니다. 각 직물 조각은 너비가 1미터이고 길이가 30미터입니다. 20개의 두루마리에 대한 결점수를 SAS 데이터셋에 저장하여 단위당 결점수관리도 생성에 사용합니다. 단위당 결점수관리도에서 각 포인트는 특정 하위그룹의 검사 단위당 결함 수를 나타냅니다. 예를 들어 첫 번째 두루마리에 12개의 결함이 있고 두루마리에는 30 평방 미터가 포함되므로 첫 번째 하위그룹에 대해 도표화된 값은 $12/30=0.4$ 입니다. 어떤 점이 관리한계(기본적으로 3 sigma)를 초과하지 않으면 단위당 결점수관리도는 직물 제조 공정이 통계적 관리 상태임을 나타냅니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	SHEWHART
필수 SAS 제품	Base SAS, SAS/QC, SAS/GRAPH
추가 SAS 제품(권장)	SAS/STAT

단위당 결점수관리도: 분석 역할에 변수 할당

이 옵션에 액세스하려면 선택 영역에서 **데이터**를 클릭합니다.

기본적으로 작업을 열기 전에 선택한 데이터 소스가 작업의 입력 데이터 소스입니다. 입력 데이터 소스의 이름은 **데이터 소스** 필드에 표시됩니다. 데이터 소스에 지정된 필터는 **필터** 필드에 표시됩니다. 입력 데이터 소스를 변경하거나 필터링하려면 **편집**을 클릭합니다.

데이터 소스를 선택한 후에는 역할에 변수를 할당할 수 있습니다.

역할 이름	설명
결점수	결점수를 갖는 변수를 지정합니다. 이 역할에는 변수를 하나 이상 할당해야 합니다.
하위그룹 식별자	데이터의 하위그룹을 식별하는 변수를 지정합니다. 변수값은 입력 테이블의 관측값이 논리적 하위그룹으로 배열되는 방법을 나타냅니다. 일반적으로 값은 다음과 같습니다. ■ 하위그룹 표본의 수집 순서를 지정하는 지수 ■ 하위그룹 표본이 수집된 날짜 또는 시간 ■ 하위그룹 표본을 고유하게 식별하는 레이블 하위그룹 식별자는 필수 변수이며 이 역할에는 변수를 하나만 할당할 수 있습니다. 식별된 하위그룹으로 데이터를 정렬하려면 하위그룹별 정렬 체크 박스를 선택합니다.
하위그룹 표본 크기	하위그룹 표본 크기를 변수값으로 지정합니다. 하위그룹 표본 크기는 반드시 입력해야 하며 이 역할에는 하나의 숫자 변수만 할당할 수 있습니다. 하위그룹 표본 크기가 고정되어 있고 크기에 상수값을 지정해야 할 때 하위그룹 표본 크기로 (변수 대신) 숫자값 사용 체크 박스를 선택합니다. 하위그룹 표본 크기 상자에 사용할 값을 입력합니다. 하위 그룹 표본 크기는 프로세스 변수의 최댓값과 99999(포함) 사이의 값이어야 합니다.

역할 이름	설명
	주: 옵션 영역에서 하위그룹 표본 크기로 (변수 대신) 숫자값 사용 체크 박스를 선택하면 이 역할은 필요하지 않습니다.
그룹 분석변수	그룹 변수에 의해 정의된 그룹의 행에 대해 개별 분석을 수행합니다. 기본적으로 데이터는 이러한 그룹 변수값의 오름차순으로 정렬됩니다. 주: 입력 데이터셋이 정렬되지 않게 하려면 변수로 정렬 체크 박스의 선택을 취소합니다. 주: 분석 역할에 이미 할당한 변수로는 분석을 그룹화할 수 없습니다.
블록 변수	데이터를 연속된 하위그룹 블록으로 그룹화하는 변수를 최대 두 개까지 지정합니다. 이 블록은 범례에서 레이블로 표시되며 각 블록 변수는 범례에서 한 종류의 레이블이 됩니다.

관리도: 관리한계 지정

이 옵션에 액세스하려면 선택 영역에서 **관리한계**를 클릭합니다.

옵션 이름	설명
Sigma 한계	관리도에 그려지는 요약 하위그룹 통계량의 표준오차(sigma)의 배수로 관리한계의 너비를 지정합니다. 너비는 양수여야 합니다. 기본 배수는 3입니다.
계산 방법 선택	관리한계의 계산 방법을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 활성 데이터에서 관리한계 계산 ■ 선택한 데이터에서 관리한계 계산 관리한계 또는 관리한계를 계산할 수 있는 모수를 포함하는 데이터셋을 지정하려면 찾아보기를 클릭합니다. 사용할 데이터셋을 선택하고 열기를 클릭합니다. 주: 선택한 데이터셋은 LIMITS= 데이터셋이어야 합니다. LIMITS= 데이터셋에 필요한 변수에 대한 자세한 내용은 생성 중인 그래프의 프로시저 문서를 참조하십시오. 모든 하위 그룹에 대한 요약통계량을 도표화하려면 하위그룹 표본 크기가 명목 관리한계 표본 크기와 동일한지 관계없이 모든 하위그룹에 대한 요약통계량 도표를 생성합니다. 체크 박스를 선택합니다. 거의 모든 하위 그룹의 표본 크기가 공통이고 Nominal 표본 크기 n에 대응하는(가변이 아닌) 고정된 관리한계를 표시하려고 할 때 이 통계량을 도표화할

옵션 이름	설명
	수 있습니다. 이 옵션을 선택하고 하위 그룹에 대한 표본 크기가 유의하게 다르면 관리한계 해석은 표본 크기가 n인 하위 그룹에 대해서만 의미가 있습니다. 표본 크기가 n과 같지 않은 데이터 포인트에 대해 특수 표식을 추가하려면 표본 크기에 해당하는 포인트에 대해 특수 표식을 추가하며, 표본 크기는 고정 관리한계의 명목표본 크기와 동일하지 않습니다. 체크 박스를 선택합니다. ■ 관리한계 지정 상한, 중심 및 하한 필드에 관리상한, 중심선 및 관리하한을 각각 지정합니다. 하위그룹의 표본 크기가 다양하면, 명목표본 크기를 지정해야 합니다. 주: 그룹 분석변수 역할에 변수를 할당하면 관리한계 값을 입력할 수 없습니다.

관리도: 수행할 검정 선택

이 옵션에 액세스하려면 선택 영역에서 **검정**을 클릭합니다.

주: 이 검정은 3 Sigma 한계가 아니거나 하위그룹 표본 크기에 따라 달라지는 관리한계에는 적용되지 않습니다.

옵션 이름	설명
검정 선택	체크 박스를 사용하여 런 검정, 패턴 검정 및 Western Electric 규칙이라고도 하는 특수 원인에 대한 검정을 하나 이상 요청합니다. 이러한 검정은 주관리도에 그려진 포인트에서 특정 비임의 패턴을 찾아냅니다. 이러한 비임의 패턴의 발생을 신호라고 합니다. 이러한 검정을 통해 특수한 변동의 원인이 존재하는지를 알 수 있습니다. 각 검정을 선택하면 해당 검정에 대한 설명이 설명 상자에 나타납니다.
레이블	텍스트 상자를 사용하여 특수 원인에 대한 검정이 양성으로 나타난 포인트에 대한 레이블을 입력합니다. 레이블의 문자 수는 16자를 초과할 수 없습니다. 각 레이블은 검정에서 신호가 발생한 모든 점에 표시됩니다.
레이블 색상	■ 드롭다운 리스트를 사용하여 지정한 레이블의 색상을 선택합니다.

옵션 이름	설명
	■ 영역 선 표시 체크 박스를 사용하여 주관리도에 A, B, C 영역을 구분하는 선을 표시합니다. 이 영역은 관리한계 사이의 간격을 여섯 개의 동일한 간격을 지닌 하위 간격으로 분할함으로써 구성됩니다. ■ 3 sigma 한계 무시 체크 박스를 사용하여 기본 배수 3 이외의 sigma 관리한계를 지정했을 때 특수한 원인을 찾는 검정을 수행할 수 있습니다. sigma 한계를 변경하려면 선택 영역에서 관리한계를 클릭합니다. ■ 모든 점을 검정의 시작점으로 사용 체크 박스를 사용하여 중복되는 점 패턴에 특수 원인 검정을 적용합니다.

관리도: 도표 모양 옵션 설정

일부 모양 옵션은 여러 관리도 작업에 공통적으로 적용됩니다. 이러한 옵션에 대한 자세한 내용은 아래 항목을 참조하십시오.

- [“축 옵션 설정” \(605페이지\)](#)
- [“참조선 추가” \(605페이지\)](#)
- [“기타 도표 옵션 설정” \(606페이지\)](#)
- [“블록 변수 옵션 설정” \(608페이지\)](#)

관리도: 분석 결과 저장

이 옵션에 액세스하려면 선택 영역에서 **테이블**을 클릭합니다.

분석 결과를 저장하기 위해 다양한 통계량을 포함하는 출력 데이터셋을 생성할 수 있습니다. 영구적으로 저장하려는 출력 데이터의 체크 박스를 선택합니다.

역할 이름	설명
하위그룹 통계량과 관리한계 데이터셋	관리도에 그려지는 정보를 포함하는 출력 데이터셋을 생성합니다. 여기에는 하위그룹 식별자값 및 해당 요약통계량과 관리한계가 포함됩니다. 주: 이 옵션은 OUTTABLE= 데이터셋을 생성합니다. SHEWHART 프로시저에서 TABLE= 옵션을 사용하여 이 데이터셋

역할 이름	설명
	을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.
하위그룹 통계량 출력 데이터셋	하위그룹 요약통계량을 포함하는 출력 데이터셋을 생성합니다. 주: 이 옵션은 OUTHISTORY= 데이터셋을 생성합니다. SHEWHART 프로시저에서 HISTORY= 옵션을 사용하여 이 데이터셋을 다시 SAS로 읽어올 수 있습니다. 자세한 내용은 SHEWHART 프로시저의 도움말을 참조하십시오.
관리한계 출력 데이터셋	관리한계를 저장하는 출력 데이터셋을 생성합니다. 추후에 이 작업 사용 시 관리한계의 계산 방법을 선택할 때 이 데이터셋을 사용할 수 있습니다.

주: SAS Enterprise Guide는 **옵션 창의 작업> 출력 라이브러리** 영역에 정의된 라이브러리 리스트를 검색하고 해당 리스트에서 첫 번째 쓰기 가능한 라이브러리의 출력 데이터를 저장합니다.

SAS Add-In for Microsoft Office는 Work 라이브러리에 출력 데이터를 저장합니다. 출력 데이터의 저장 위치를 달리 지정하려면 **찾아보기**를 클릭합니다.

제목 및 각주 지정

이 옵션에 액세스하려면 선택 영역에서 **제목**을 클릭합니다.

결과에 대해 기본 제목과 주석 텍스트를 사용하면 별도의 작업을 수행할 필요가 없습니다.

제목 및 주석을 지정하는 방법:

- 1 **섹션** 상자에서 제목 또는 주석을 선택합니다.
- 2 **기본 텍스트 사용** 체크 박스의 선택을 취소합니다.
- 3 텍스트 상자에 표시된 제목 또는 주석 텍스트를 편집합니다.

주: 제목과 각주에는 매크로 변수를 사용할 수 있습니다.

속성 보기

이 옵션에 액세스하려면 선택 영역에서 **속성**을 클릭합니다.

속성 패널에서 작업 속성에 대한 요약을 볼 수 있습니다. 일부 작업에서는 편집을 클릭하여 이러한 속성을 **편집**할 수 있습니다. **속성** 대화 상자가 나타납니다.

속성 패널에서 다음과 같은 정보를 볼 수 있습니다.

- 작업 레이블
- 작업을 실행하는 데 사용된 서버
- 작업이 생성된 날짜 및 시간
- 작업이 마지막으로 수정된 날짜 및 시간
- 작업이 마지막으로 실행된 시간
- 데이터 제한
- 사용된 프롬프트
- 결과 출력형식
- 결과의 그래픽 출력에 대한 출력형식

주: 기존 작업의 레이블을 변경하는 경우 **속성** 대화 상자에서 **확인**을 클릭하면 새로운 레이블이 저장됩니다. 그러면 작업을 취소해도 새로운 레이블이 사용됩니다.

라이브러리 메타데이터 업데이트

라이브러리 메타데이터 업데이트 마법사 정보	583
라이브러리 메타데이터 업데이트: SAS 서버 및 라이브러리 선택	584
라이브러리 메타데이터 업데이트: 메타데이터에서 리포트 실행 또는 메타데이터 업데이트	584

라이브러리 메타데이터 업데이트 마법사 정보

라이브러리 메타데이터 업데이트 마법사에서는 메타데이터 레파지토리에 있는 테이블 정의에 대한 리포트를 실행하거나 이러한 테이블 정의를 업데이트할 수 있습니다. 메타데이터 레파지토리의 테이블 정의를 기준으로 라이브러리 콘텐츠를 나열하는 리포트나 물리적 라이브러리의 테이블과 메타데이터 레파지토리의 테이블 정의 사이의 차이점을 나열하는 리포트를 생성할 수 있습니다. 리포트를 실행하는 것 외에 **라이브러리 메타데이터 업데이트** 마법사를 사용하여 메타데이터 레파지토리의 테이블 정의를 업데이트, 추가 또는 삭제할 수도 있습니다.

기본적으로 **SAS Enterprise Guide**에는 메타데이터에 등록된 테이블만 표시됩니다. **라이브러리 메타데이터 업데이트** 마법사는 METALIB 프로시저를 사용하여 이러한 테이블을 메타데이터 레파지토리에 프로그래밍 방식으로 등록하고 이러한 테이블에 대한 메타데이터를 동기화합니다.

주: 이 작업은 **SAS Enterprise Guide**에서만 사용 가능합니다.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	METALIB
필수 SAS 제품	Base SAS

요구 사항 이름	프로시저 및 제품 이름
추가 SAS 제품(권장)	없음

라이브러리 메타데이터 업데이트: SAS 서버 및 라이브러리 선택

- 1 사용 가능한 서버 리스트에서 서버를 선택합니다.
- 2 라이브러리를 선택합니다.
- 3 다음을 클릭하여 선택한 라이브러리에 대해 리포트를 실행하거나 메타데이터를 업데이트합니다.

라이브러리 메타데이터 업데이트: 메타데이터에서 리포트 실행 또는 메타데이터 업데이트

- 1 리포트를 실행하거나 메타데이터를 업데이트하려면 먼저 SAS 서버 및 라이브러리를 선택해야 합니다.
- 2 다음 옵션 중 하나를 선택합니다.

옵션 이름	설명
메타데이터 레파지토리에 등록된 라이브러리 콘텐츠에 대한 리포트 생성	메타데이터 레파지토리의 테이블 정의를 기준으로 라이브러리 콘텐츠에 대한 리포트를 생성합니다.
물리적 테이블과 메타데이터 레파지토리 간의 차이에 대한 리포트 생성	다음 정보가 포함된 리포트를 생성합니다. ■ 해당 물리적 테이블과 일치하도록 업데이트되는 메타데이터 레파지토리의 테이블 정의 ■ 메타데이터 레파지토리에 테이블 정의가 없는 물리적 라이브러리의 테이블

옵션 이름	설명
	■ 물리적 라이브러리에 더 이상 없지만 메타데이터 레파지토리에 테이블 정의가 있는 테이블
메타데이터에 있는 테이블 정의 추가 및 삭제 테이블 및 칼럼으로 업데이트	다음 단계를 수행합니다. ■ 메타데이터가 있는 물리적 라이브러리의 테이블이면 메타데이터 레파지토리의 테이블 정의가 해당하는 물리적 테이블과 일치하도록 업데이트됩니다. 이러한 업데이트에는 테이블 칼럼, 인덱스, 고유 키, 외래 키 및 키 연관에 대한 정보가 포함됩니다. ■ 메타데이터가 없는 물리적 라이브러리의 테이블이면 해당 테이블 정의가 메타데이터 레파지토리에 추가됩니다.
메타데이터에 있는 기존 테이블 정의만 현재 칼럼 정보로 업데이트	메타데이터가 있는 물리적 라이브러리의 테이블에 대한 메타데이터 레파지토리의 테이블 정의를 업데이트합니다.
메타데이터 라이브러리에서 오래된 엔트리 (더 이상 존재하지 않는 테이블) 삭제	물리적 라이브러리에 더 이상 존재하지 않는 테이블에 대한 테이블 정의를 메타데이터 레파지토리에서 삭제합니다.

- 3 사용자 ID와 암호가 필요하지 않으면 이러한 옵션을 사용할 수 없습니다. 사용자 ID 및 암호가 필요하지 않으면 이 옵션을 사용할 수 없습니다.
- 4 **마침**을 클릭하여 리포트를 실행하거나 라이브러리의 메타데이터를 업데이트합니다.

CAS로 업로드

CAS로 업로드 작업 정보	587
SAS CAS Analytic Server로 데이터 업로드	588

CAS로 업로드 작업 정보

CAS로 업로드 작업으로 SAS 9.4 데이터를 CAS(Cloud Analytics Services) 서버에 로드할 수 있습니다. 이 작업은 SAS/CONNECT를 사용하여 데이터를 원격 서버로 이동합니다. SAS 9.4M5부터 여러 가지 효율적인 방법으로 데이터를 CAS 서버로 이동합니다. 예를 들어 CAS LIBNAME 엔진 또는 PROC CASUTIL을 사용할 수 있습니다. 자세한 내용은 [SAS Cloud Analytic Services: User's Guide](#)를 참조하십시오.

주: **SAS Enterprise Guide** 또는 **SAS Add-In for Microsoft Office**에서 이 작업을 실행하려면 먼저 환경을 구성해야 합니다. 자세한 내용은 [SAS Intelligence Platform: Desktop Application Administration Guide](#)를 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	없음
필수 SAS 제품	Base SAS, SAS Viya
추가 SAS 제품(권장)	없음

SAS CAS Analytic Server로 데이터 업로드

- 1 CAS 서버를 선택합니다.
- 2 데이터를 저장할 라이브러리를 선택합니다.
- 3 테이블의 멤버 이름을 지정합니다. 기본적으로 멤버 이름은 데이터셋 이름과 동일합니다. 분산 로딩을 사용하면 멤버 이름을 변경할 수 없습니다.
- 4 선택한 테이블에 대한 자격 증명을 찾을 위치를 지정합니다. SAS Metadata Server에 연결되어 있으면, 메타데이터에 저장된 자격 증명을 사용할 수 있습니다. CAS 서버로 로컬 데이터를 업로드할 때 CAS 서버의 사용자 이름과 암호를 입력해야 합니다.

데이터 파일을 서버에 업로드

데이터 파일을 서버로 업로드 작업 정보	589
데이터 파일을 서버에 업로드: 업로드할 SAS 데이터셋 선택	589
데이터 파일을 서버에 업로드: 업로드 옵션 지정	590

데이터 파일을 서버로 업로드 작업 정보

데이터 파일을 서버로 업로드 작업은 사용자가 로컬 Microsoft Windows 파일 시스템에서 하나 이상의 SAS 데이터 파일(*.sas7bdat)을 선택하여 SAS 서버로 복사할 수 있도록 합니다.

이 작업은 복사된 데이터 파일, 바이트 크기, 각 파일 및 그룹인 모든 파일에 대한 시간, 전송 시 발생한 오류 등이 요약된 로그를 생성합니다. 기본적으로 모든 데이터 파일은 SAS Enterprise Guide 프로젝트의 출력 데이터로 추가되므로 프로젝트에서 사용할 수 있습니다.

데이터 파일을 서버에 업로드: 업로드할 SAS 데이터셋 선택

- 1 **추가**를 클릭하여 업로드할 데이터 파일을 선택합니다.
- 2 복사할 파일의 이름을 선택합니다.

.....
주: SAS7BDAT 파일을 선택해야 합니다.
.....

- 3 **열기**를 클릭하여 복사할 데이터 파일을 선택합니다.

주: 타겟 서버에서 복사되는 모든 파일에는 소문자 이름이 지정됩니다.

- 4 다음을 클릭하여 업로드 옵션을 지정합니다.

데이터 파일을 서버에 업로드: 업로드 옵션 지정

- 1 업로드 옵션을 지정하기 전에 먼저 업로드할 데이터 파일을 선택해야 합니다.
- 2 **SAS 서버 선택** 드롭다운 리스트에서 선택한 데이터 파일을 복사해 놓을 서버를 선택합니다.
- 3 **대상 라이브러리 선택** 드롭다운 리스트에서 사용할 SAS 라이브러리를 선택합니다.
- 4 복사할 파일과 이름이 같은 SAS 서버의 기존 파일을 바꾸려면 **동일한 이름의 기존 파일 바꾸기** 체크 박스를 선택합니다. 기본적으로는 기존 파일이 바뀌지 않습니다.
- 5 복사 중인 데이터 파일을 현재 **SAS Enterprise Guide** 프로젝트에 추가하지 않으려면 **현재 프로젝트에 데이터 파일 추가** 체크 박스를 선택 취소합니다. 기본적으로 복사된 데이터 파일은 현재 프로젝트에 추가됩니다.
- 6 **마침**을 클릭하여 데이터 파일을 복사합니다.

주: 대상 라이브러리가 단일 파일 경로의 BASE 엔진 라이브러리이면 파일이 해당 라이브러리 경로로 직접 전송됩니다. 타겟 라이브러리가 Base SAS 엔진 라이브러리가 아니거나 연결된 라이브러리인 경우 파일 전송은 먼저 Work의 Staging 영역으로 이동합니다. 그런 다음 PROC SQL을 사용하여 타겟 라이브러리에 복사됩니다.

LASR로 업로드

LASR로 업로드 작업 정보	591
SAS LASR Analytic Server로 데이터 업로드	592

LASR로 업로드 작업 정보

SAS LASR Analytic Server는 분석 계산을 가속화하기 위해 특별히 고안된 In-Memory 분석 플랫폼입니다. 서버는 대용량 데이터가 없는 배포에 대한 비분산 구성에서 사용할 수 있습니다. 분산 서버는 분산 계산 환경 및 대용량 데이터셋 분석에 대한 병렬 프로세스를 사용합니다. LASR로 업로드 작업은 **SAS Visual Analytics**에서 데이터에 액세스할 수 있는 LASR Analytic Server에 데이터를 업로드할 수 있습니다.

주: 이 작업을 실행하려면 **SAS Enterprise Guide** 또는 **SAS Add-In for Microsoft Office**가 **SAS Visual Analytics**로 동일한 메타데이터 서버에 연결되어야 합니다.

SAS LASR Analytic Server의 연결 설정 방법에 대한 자세한 내용은 [SAS Intelligence Platform: Desktop Application Administration Guide](#)를 참조하십시오.

이 작업을 실행하려면 다음 요구 사항을 충족해야 합니다.

요구 사항 이름	프로시저 및 제품 이름
사용되는 SAS 프로시저	없음
필수 SAS 제품	Base SAS, SAS LASR Analytic Server
추가 SAS 제품(권장)	SAS Visual Analytics

SAS LASR Analytic Server로 데이터 업로드

- 1 데이터를 저장할 라이브러리를 선택합니다.
- 2 테이블의 멤버 이름을 지정합니다. 기본적으로 멤버 이름은 데이터셋 이름과 동일합니다. 분산 로딩을 사용하면 멤버 이름을 변경할 수 없습니다.
- 3 업로드한 데이터에 대한 디렉터리 경로를 지정합니다. 기본적으로 각 라이브러리는 사전 정의된 위치를 가지고 있습니다.

부록 1

그래프 작업에 대한 공통 모양 옵션

축 옵션 설정	593
참고 항목.....	595
참조선 추가	595
참고 항목.....	596
범례 옵션 설정	596
참고 항목.....	597
그래프 영역 사용자 정의	597
참고 항목.....	598
막대 모양 지정	598
참고 항목.....	599
막대 옵션 설정	599
참고 항목.....	600
선 도표 옵션 설정	600
참고 항목.....	600
보간법 선택	601
참고 항목.....	601
고급 옵션 설정	602
참고 항목.....	603

축 옵션 설정

선택 영역에서 이 옵션은 모양 > 축 머리글 아래에 표시됩니다.

주: GIF를 그래픽 출력형식으로 사용하여 표면을 생성하면 축 옵션이 적용되지 않습니다. 표면에 축 옵션을 적용하려면 ActiveX를 그래픽 출력형식으로 사용합니다.

역할 이름	설명
축과 눈금 숨기기	축과 눈금 표시 숨기기
축	가로축 또는 세로축을 사용자 정의합니다. ■ 일부 도표에서는 오른쪽 세로축에 대한 옵션을 설정할 수도 있습니다. ■ 3D 산점도, 3D 바늘 도표 및 3D 표면도에서는 깊이축에 대해 이 옵션을 지정할 수 있습니다. ■ 그룹 도표 또는 다중 축도 막대 그래프에서는 그룹축에 대해 이 옵션을 지정할 수 있습니다. 축 모양을 사용자 정의하는 방법 ■ 가로축, 세로축 및 깊이축의 선에 대해 색상, 너비 및 스타일을 지정합니다. ■ 수평축 또는 수직축의 값 순서를 바꾸려면 축 바꾸기 옵션을 선택합니다. 수평축의 체크 박스를 선택하면 X값의 순서가 바뀝니다. 수직축의 체크 박스를 선택하면 Y값의 순서가 바뀝니다. 주: 이 옵션은 영역 도표, 상자 도표, 버블 도표, 등고선 도표, 선 도표 및 2D 산점도에서만 사용할 수 있습니다. ■ 축 레이블을 사용자 정의하려면 레이블 탭을 클릭합니다. 레이블 상자에 사용자 정의 레이블의 텍스트를 입력합니다. 레이블 회전 드롭다운 리스트를 사용하여 축에 대한 레이블 표시 방법을 설정합니다. 또한 글꼴 유형, 크기 및 색상도 지정할 수 있습니다. 텍스트에 밑줄, 굵게 또는 기울임을 설정할 수 있습니다. 양쪽 맞춤을 변경하려면 이 버튼을 사용합니다. 주: 다중 축도 막대 그래프를 생성하는 중이면 반응축의 기본 레이블은 각 막대의 길이를 계산하는 데 사용되는 통계량입니다. ■ 축 값을 사용자 정의하려면 값 탭을 클릭합니다. 값의 글꼴에 대한 유형, 크기 및 색상을 지정할 수 있습니다. 텍스트에 밑줄, 굵게 또는 기울임을 설정할 수 있습니다. 양쪽 맞춤을 변경하려면 이 버튼을 사용합니다. 축에 대한 값 표시 방법을 지정하려면 값 회전 드롭다운 리스트를 사용합니다.
주 눈금	가로축 또는 세로축에 대한 주 눈금 표시를 지정합니다. 일부 도표에서는 오른쪽 세로축에 대한 옵션을 설정할 수도 있습니다. 3D 산점도, 3D 바늘 도표 및 3D 표면도에서는 깊이축에 대한 눈금 표시를 생성할 수 있습니다. 막대 그래프에서는 반응축에만 눈금 표시를 지정할 수 있습니다. 예를 들어 세로축 도표를 만들 때 가로축이 아닌 세로축에 눈금 표시를 지정할 수 있습니다. 축 이름 주 눈금 표시를 선택하고 사용할 눈금 개수를 결정하는 방법을 지정합니다. ■ 주 눈금 표시의 수와 빈도를 자동으로 결정하려면 자동을 선택합니다. 영역 도표, 버블 도표, 선 도표 또는 2D 산점도를 생성 중이라면 0에서 시작 체크 박스를 선택하여 눈금 표시값이 0에서 시작하도록 합니다.

역할 이름	설명
	■ 사용할 눈금 표식 수를 지정하려면 사용을 선택합니다. 원하는 수를 텍스트 상자에 입력합니다. ■ 로그 눈금 패턴을 지정하려면 로그를 선택합니다. 밑과 먹을 선택하고 밑 또는 먹의 값으로 눈금 표식을 배열할 것인지를 선택합니다. ■ 눈금 표식에 대한 특정 위치를 지정하려면 지정을 선택합니다. 텍스트 상자에 눈금 표식 값을 입력하고 추가를 클릭합니다. 값을 변경하려면 리스트에서 해당 값을 선택한 후 편집을 클릭합니다. 변경한 후 저장을 클릭합니다.
보조 눈금	가로축 또는 세로축에 대한 보조 눈금 표식을 지정합니다. ■ 일부 도표에서는 오른쪽 세로축에 대한 옵션을 설정할 수도 있습니다. ■ 3D 산점도, 3D 바늘 도표 및 3D 표면도에서는 깊이축에 대한 눈금 표식을 생성할 수 있습니다. ■ 막대 그래프에서는 반응축에만 눈금 표식을 지정할 수 있습니다. 예를 들어 세로축 도표를 만들 때 가로축이 아닌 세로축에 눈금 표식을 지정할 수 있습니다.

참고 항목

- 3장, “영역 도표” (11페이지)
- 6장, “막대 그래프” (31페이지)
- 8장, “막대-선 그래프” (41페이지)
- 11장, “상자 도표” (59페이지)
- 12장, “버블 도표” (65페이지)
- 19장, “등고선 도표” (107페이지)
- 42장, “선 도표” (245페이지)
- 77장, “산점도” (493페이지)
- 87장, “표면도” (543페이지)

참조선 추가

선택 영역에서 이 옵션은 모양 > 참조 머리글 아래에 표시됩니다.

주: 이 옵션은 3D 산점도에서는 사용할 수 없습니다.

역할 이름	설명
격자선 표시	격자선을 표시합니다. 주: 이 옵션은 막대 그래프에서는 사용할 수 없습니다.
참조선	가로축 또는 세로축에 참조선을 추가합니다. 일부 도표에서는 오른쪽 세로축에 대한 참조선을 설정할 수도 있습니다. 주: 막대 그래프에서는 참조선이 반응축에 대해 그려집니다. 수직 막대 그래프에서는 참조선이 가로선으로 표시되고 수평 막대 그래프에서는 참조선이 세로선으로 표시됩니다. 3D 그래프에서는 참조선을 사용할 수 없습니다.

참고 항목

- 3장, “영역 도표” (11페이지)
- 6장, “막대 그래프” (31페이지)
- 8장, “막대-선 그래프” (41페이지)
- 11장, “상자 도표” (59페이지)
- 12장, “버블 도표” (65페이지)
- 19장, “등고선 도표” (107페이지)
- 42장, “선 도표” (245페이지)
- 77장, “산점도” (493페이지)

범례 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **범례**를 클릭합니다.

다중 영역 도표, 다중 선 도표, 막대-선 그래프, 등고선 도표, 지도, 누적형 막대 그래프, 원 그래프 및 도넛 그래프에 대한 범례를 생성할 수 있습니다. **누적** 역할에 변수를 할당하면 다중 축도 막대 그래프에 대한 범례를 사용할 수 있습니다.

막대-선 그래프의 경우 두 개의 범례를 생성할 수 있습니다. 하나는 막대 그래프용이고 다른 하나는 선 도표용입니다.

산점도를 생성할 때 **세로** 및 **오른쪽 세로** 역할에 변수를 할당할 때도 범례를 사용할 수 있습니다. 같은 그래프에 여러 산점도가 있으므로 범례를 사용하여 **세로** 변수의 데이터 포인트와 **오른쪽 세로** 변수의 데이터 포인트를 구분합니다.

참고 항목

- 3장, “영역 도표” (11페이지)
- 6장, “막대 그래프” (31페이지)
- 8장, “막대-선 그래프” (41페이지)
- 19장, “등고선 도표” (107페이지)
- 30장, “도넛 그래프” (165페이지)
- 42장, “선 도표” (245페이지)
- 61장, “원 그래프” (383페이지)

그래프 영역 사용자 정의

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **그래프 영역**을 클릭합니다.

역할 이름	설명
사용자 정의 그래프 크기 지정	이미지 크기를 지정합니다. 참조를 위해 현재 화면 크기가 표시됩니다.
그래프 배경 색상	배경 색상을 지정합니다. 주: 지도 그래프를 생성하면 지도 배경 색상 드롭다운 리스트에서 지도의 배경 색상을 선택할 수 있습니다.
도표 주위에 프레임 그리기	그래프 영역 주위에 프레임을 포함합니다. 영역 배경 색상 드롭다운 리스트에서 도표 영역의 배경 색상을 선택합니다. 주: 이 옵션은 도넛 그래프, 지도 그래프, 원 그래프, 3D 그룹형 막대 그래프, 3D 산점도 또는 표면도에 대해 사용할 수 없습니다.
그래프 팁 표시	그래프 팁을 표시할지를 지정합니다. 그래프 팁은 전체 그래프에 적용되며 막대나 선과 같은 그래프의 개별 성분에는 적용되지 않습니다. ActiveX 또는 Java를 그래프 출력형식으로 선택하거나 지도 그래프를 사용하면 그래프 팁이 표시되지 않습니다. 주: 지도 그래프에서는 그래프 팁이 표시되지 않거나 그래프 출력 출력형식으로 ActiveX 또는 Java를 선택한 경우에는 그래프 팁이 표시되지 않습니다.

참고 항목

- 3장, “영역 도표” (11페이지)
- 6장, “막대 그래프” (31페이지)
- 8장, “막대-선 그래프” (41페이지)
- 11장, “상자 도표” (59페이지)
- 12장, “버블 도표” (65페이지)
- 19장, “등고선 도표” (107페이지)
- 30장, “도넛 그래프” (165페이지)
- 42장, “선 도표” (245페이지)
- 61장, “원 그래프” (383페이지)
- 69장, “레이더 그래프” (447페이지)
- 77장, “산점도” (493페이지)
- 87장, “표면도” (543페이지)

막대 모양 지정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **막대**를 클릭합니다.

역할 이름	설명
막대 색상 지정	막대 색상을 지정합니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 기본 색상 구성 - 막대에 기본 색상 구성을 사용합니다. 기본 색상 구성은 현재 스타일에 따라 결정됩니다. 스타일 관리자나 결과에 대한 속성 대화 상자를 사용하여 스타일을 지정할 수 있습니다. SAS Enterprise Guide의 옵션 대화 상자에서 각 결과 유형에 대한 스타일을 지정할 수도 있습니다. SAS Add-In for Microsoft Office에서 SAS Add-In 옵션 대화 상자를 사용하여 SAS 리포트 및 HTML 출력형식에 스타일을 적용할 수 있습니다. ■ 기존 색상(이전 버전에서 지정됨) - 이전 버전의 SAS Enterprise Guide 또는 SAS Add-In for Microsoft Office에서 지정한 색상을 사용할 수 있습니다. 이 옵션은 이전 릴리스에서 마이그레이션된 SAS 콘텐츠에 대해서만 사용 가능합니다. ■ 사용자 정의 색상 - 막대에 사용할 사용자 정의 색상을 지정할 수 있습니다. 막대 그래프, 누적 그래프, 인터리브 그래프와 같이 둘 이상의 색상을 사용해야 하는 막대 그래프에 대해 사용자 정의 색상을 지정할 수 있습니다. 드롭다운 리스트를 사용하여 막대

역할 이름	설명
	그래프에 최대 12개의 색상을 지정할 수 있습니다. 그래프에 12개가 넘는 색상을 사용해야 할 때는 현재의 기본 색상 구성에서 추가 색상이 선택됩니다. 사용자 정의 색상을 기본값으로 돌려려면 재설정을 클릭합니다. 주: 막대 그래프 유형에 따라 일부 사용자 지정 색상을 사용하지 못할 수도 있습니다.
막대 개수 결정	그래프의 막대 수를 지정합니다. 기본적으로 막대 수는 자동으로 결정됩니다. 숫자 칼럼이면 값이 여러 범위로 나뉘며 각 범위마다 막대가 한 개씩 표시됩니다. 숫자 칼럼이 이산 값을 나타내면 고유 데이터 값마다 막대 하나 옵션을 사용할 수 있습니다. 주: 다중 축도 막대 그래프에 대해서는 막대 수를 지정할 수 없습니다. 그래프에 표시할 막대 수를 지정하려면 막대 개수 지정 체크 박스를 선택합니다.

참고 항목

- 6장, “막대 그래프” (31페이지)
- 8장, “막대-선 그래프” (41페이지)

막대 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **막대 옵션**을 클릭합니다.

순서 드롭다운 리스트를 사용하여 막대의 정렬 순서를 지정합니다. 기본적으로 정렬 순서는 자동으로 결정됩니다. 막대를 길이나 높이에 따라 오름차순 또는 내림차순으로 배열할 수 있습니다.

각 막대에 외곽선을 표시하려면 **외곽선 색상** 체크 박스를 선택합니다. 드롭다운 리스트에서 외곽선 색상을 선택할 수 있습니다.

막대 크기 드롭다운 리스트를 사용하여 **막대 크기**를 설정합니다. 기본적으로 막대 크기는 자동으로 결정됩니다. 그러나 텍스트 상자에 막대 너비 또는 막대 간격을 총 그래프 너비의 백분율로 지정할 수 있습니다.

참고 항목

8장, “막대-선 그래프” (41페이지)

선 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **도표**를 클릭합니다.

주: **선 합계** 역할에 하나 이상의 변수를 할당할 때만 이러한 옵션을 사용할 수 있습니다. **선 합계** 역할에 변수를 할당하지 않으면 **고급** 패널의 **선을 그릴 때 사용되는 통계량** 드롭다운 리스트를 사용하여 막대-선 그래프에서 선 도표를 생성할 때 사용할 통계량을 선택할 수 있습니다.

특정 변수에 대한 선 도표 모양을 제어하는 방법

- 1 리스트에서 변수를 선택합니다. 리스트의 변수는 **선 합계** 역할에 할당한 변수입니다.
- 2 **선** 영역에서 선 유형, 선 두께 및 선 색상을 지정합니다.
- 3 외곽선 색상을 선택합니다.
- 4 **데이터 포인트 표식** 영역에서 표식 유형, 사용할 기호, 각 표식의 높이 및 데이터 포인트 표식의 색상을 지정합니다.

주: 일부 기호는 모든 그래픽 출력형식에서 지원되지 않습니다. 결과 출력형식에서 지원되지 않는 기호를 선택하면 지원되는 다른 기호가 대신 표시됩니다.

- 5 **선을 그릴 때 사용되는 통계량** 드롭다운 리스트에서 막대-선 그래프에서 선을 그릴 때 사용할 통계량을 선택합니다. 다음 통계량 중에서 선택할 수 있습니다.
 - **평균**은 **선 합계** 역할에 할당된 변수값의 평균입니다. 이 평균은 선에서 점으로 표시된 각 고윳값에 대해 계산됩니다. 예를 들어 데이터셋의 각 도시에 대한 평균 판매를 표시하려면 판매 변수를 **선 합계** 역할에 할당하고 **평균**을 선 도표를 생성하는 데 사용할 통계량으로 선택합니다.
 - **합계**는 **선 합계** 역할에 할당된 변수값의 합계입니다. 이 합계는 선에서 점으로 표시된 각 고윳값에 대해 계산됩니다. 예를 들어 판매를 **선 합계** 역할에 할당하고 덴버의 판매 변수값이 8734, 982, 1504, 3207, 4502, 624 및 918이면 덴버의 합 통계량은 20,471입니다.

참고 항목

8장, “막대-선 그래프” (41페이지)

보간법 선택

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **보간법**을 클릭합니다.

보간법을 선택하는 방법

- 1 변수를 선택합니다. 선 도표 또는 2D 산점도를 생성 중이라면 리스트의 변수는 **세로** 및 **세로(오른쪽)** 역할에 할당된 변수입니다. **그룹 칼럼별 다중 선 도표**를 선 도표 유형으로 선택했으면 이 리스트에는 **그룹** 역할에 할당된 고유한 변수값이 포함됩니다.
- 2 드롭다운 리스트에서 보간법을 선택합니다. 선택한 보간법에 따라 추가 설정이 필요할 수도 있습니다. 다음과 같은 보간법을 사용할 수 있습니다.
 - **라그랑지** - 데이터 포인트를 지나는 매끄러운 곡선을 그리는 방법입니다. 이 방법은 데이터가 테이블식의 정확한 값으로 구성되었을 때 주로 사용됩니다. 지정된 차수 (1,3,5)의 다항식이 가장 근사한 2, 4, 6개의 포인트에 적용됩니다. 일반적으로 1차 미분계수는 연속적이지 않습니다. X 변수의 값이 반드시 증가하지만은 않을 때 해당 모수적 방법이 사용됩니다.
 - **선** - 데이터 포인트를 직선으로 연결합니다. 포인트는 입력 데이터에 나타난 순서대로 연결됩니다.
 - **바늘** - 각 데이터 포인트에서 세로축이 0 값을 갖는 수평선까지, 수평선이 0보다 크면 세로축의 최솟값까지 수직선을 그립니다. 수평선은 자동으로 표시됩니다.
 - **회귀** - 회귀분석 도표를 생성합니다. 이 보간법은 숫자 변수를 가로 및 세로 역할에 할당할 때만 사용할 수 있습니다.
 - **산점도** - 보간을 생략합니다. 데이터 포인트 도표가 생성됩니다.
 - **평활** - 스플라인 루틴을 사용하여 데이터에 평활선을 적합시킵니다. 잡음 데이터를 평활화하기 위한 방법입니다. 도표에서 점은 반드시 선에 위치하지 않아도 됩니다.
 - **스플라인** - 스플라인 루틴을 사용하여 선 도표를 그립니다.
 - **표준편차** - 평균 Y 값과 각 X에 대한 $\pm 1, 2$ 또는 3 표준편차를 실선으로 연결합니다.
 - **스텝** - 스텝함수를 사용하여 데이터를 도표로 그립니다.
- 3 (선택 사항) 동일한 보간법 및 속성 설정을 리스트의 모든 변수에 사용하려면 **모두 적용** 체크 박스를 선택합니다. 이 옵션은 리스트에 둘 이상의 변수가 있을 때만 사용할 수 있습니다.

참고 항목

- 42장, “선 도표” (245페이지)
- 77장, “산점도” (493페이지)

고급 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **고급**을 클릭합니다.

역할 이름	설명
막대 계산에 사용되는 통계량	막대 길이를 계산하는 방법을 지정합니다. 사용할 수 있는 통계량은 생성하는 그래프에 따라 다릅니다. ■ 막대 그래프 작업의 합계 칼럼 역할 또는 막대-선 그래프 작업의 막대 합계 역할에 변수를 할당하지 않으면 빈도, 누적빈도, 백분율 및 누적백분율 통계량을 사용할 수 있습니다. ■ 합계 칼럼 또는 막대 합계 역할에 변수를 할당했거나 다중 측도 그룹을 막대 그래프 유형으로 선택했으면 합계 및 평균 통계량을 사용할 수 있습니다.
선을 그릴 때 사용되는 통계량	막대-선 그래프에서 선을 그릴 때 사용할 통계량을 지정합니다. 빈도, 누적빈도, 백분율 및 누적백분율 통계량을 사용할 수 있습니다. 막대-선 그래프 작업에서 선 합계 역할에 변수를 할당하면 고급 패널에서 이 옵션을 사용할 수 없습니다. 대신 도표 패널에서 이 옵션을 사용하여 선 합계 역할에 둘 이상의 변수를 할당할 때 각 선에 대한 통계량을 관리할 수 있습니다.
오차 막대 표시	막대에 신뢰구간을 표시합니다. 유형 드롭다운 리스트에서 다음 옵션을 선택합니다. ■ 막대는 오차 막대를 기본 막대 너비의 절반으로 막대로 그립니다. ■ 둘 다는 선으로 연결된 두 개의 눈금으로 오차 막대를 그립니다. ■ 맨 위는 오차 막대를 선으로 막대의 상단에 결합한 신뢰상한에 대한 눈금으로 그립니다. 오차 막대에 색상을 지정하려면 CLM 상자를 사용하여 신뢰수준을 지정합니다. 주: 다음 조건에 맞을 때만 오차 막대를 표시할 수 있습니다. ■ 평균 또는 백분율을 막대 길이를 계산하는 데 사용되는 통계량으로 선택했을 때 ■ 3차원 그룹형 막대 그래프를 생성하지 않았을 때 ■ 누적 역할에 변수를 할당하지 않았을 때
결측값 사용	결측값을 유효한 그래프 칼럼 값으로 사용하도록 지정합니다.
0 값 제외	막대 길이가 0인 그래프 칼럼의 값을 표시하지 않습니다.

역할 이름	설명
그룹별 백분율 및 누적 백분율 계산(G100)	이 옵션을 선택하면, 개별 백분율이 그룹에 대한 중간점 기여도를 반영합니다. 이 옵션을 선택하면 개별 백분율에 그룹에 대한 중간점의 기여도가 반영됩니다. 각 그룹의 합계는 100%입니다. 주: 이 옵션은 그룹형 막대 그래프의 그룹화 칼럼 역할에 변수를 할당할 때만 사용할 수 있습니다.
막대 옆에 통계량 표시	막대 계산에 사용되는 통계량 값을 표시합니다. 주: 이 옵션은 수평 막대 그래프에서만 사용할 수 있습니다.
막대에 대해 표시할 통계 값 지정	드롭다운 리스트에서 선택한 통계량의 값을 표시합니다. 통계량은 막대 길이 계산에 사용되는 통계량과 같을 수도 있고 다를 수도 있습니다. 2차원 막대 그래프를 생성할 때에는 통계량 값을 막대의 내부에 표시할 수 있습니다.

참고 항목

- 6장, “막대 그래프” (31페이지)
- 8장, “막대-선 그래프” (41페이지)
- 19장, “등고선 도표” (107페이지)

부록 2

관리도 작업에 대한 공통 모양 옵션

축 옵션 설정	605
참조선 추가	605
인셋 추가	606
기타 도표 옵션 설정	606
블록 변수 옵션 설정	608

축 옵션 설정

이 옵션에 액세스하려면 선택 영역의 **모양** ⇨ **축** 머리글 아래에서 **축**을 클릭합니다.

다음 옵션을 설정할 수 있습니다.

- 축 색상
- 축 너비
- 눈금 표시 색상
- 가로축 및 세로축 레이블

참조선 추가

이 옵션에 액세스하려면 선택 영역의 **모양** ⇨ **축** 머리글 아래에서 **가로** 및 **세로**를 클릭합니다.

참조선을 사용하려면 **참조선 사용**을 선택합니다. 참조선의 출력형식을 지정하고 각 참조선의 위치를 지정할 수 있습니다.

인셋 추가

이 옵션에 액세스하려면 선택 영역의 **모양** 머리글 아래에서 **인셋**을 클릭합니다.

요약통계량 상자 또는 테이블(인셋)을 그래프에 직접 추가하여 도표를 보기 좋게 만들 수 있습니다. 비교 히스토그램을 생성하면 각각의 작은 히스토그램에 인셋이 표시됩니다.

인셋 포함을 선택하여 도표에 요약통계량을 추가합니다. 인셋에 포함할 통계량을 선택합니다.

주: 공정능력지수를 계산하기 위한 통계량은 **분석변수** 역할에 할당한 변수에 대한 규격한계를 생성할 때만 사용할 수 있습니다. 히스토그램, CDF 도표, 확률 도표 및 Q-Q 도표에 대한 규격한계를 생성할 수 있습니다. 공정능력지수 계산에 대한 자세한 내용은 CAPABILITY 프로시저에 대한 도움말을 참조하십시오.

기타 도표 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **옵션**을 클릭합니다.

옵션 이름	설명
관리한계 내부, 관리한계, 연결선, 한계 외부 선, 한계 외부 영역 및 프레임	이러한 그래프 요소들의 색상을 지정합니다. ■ 관리한계 내부 - 관리상한 및 관리하한 내부 색상을 지정합니다. ■ 관리한계 - 관리한계 및 중심선의 색상을 지정합니다. 이들 선에 대한 레이블 색상도 지정할 수 있습니다. ■ 연결선 - 도표에서 포인트를 연결하는 선의 색상을 지정합니다. ■ 한계 외부 선 - 관리한계 외부에 놓여있는 선의 연결 부위 및 기호를 도표로 그리기 위한 색상을 지정합니다. 이 옵션은 이상상태점을 강조 표시하는 데 도움이 됩니다. ■ 한계 외부 영역 - 연결된 포인트와 관리한계 사이에 존재하며 연결선에 의해 경계가 지어진 관리한계 외부 영역에 대한 채우기 색상을 지정합니다. 이 옵션은 이상상태점을 강조 표시하는 데 도움이 됩니다. ■ 프레임 - 축과 프레임으로 둘러싸인 직사각형을 채울 색상을 지정합니다.
기호	데이터 포인트에 대한 기호를 지정합니다.

옵션 이름	설명
표시할 페이지 수	그래프를 표시하는 데 사용할 페이지 수를 지정합니다.
이동 범위	이동 범위를 계산하는 연속 측도 수를 지정합니다. 지정된 값은 2와 10(경계 포함) 사이여야 합니다. 기본값은 2입니다. 이 작업의 범위는 2 - 10으로 정해져 있습니다. SAS의 SHEWHART 프로시저를 사용하고 있으면, 더 높은 값을 지정할 수 있습니다. 자세한 내용은 Base SAS Procedures Guide의 SHEWHART 프로시저를 참조하십시오. 주: 이 옵션은 개별 측정값관리도 작업에서만 사용할 수 있습니다.
하위그룹 표본 크기 범례 생략	하위그룹 표본크기에 대한 기본 범례를 숨깁니다. 이 옵션은 관리한계 표본크기가 자동으로 그래프의 오른쪽 상단에 나타나기 때문에 하위그룹 표본크기가 일정하고 관리한계 표본크기와 동일할 때 유용합니다.
하위그룹 표본 크기로 (변수 대신) 숫자값 사용	크기에 대한 상수 값을 지정합니다. 하위그룹 표본 크기가 고정되어 있을 때 이 옵션을 사용합니다. 하위그룹 표본 크기 상자에 사용할 값을 입력합니다. 하위 그룹 표본 크기는 프로세스 변수의 최댓값과 99999(포함) 사이의 값이어야 합니다. 주: 하위그룹 표본 크기 역할에 변수를 할당하면 이 옵션 선택이 취소되고 이 역할에 할당된 변수가 하위그룹 표본 크기에 사용됩니다. 주: 이 옵션은 불량개수관리도, 불량률관리도 및 단위당 결점수관리도 작업에서만 사용할 수 있습니다.
두 번째 그래프 생략	두 번째 그래프가 표시되지 않게 합니다. 이 옵션은 일반적으로 개별 측정값관리도와 함께 사용되며 이때, 개별 측정값관리도는 생성되지만 그와 함께 생성되는 이동 범위관리도는 표시되지 않습니다. 주: 이 옵션은 개별 측정값관리도, 평균 및 범위관리도, 평균 및 표준편차관리도 작업에서만 사용할 수 있습니다.
여백 도표 유형	개별 측정값관리도에 표시되는 여백 도표의 유형을 지정합니다. 여백 도표는 관리도 통계량의 일변량 도표입니다. Digidot 도표, 히스토그램 또는 상자수염도를 선택할 수 있습니다. 이 도표는 개별 측정값관리도의 왼쪽 또는 오른쪽에 나타납니다. 기본적으로 No margin plot이 선택됩니다. 주: 이 옵션은 개별 측정값관리도 작업에서만 사용할 수 있습니다.

블록 변수 옵션 설정

이 옵션에 액세스하려면 선택 영역의 모양 머리글 아래에서 **블록 옵션**을 클릭합니다.

주: 이 옵션은 **블록 변수** 역할에 변수를 할당할 때만 사용할 수 있습니다.

역할 이름	설명
블록 레이블	블록 변수에 대한 레이블을 지정합니다.
범례 위치	블록 변수값에 대한 범례의 수직 위치를 지정합니다.
긴 값	블록 범례에 표시할 공간이 부족할 때 긴 블록 변수값을 처리하는 방법을 지정합니다. 기본적으로 긴 값은 표시되지 않습니다. 일정한 길이로 줄이기 옵션을 선택하면 값의 텍스트 길이가 범례에 맞게 줄어듭니다. 오른쪽 절단 옵션을 선택하면 긴 값의 오른쪽이 범례에 맞춰 잘립니다.
레이블 위치	블록 범례에서의 블록 변수 레이블 위치를 지정합니다. 범례 위치에 따라 사용할 수 있는 옵션이 다릅니다. 다음 옵션 중에서 선택할 수 있습니다. ■ 위 - 레이블을 범례 바로 위에 배치합니다. 기본값입니다. ■ 왼쪽 - 레이블을 범례 왼쪽에 배치합니다. ■ 오른쪽 - 레이블을 범례 오른쪽에 배치합니다. 범례를 관리도 위에 배치했을 때 이 옵션을 사용할 수 없습니다. 그래프 각 측면의 여백에 맞는 레이블에 대해서만 레이블 위치로 왼쪽 또는 오른쪽을 지정해야 합니다. 너무 긴 레이블은 잘립니다.
블록 레이블 색상	블록 레이블 텍스트의 색상을 지정합니다. 주: 이 옵션은 레이블 위치에 대해 위를 선택할 때만 사용할 수 있습니다.
범례에 반복된 값 표시	모든 하위그룹에 대한 블록 변수값을 표시합니다. 기본적으로 블록의 첫 번째 블록 변수값만이 표시되며 반복된 블록 변수값은 표시되지 않습니다.

부록 3

모델 생성

모델 정보	609
참고 항목	609
주효과 지정	610
교차효과(교호작용) 지정	610
지분 효과 지정	610
두 개 이상의 선택 변수에 대한 다중 교차 생성	610
양적변수의 다항 교차 생성	611
분석에서 절편 효과 제거	611

모델 정보

이 옵션에 액세스하려면 선택 영역에서 **모델**을 클릭합니다.

기본적으로 아무 효과도 지정되지 않으므로 절편 전용 모델이 생성됩니다. 효과를 지정하려면 한 개 이상의 변수를 **양적변수** 또는 **분류변수** 역할에 할당해야 합니다. 변수 조합을 선택하여 교차, 지분, 팩토리얼 또는 다항 효과를 생성할 수 있습니다.

참고 항목

- 37장, “High-Performance 선형회귀” (203페이지)
- 38장, “High-Performance 로지스틱 회귀” (213페이지)
- 48장, “로지스틱 회귀” (283페이지)

주효과 지정

- 1 **분류변수 및 양적변수** 상자에서 변수 이름을 선택합니다.
- 2 **주**를 클릭하여 **효과** 상자에 변수를 추가합니다.

교차효과(교호작용) 지정

- 1 하나 이상의 변수 이름을 선택합니다. 둘 이상의 변수를 선택하려면 Ctrl 키를 누른 상태에서 추가 변수 이름을 선택합니다.
- 2 **교차**를 클릭합니다. 변수를 한 개만 선택하면 변수는 자체 교차됩니다.

지분 효과 지정

- 1 중첩할 주효과 또는 교차효과를 하나 이상 선택합니다.
- 2 주효과 또는 교차효과를 중첩할 분류변수를 하나 이상 선택합니다.
- 3 **지분효과**를 클릭합니다. 중첩 효과는 **효과** 상자에 **변수 1(변수 2)** 형식으로 나타납니다.

두 개 이상의 선택 변수에 대한 다중 교차 생성

- 1 **차수** 필드의 숫자를 조정하여 더 높은 차수의 교차를 지정합니다.
- 2 **팩토리얼**을 클릭하여 **효과** 상자에 팩토리얼 효과를 추가합니다. 기본적으로 두 변수의 각 조합이 사용됩니다.

양적변수의 다항 교차 생성

- 1 차수 필드의 숫자를 조정하여 더 높은 차수의 교차를 지정합니다.
- 2 다항을 클릭하여 효과 상자에 다항 효과를 추가합니다.

분석에서 절편 효과 제거

절편 포함 체크 박스의 선택을 취소합니다. 이 체크 박스의 선택을 취소하면 한 개 이상의 효과를 지정해야 합니다. 절편은 기본적으로 포함됩니다.

부록 4

빌트인 작업

빌트인 작업, 관련 SAS 프로시저 및 라이선스 613

빌트인 작업, 관련 SAS 프로시저 및 라이선스

다음 테이블에는 **SAS Enterprise Guide** 및 **SAS Add-In for Microsoft Office**의 빌트인 작업이 나열되어 있습니다. 각 작업에는 해당 작업에 대한 기본 SAS 프로시저와 작업을 실행하는 데 필요한 SAS 소프트웨어 라이선스가 나열되어 있습니다.

주: **SAS Enterprise Guide** 또는 **SAS Add-In for Microsoft Office** 버전은 여기에 나열된 모든 작업을 포함하지 않을 수 있습니다. 작업 사용은 사이트에 필요한 SAS 소프트웨어의 라이선스 유무와 설치 여부에 따라 달라질 수 있습니다. 자세한 내용은 시스템 관리자에게 문의하십시오.

사이트에 **SAS Studio** 라이선스가 있고 설치되어 있으면, 추가 작업이 가능합니다. 자세한 내용은 SAS Studio 작업 및 관련 SAS 프로시저 및 라이선스를 참조하십시오.

빌트인 작업	사용되는 SAS 프로시저	SAS 라이선스 필요
테이블 추가	SQL	Base SAS
영역 도표	GPLOT	Base SAS, SAS/GRAPH
ARIMA 모형화 및 예측	ARIMA	Base SAS, SAS/ETS
프로젝트 라이브러리 할당	LIBNAME	Base SAS

빌트인 작업	사용되는 SAS 프로시저	SAS 라이선스 필요
막대 그래프	GCHART	Base SAS, SAS/GRAPH
막대 그래프 마법사 (37페이지)	GCHART	Base SAS, SAS/GRAPH
막대-선 그래프	GBARLINE	Base SAS, SAS/GRAPH
기본 예측 모형	FORECAST	Base SAS, SAS/ETS
상자관리도	SHEWHART	Base SAS, SAS/QC
상자 도표	GPLOT	Base SAS, SAS/GRAPH
버블 도표	GPLOT	Base SAS, SAS/GRAPH
결점수관리도 (75페이지)	SHEWHART	Base SAS, SAS/QC
정준상관분석	CANCORR	Base SAS, SAS/STAT
CDF 도표	CAPABILITY	Base SAS, SAS/QC
데이터 특성화	CONTENTS, UNIVARIATE, FREQ	Base SAS
군집분석	CLUSTER, FASTCLUS, TREE	Base SAS, SAS/STAT
데이터 비교	COMPARE	Base SAS
등고선 도표	GCONTOUR, G3GRID	Base SAS, SAS/GRAPH
상관분석	CORR	Base SAS
출력형식 생성	FORMAT	Base SAS
데이터셋에서 출력형식 생성 (129페이지)	FORMAT	Base SAS
지도용 테이블 생성 (133페이지)	GPROJECT	Base SAS, SAS/GRAPH
시계열 데이터 생성	TIMESERIES	Base SAS, SAS/ETS

빌트인 작업	사용되는 SAS 프로시저	SAS 라이선스 필요
데이터셋 속성	DATASETS	Base SAS
데이터셋 및 출력형식 삭제	SQL, CATALOG	Base SAS
판별분석	DISCRIM	Base SAS, SAS/STAT
분포분석	UNIVARIATE	Base SAS
도넛 그래프	GCHART	Base SAS, SAS/GRAPH
데이터 파일을 PC로 다운로드	SQL	Base SAS
요인분석	FACTOR	Base SAS, SAS/STAT
필터 및 정렬*	SQL	Base SAS
Forecast Studio 프로젝트 생성	없음	SAS Forecast Server
Forecast Studio 프로젝트 열기	없음	SAS Forecast Server
Forecast Studio 프로젝트 Override (189페이지)	없음	SAS Forecast Server
일반화선형모델	GENMOD	Base SAS, SAS/STAT
High-Performance 선형회귀	HPREG	Base SAS, SAS High Performance Analysis Server, SAS/STAT
High-Performance 로지스틱 회귀	HPLOGISTIC	Base SAS, SAS High Performance Analysis Server, SAS/STAT
히스토그램 (221페이지)	CAPABILITY	Base SAS, SAS/QC
데이터 가져오기(Microsoft Excel, 텍스트 파일, Microsoft Access)*	DATA 스텝, IMPORT	Base SAS, SAS/ACCESS to PC File Formats
JMP 파일 가져오기*	DATA 스텝, IMPORT	Base SAS, SAS/ACCESS to PC File Formats

빌트인 작업	사용되는 SAS 프로시저	SAS 라이선스 필요
SPSS 파일 가져오기*	DATA 스텝, IMPORT	Base SAS, SAS/ACCESS to PC File Formats
Stata 파일 가져오기*	DATA 스텝, IMPORT	Base SAS, SAS/ACCESS to PC File Formats
개별 측정값관리도	SHEWHART	Base SAS, SAS/QC
JMP 스토어드 프로세스 패키지*	DATA 스텝	JMP
생명표 분석	LIFETEST	Base SAS, SAS/STAT
선 도표	GPLOT	Base SAS, SAS/GRAPH
선 도표 마법사	GPLOT	Base SAS, SAS/GRAPH
선형모델	GLM	Base SAS, SAS/STAT
선형회귀	REG	Base SAS, SAS/STAT
데이터 리스트	PRINT	Base SAS
리스트 리포트 마법사 (279페이지)	REPORT	Base SAS
로지스틱 회귀	LOGISTIC	Base SAS, SAS/STAT
지도 그래프	GMAP	Base SAS, SAS/GRAPH
평균 및 범위관리도	SHEWHART	Base SAS, SAS/QC
평균 및 표준편차관리도	SHEWHART	Base SAS, SAS/QC
혼합모형	MIXED	Base SAS, SAS/STAT
모델 스코어링	없음	SAS Enterprise Miner
비선형회귀	NLIN	Base SAS, SAS/STAT
비모수 일원분산분석	NPAR1WAY	Base SAS, SAS/STAT
불량개수관리도	SHEWHART	Base SAS, SAS/QC

빌트인 작업	사용되는 SAS 프로시저	SAS 라이선스 필요
일원분산분석	ANOVA	Base SAS, SAS/STAT
일원빈도분석	FREQ	Base SAS
SAS ODS 통계적 그래프	없음	ODS Graphics Designer
불량률관리도	SHEWHART	Base SAS, SAS/QC
파레토 그래프	PARETO	Base SAS, SAS/QC
원 그래프	GCHART	Base SAS, SAS/GRAPH
원 그래프 마법사	GCHART	Base SAS, SAS/GRAPH
P-P 도표	CAPABILITY	Base SAS, SAS/QC
시계열 데이터 준비	EXPAND	Base SAS, SAS/ETS
주성분분석	PRINCOMP	Base SAS, SAS/STAT
확률 도표	CAPABILITY	Base SAS, SAS/QC
비례위험분석	PHREG	Base SAS, SAS/STAT
Q-Q 도표	CAPABILITY	Base SAS, SAS/QC
질의 빌더*	SQL	Base SAS
레이더 그래프	GRADAR	Base SAS, SAS/GRAPH
확률표본	SURVEYSELECT	Base SAS, SAS/STAT
데이터 순위화	RANK	Base SAS
Rapid Predictive Modeler	SAS Enterprise Miner 매크로	SAS Enterprise Miner
RFM(Recency, Frequency, Monetary) 분석	SAS 매크로	Base SAS
패널 데이터의 회귀분석	TSCSREG	Base SAS, SAS/ETS

빌트인 작업	사용되는 SAS 프로시저	SAS 라이선스 필요
자기회귀 오차가 있는 회귀분석	AUTOREG	Base SAS, SAS/ETS
산점도	GPLOT, G3D	Base SAS, SAS/GRAPH
산점도 행렬	SGSCATTER	Base SAS, SAS/GRAPH
ODS 통계적 그래프 표시	SGDESIGN	Base SAS
데이터 정렬	SORT	Base SAS
칼럼 분할	TRANSPOSE	Base SAS
칼럼 세로 배열	TRANSPOSE	Base SAS
데이터 표준화	STANDARD	Base SAS
요약통계량	MEANS	Base SAS
요약통계량 마법사	MEANS	Base SAS
요약테이블 (533페이지)	TRANSPOSE	Base SAS
요약테이블 마법사	TRANSPOSE	Base SAS
표면도	G3D	Base SAS, SAS/GRAPH
t 검정	TTEST	Base SAS, SAS/STAT
테이블 분석	FREQ	Base SAS
타일 그래프	GTILE	Base SAS, SAS/GRAPH
전치	TRANSPOSE	Base SAS
단위당 결점수관리도	SHEWHART	Base SAS, SAS/QC
라이브러리 메타데이터 업데이트*	METALIB	Base SAS
데이터 파일을 서버에 업로드*	SQL	Base SAS

주: 별표가 뒤에 오는 작업은 SAS Enterprise Guide에서만 사용할 수 있습니다.
