버전 16

JMP 살펴보기

"진정 무엇인가를 발견하는 여행은 새로운 풍경을 바라보는
것이 아니라 새로운 눈을 가지는 데 있다 ."

Marcel Proust

JMP, A Business Unit of SAS
SAS Campus Drive
Cary, NC 27513

16.0

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2020–2021. JMP® 살펴보기 16. Cary, NC: SAS Institute Inc.

JMP 살펴보기® 16

Copyright © 2020-2021, SAS Institute Inc., Cary, NC, USA

All rights reserved. Produced in the United States of America.

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, North Carolina 27513-2414.

2021 년 3 월

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.

SAS software may be provided with certain third-party software, including but not limited to open-source software, which is licensed under its applicable third-party software license agreement. For license information about third-party software distributed with SAS software, refer to http://support.sas.com/thirdpartylicenses.

JMP 활용하기

JMP를 처음 사용하든 오랫동안 사용해왔든, JMP와 관련해서 배워야 하는 것은 언제나 있습니다. JMP.com 을 방문하면 다음과 같은 유용한 자료를 볼 수 있습니다.

- JMP 시작 방법에 대한 라이브 및 녹화 웹 캐스트
- 새로운 기능과 고급 기법에 대한 비디오 데모 및 웹 캐스트
- JMP 교육 과정 등록에 대한 상세 정보
- 현지에서 개최되는 세미나 일정
- 다른 사용자들이 JMP 를 사용하는 방법을 보여주는 성공 사례
- JMP 직원의 팁 , 힌트 및 사례가 포함된 블로그
- 다른 사용자와 JMP 에 대해 토론할 수 있는 포럼

https://www.jmp.com/getstarted

목차

JMP 살펴보기

설명서 소개 9

JMP 그래프 갤러리 11

1 JMP 알아보기 31

설명서 및 추가 리소스

JMP 설명서에 사용되는 서식 규칙 33
JMP 도움말 34
JMP 설명서 라이브러리 34
JMP 학습을 위한 추가 리소스 40
JMP 자습서 40
샘플 데이터 테이블 40
통계 및 JSL 용어에 대해 알아보기 41
JMP 팁 및 힌트 알아보기 41
JMP 툴팁 41
JMP 사용자 커뮤니티 41
무료 온라인 Statistical Thinking 교육 과정 42
새로운 사용자를 위한 JMP 입문 키트 42
Statistics Knowledge 포털 42
JMP 교육 과정 42
사용자가 작성한 JMP 설명서 42
JMP 시작하기 창 42
JMP 기술 지원 43

2 JMP 소개 45

기본 개념

JMP 필수 개념 47
JMP 시작하기 47
JMP 시작 48
샘플 데이터 사용 49
데이터 테이블 이해 51
JMP 워크플로우 이해 52
1 단계 : JMP 플랫폼 시작 및 결과 보기 53
2 단계 : JMP 보고서에서 상자 그림 제거 54
3 단계 : 추가 JMP 출력 요청 55

4 단계 : JMP 플랫폼 결과에서의 상호 작용 56
JMP 와 Excel 의 차이점 57
데이터 테이블의 구조 57
JMP 의 계산식 58
JMP 분석 및 그래프 생성 59

3 데이터 작업 61
그래프 생성 및 분석을 위한 데이터 준비

JMP 로 데이터 가져오기 63
데이터 테이블에 데이터 복사 및 붙여넣기 63
데이터 테이블로 데이터 가져오기 63
데이터 테이블에 데이터 입력 66
Excel 에서 JMP 로 데이터 전송 68
데이터 테이블 작업 69
데이터 테이블에서 데이터 편집 70
데이터 테이블에서 값 선택 , 선택 취소 및 찾기 72
데이터 테이블에서 열 정보 보기 또는 변경 75
계산식으로 값 계산의 예 77
보고서 데이터 필터링의 예 78
데이터 재구성의 예 80
요약 통계량 보기 80
부분집합 생성의 예 85
예 : 데이터 테이블 결합 87
데이터 정렬 88

4 데이터 시각화 91
일반적인 그래프

단변량 그래프의 단일 변수 분석 93
연속형 변수에 히스토그램 사용 93
범주형 변수에 막대 차트 사용 96
다중 변수 비교 98
산점도를 사용하여 여러 변수 비교 99
산점도 행렬을 사용하여 여러 변수 비교 102
병렬 상자 그림을 사용하여 여러 변수 비교 105
그래프 빌더를 사용하여 여러 변수 비교 107
버블 그림을 사용하여 여러 변수 비교 113
중첩 그림을 사용하여 여러 변수 비교 118
계량형 차트를 사용하여 여러 변수 비교 122

5 데이터 분석 127
분포 , 관계 및 모형

내용 소개 129
데이터 그래프 생성의 중요성 129

모델링 유형 이해 132
예 : 모델링 유형 결과 보기 133
모델링 유형 변경 134
분포 분석 136
연속형 변수의 분포 137
범주형 변수의 분포 139
관계 분석 142
하나의 예측 변수가 있는 회귀 사용 143
한 변수에 대한 평균 비교 148
비율 비교 151
여러 변수의 평균 비교 153
다중 예측 변수가 있는 회귀 사용 157

6 전체를 보는 시각 163
여러 플랫폼에서 데이터 탐색

재미있는 사실 : 연결된 분석 165
여러 플랫폼에서 데이터 탐색 165
분포 플랫폼에서 분포 분석 165
다변량 플랫폼에서 패턴 및 관계 분석 169
군집화 플랫폼에서 유사 값 분석 172

7 작업 저장 및 공유 177
결과 저장 및 다시 생성

프로젝트 작업 179
새 프로젝트 생성 180
프로젝트에서 파일 열기 180
프로젝트에서 파일 재배열 180
프로젝트 저장 182
프로젝트 작업 영역 182
프로젝트 책갈피 183
프로젝트 내용 184
프로젝트의 최근 사용한 파일 184
프로젝트 로그 185
파일을 프로젝트로 이동 185
프로젝트 내의 열기 스크립트 작성 186
새 프로젝트 생성의 예 186
플랫폼 결과를 저널로 저장 189
예 : 저널 생성 189
저널에 분석 추가 190
스크립트 저장 및 실행 191
예 : 스크립트 저장 및 실행 191
스크립트 및 JSL 192

보고서를 대화식 HTML 로 저장 192
대화식 HTML 의 데이터 포함 193
예 : 대화식 HTML 생성 194
보고서를 대화식 HTML 로 공유 195
보고서를 PowerPoint 프레젠테이션으로 저장 199
대시보드 생성 199
예 : 창 결합 200
예 : 두 개의 보고서가 포함된 대시보드 생성 201

8 특수 기능 203
자동 분석 업데이트 및 SAS Integration
분석 및 그래프 자동 업데이트 205
예 : 자동 재계산 사용 205
환경 설정 변경 209
예 : 환경 설정 변경 209
JMP 와 SAS 통합 212
예 : SAS 코드 생성 212
예 : SAS 코드 전송 213

A 기술 공지사항 215

설명서 소개

JMP 살펴보기에서는 JMP 소프트웨어와 관련한 일반적인 내용을 소개합니다. 이 설명서에서는 사용자가 JMP에 대한 지식이 없다고 가정합니다. 이 설명서에서는 분석가, 연구원, 학생, 교수 또는 통계학자 등을 대상으로 JMP의 사용자 인터페이스 및 기능에 대한 일반적인 개요를 제공합니다.

이 설명서에서는 다음과 같은 정보를 소개합니다.

- JMP 시작
- JMP 창의 구조
- 데이터 준비 및 조작
- 대화식 그래프를 기반으로 한 데이터 파악
- 간단한 분석을 통해 그래프 보강
- JMP 및 특수 기능 사용자 정의
- 결과 공유

이 설명서에는 6 개의 장이 있습니다. 각 장에서는 해당 장에서 제시된 개념을 보강하는 예를 소개합니다. 모든 통계적 개념은 소개 수준으로 제공됩니다. JMP 살펴보기에 사용된 샘플 데이터는 소프트웨어에 포함되어 있습니다. 다음은 각 장에 대한 설명입니다.

- 2 장, "JMP 소개" 에서는 JMP 응용 프로그램의 개요를 제공합니다. 콘텐츠가 구성된 방식과 소프트웨어를 탐색하는 방법을 안내합니다.
- 3 장, " 데이터 작업 " 에서는 다양한 소스에서 데이터를 가져와 분석을 위해 준비하는 방법을 설명합니다. 데이터 조작 도구에 대한 개요도 제공합니다.
- 4 장, " 데이터 시각화 " 에서는 데이터를 시각화하고 이해하기 위해 사용할 수 있는 그래프 및 차트에 대해 설명합니다. 단일 변수가 포함된 간단한 분석에서 여러 변수 간의 관계를 볼 수 있게 해 주는 다중 변수 그래프에 이르기까지 다양한 예제가 있습니다.
- 5 장, " 데이터 분석 " 에서는 일반적으로 사용되는 여러 가지 분석 기법을 설명합니다. 통계적 방법을 사용하지 않아도 되는 간단한 기법에서부터 통계 지식이 유용하게 사용되는 고급 기법에 이르기까지 다양합니다.
- 6 장, " 전체를 보는 시각 " 에서는 여러 플랫폼에서 분포, 패턴 및 유사한 값을 분석하는 방법을 보여 줍니다.
- 7 장, " 작업 저장 및 공유 " 에서는 PowerPoint 프레젠테이션 및 대화식 HTML 에서 JMP 를 사용하지 않는 사람들과 작업을 공유하는 방법을 설명합니다. 분석을 스크립트로 저장하고 JMP 사용자를 위해 저널 및 프로젝트에 작업을 저장하는 방법도 다룹니다.

- 8 장 , " 특수 기능 " 에서는 데이터가 변경될 때 그래프 및 분석을 자동으로 업데이트하는 방법 , 환경 설정을 사용하여 보고서를 사용자 정의하는 방법 , JMP 가 SAS 와 상호 작용하는 방법에 대해 설명합니다 .

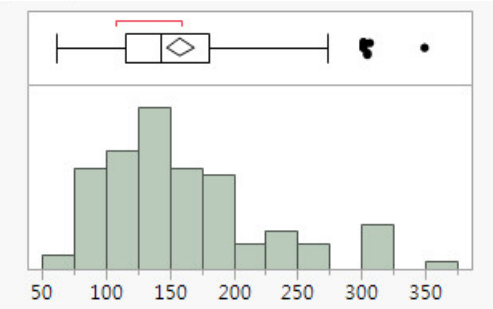
이 설명서를 검토한 후에는 JMP에서 편안하게 데이터를 탐색하고 작업할 수 있습니다.

JMP는 Windows 및 macOS 운영 체제 모두에서 사용할 수 있습니다. 그러나 이 설명서의 내용은 Windows 운영 체제를 기반으로 합니다.

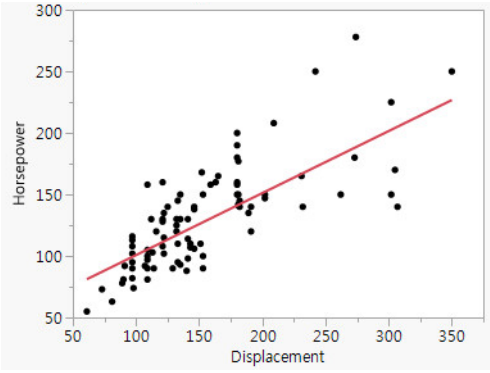
JMP 그래프 갤러리

다양한 그래프 및 플랫폼

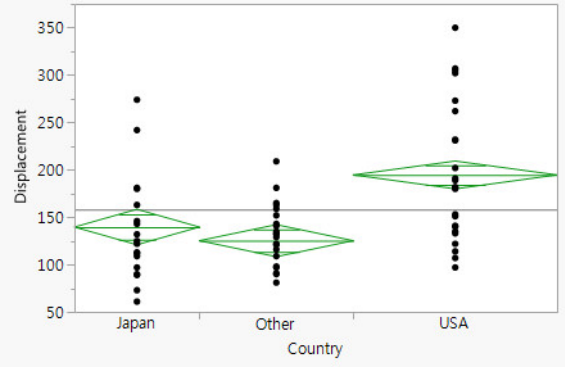
다음은 JMP로 생성할 수 있는 여러 가지 그래프의 그림입니다. 각 그림의 라벨에는 해당 그림을 생성하는 데 사용되는 플랫폼이 지정되어 있습니다. 플랫폼과 이러한 그래프 및 기타 그래프에 대한 자세한 내용은 **도움말 > 사용 설명서** 메뉴의 문서에서 확인하십시오.



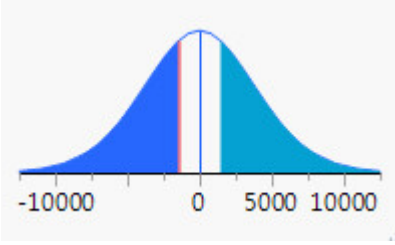
히스토그램
분석 > 분포



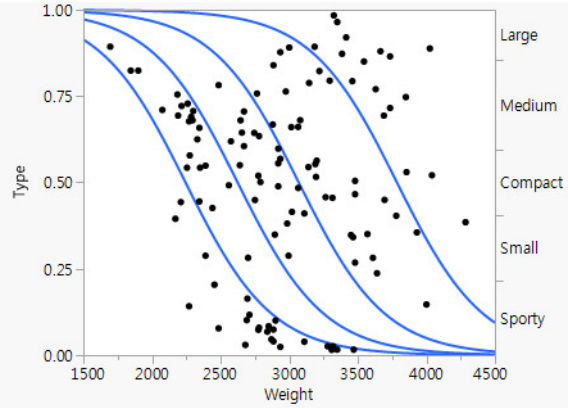
이변량
분석 > X 로 Y 적합



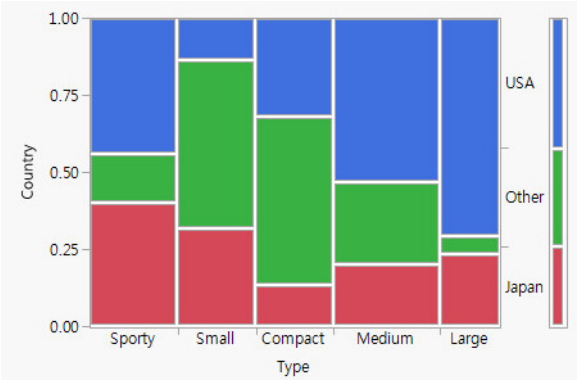
일원 분석
분석 > X 로 Y 적합



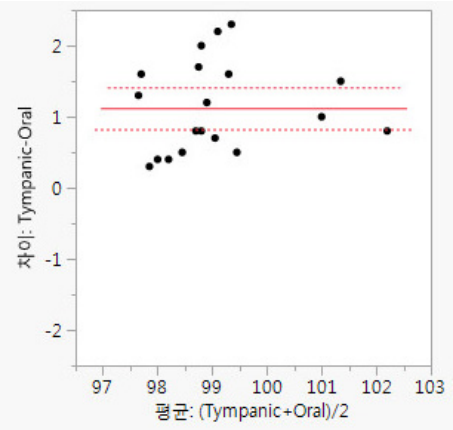
일원 t- 검정
분석 > X 로 Y 적합



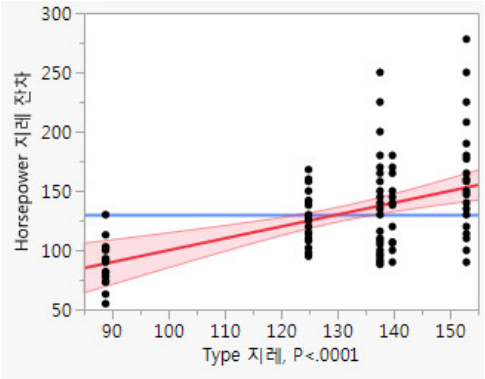
로지스틱
분석 > X 로 Y 적합



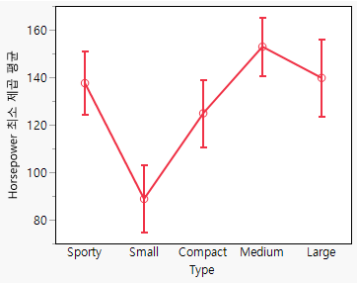
모자이크 그림
분석 > X 로 Y 적합



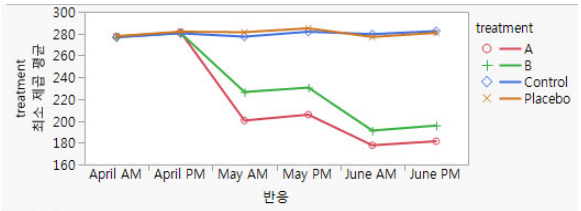
매칭 쌍
분석 > 전문 모형 > 매칭 쌍



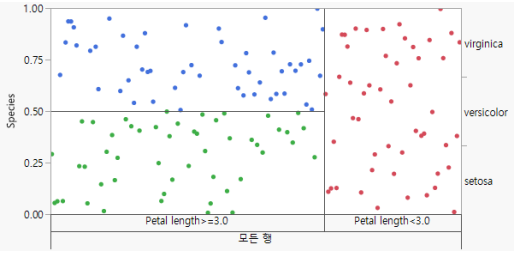
레버리지 그림
분석 > 모형 적합



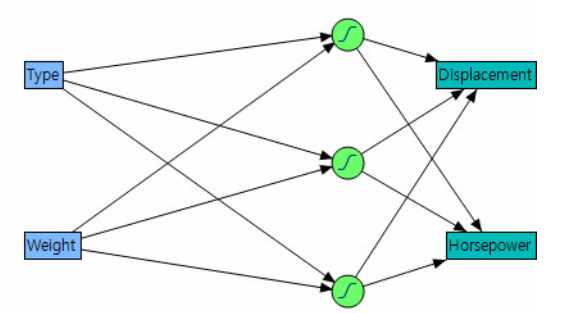
최소 제곱 평균 그림
분석 > 모형 적합



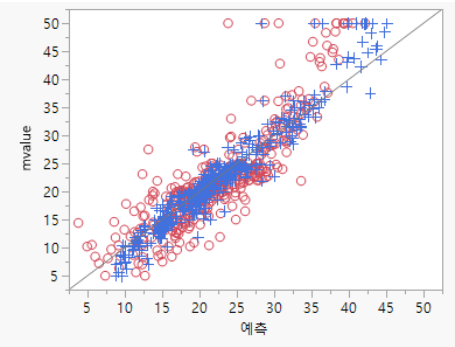
MANOVA
분석 > 모형 적합



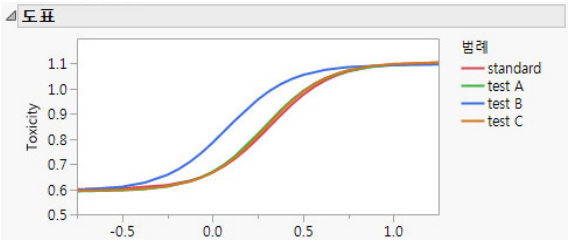
파티션
분석 > 예측 모형 > 파티션



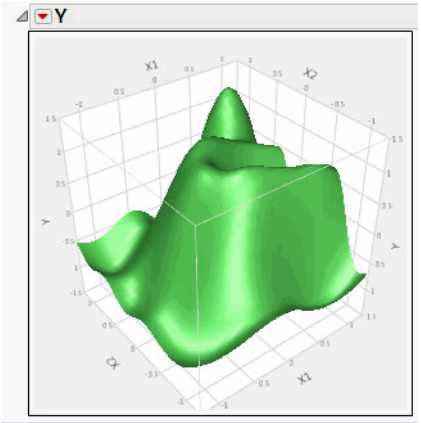
신경망 다이어그램
분석 > 예측 모형 > 신경망



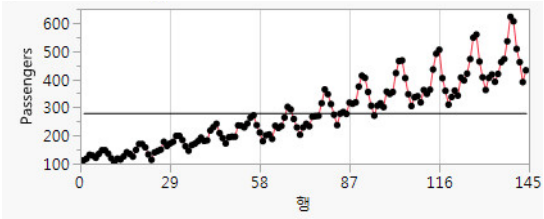
실제값 대 예측값
분석 > 예측 모형 > 모형 비교



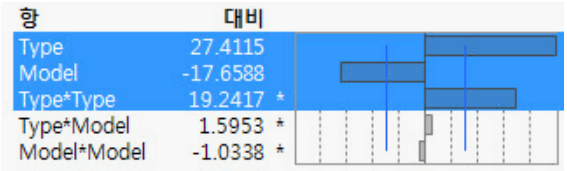
비선형 적합
분석 > 전문 모형 > 비선형



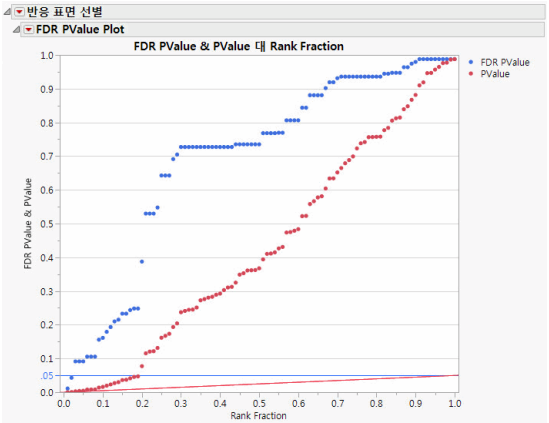
표면 프로파일러
분석 > 전문 모형 > 가우스 과정



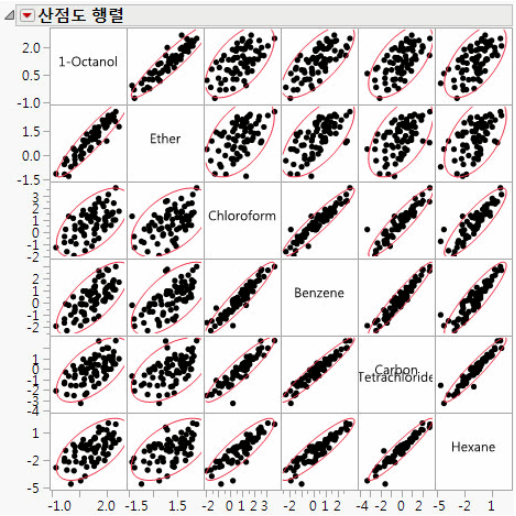
시계열
분석 > 전문 모형 > 시계열



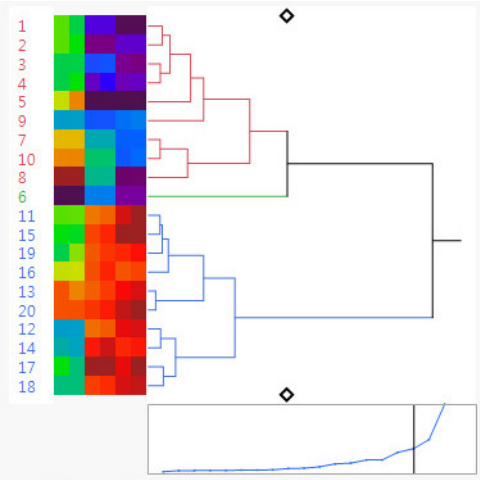
선별
분석 > 전문 모형 > 특수 DOE 모형 > 2 수준 선별 적합



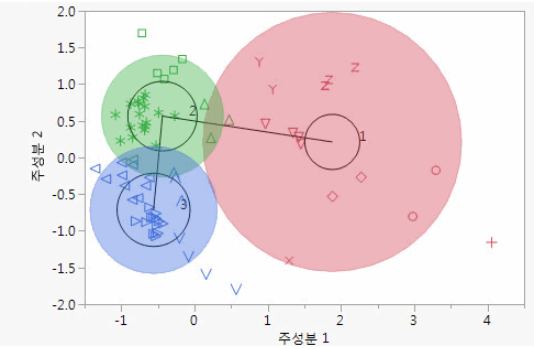
FDR p 값 그림
분석 > 선별 > 반응 변수 선별



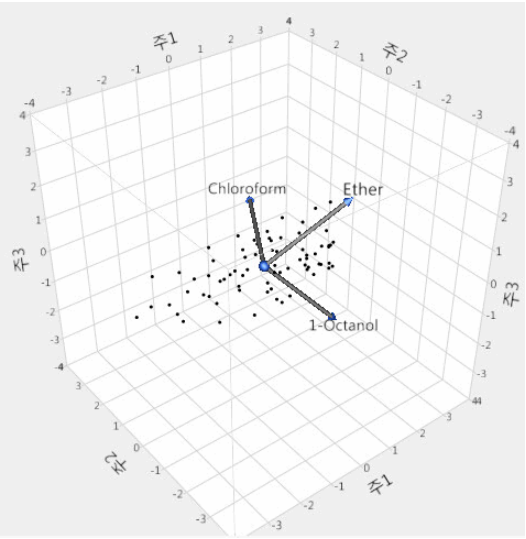
산점도 행렬
분석 > 다변량 방법 > 다변량



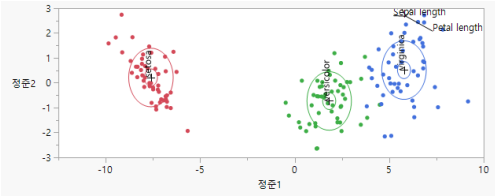
덴드로그램
분석 > 군집화 > 계층적 군집화



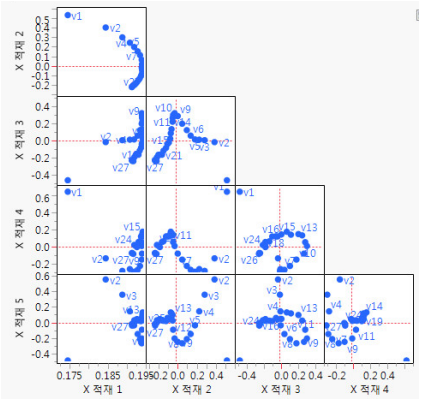
자기 조직화 지도
분석 > 군집화 > K 평균 군집화



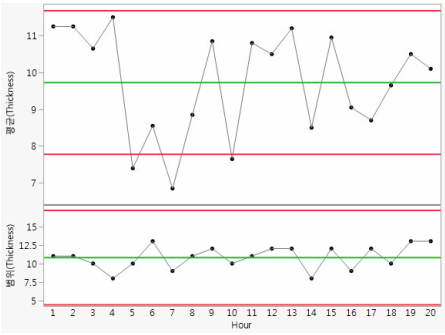
주성분
분석 > 다변량 방법 > 주성분



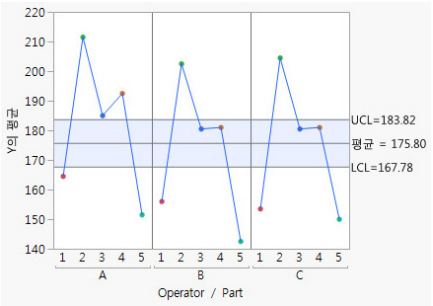
정준 그림
분석 > 다변량 방법 > 판별



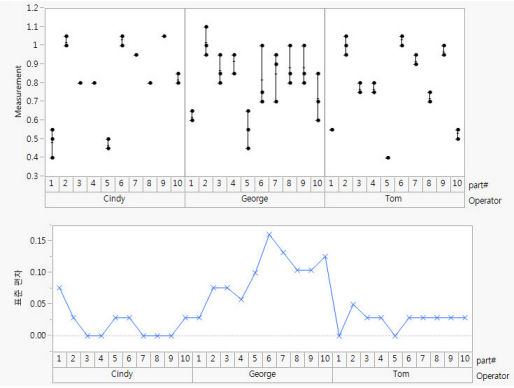
적재 그림
분석 > 다변량 방법 > 부분 최소 제곱



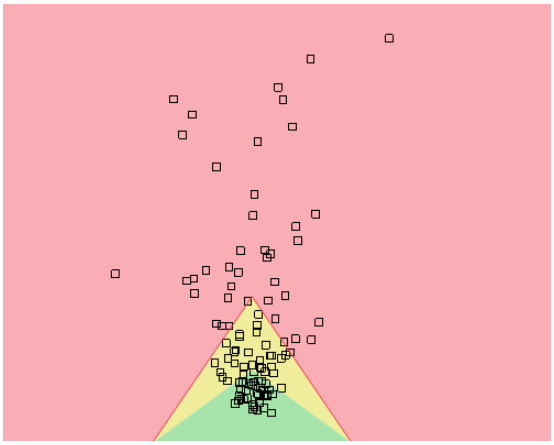
Xbar-R 차트
분석 > 품질 및 공정 > 관리도 빌더



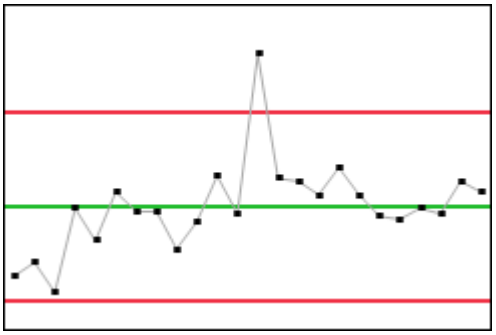
평균 차트
분석 > 품질 및 공정 > 측정 시스템 분석



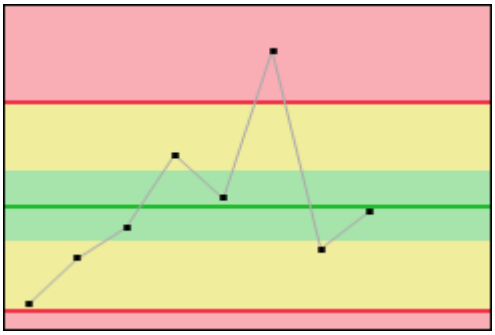
계량형 차트
분석 > 품질 및 공정 > 계량형 / 계수형 게이지 차트



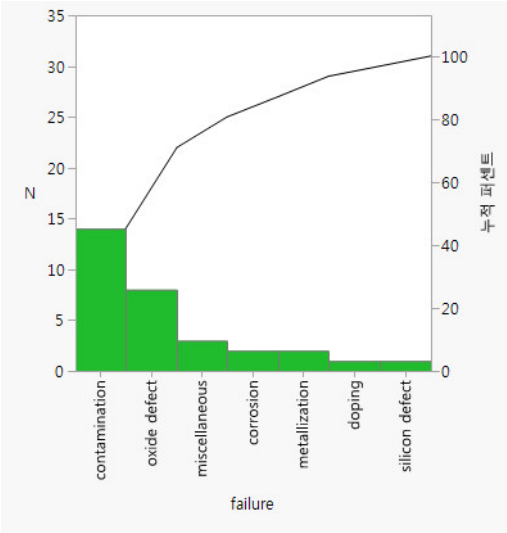
목표 그림
분석 > 품질 및 공정 > 공정 능력



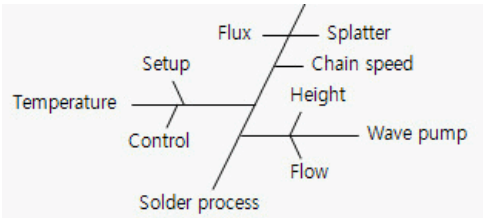
개별 측정값 차트
이동 범위 차트
분석 > 품질 및 공정 > 관리도 > IR



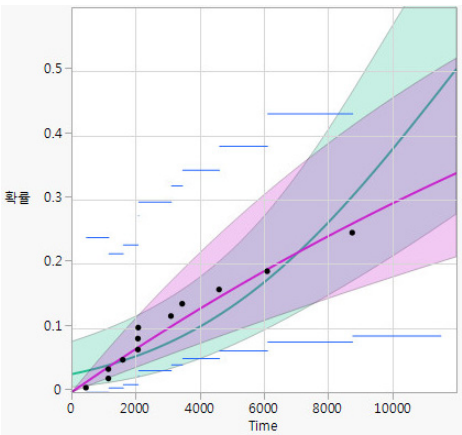
Xbar 차트
분석 > 품질 및 공정 > 관리도 > Xbar



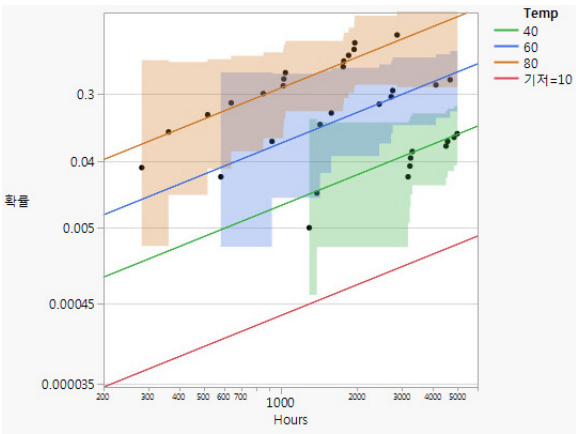
파레토도
분석 > 품질 및 공정 > 파레토도



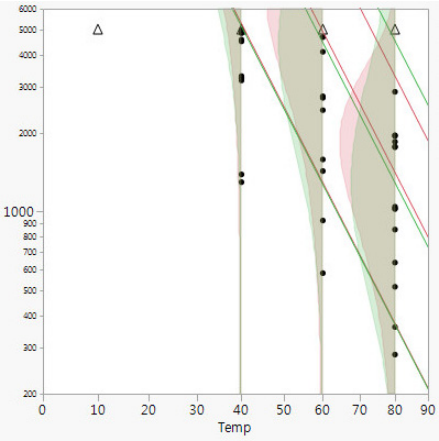
Ishikawa 차트
Fishbone 차트
분석 > 품질 및 공정 > 다이어그램



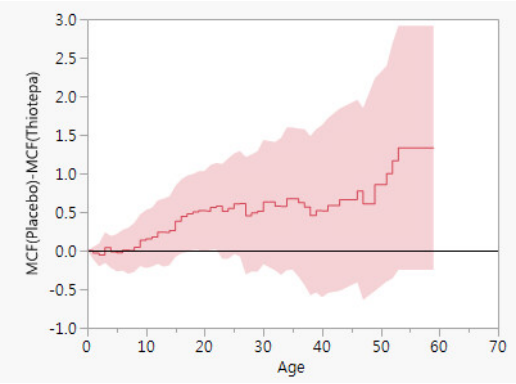
분포 비교
분석 > 신뢰성 및 생존 > 수명 분포



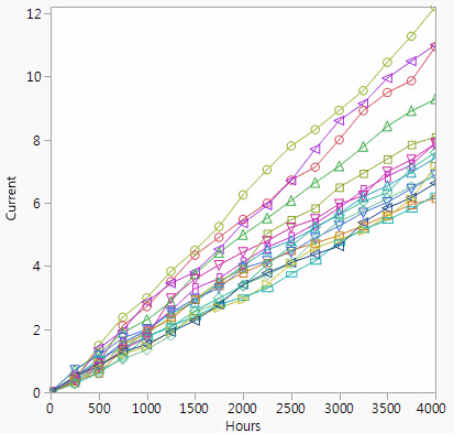
비모수 중첩
분석 > 신뢰성 및 생존 > 수명 분포 적합



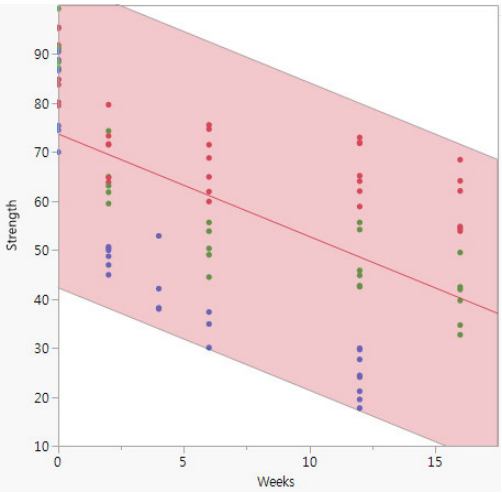
산점도
분석 > 신뢰성 및 생존 > 수명 분포 적합



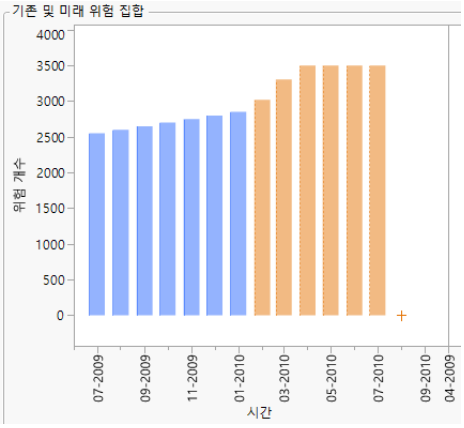
MCF 그림
분석 > 신뢰성 및 생존 > 재발 분석



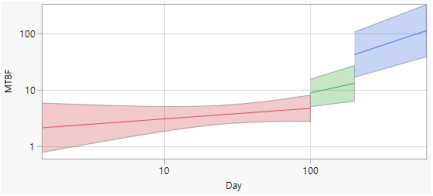
중첩
분석 > 신뢰성 및 생존 > 열화



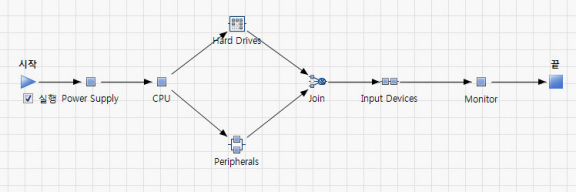
예측 구간
분석 > 신뢰성 및 생존 > 파괴 열화



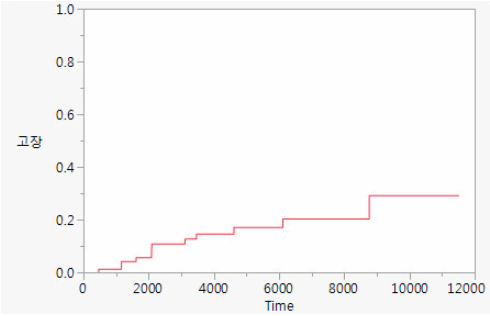
예측
분석 > 신뢰성 및 생존 > 신뢰도 예측



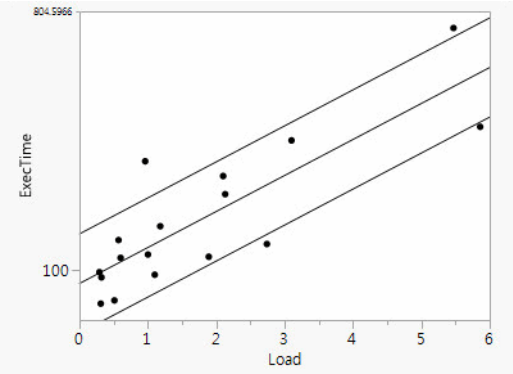
구간적 Weibull NHPP
분석 > 신뢰성 및 생존 > 신뢰도 성장



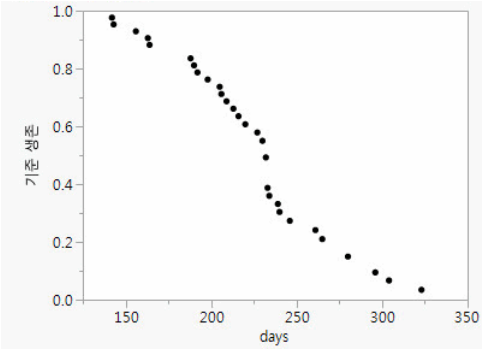
신뢰도 블록 다이어그램
분석 > 신뢰성 및 생존 > 신뢰도 블록 다이어그램



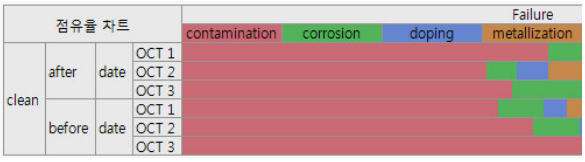
고장 그림
분석 > 신뢰성 및 생존 > 생존



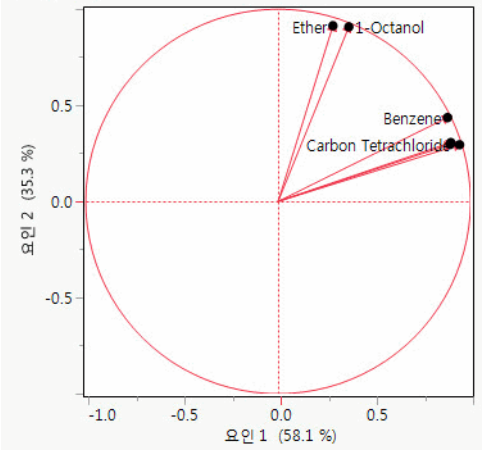
생존 분위수
분석 > 신뢰성 및 생존 > 모수 생존 모형 적합



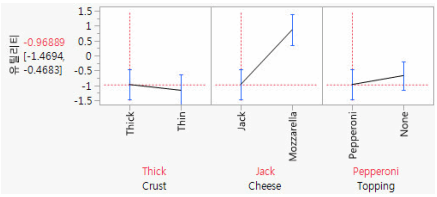
기준 생존율
분석 > 신뢰성 및 생존 > 비례 위험 모형 적합



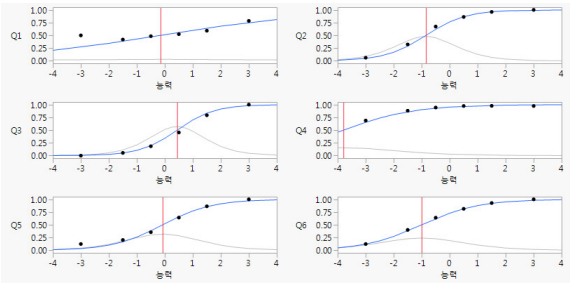
혼합물 프로파일러
분석 > 소비자 조사 > 범주형



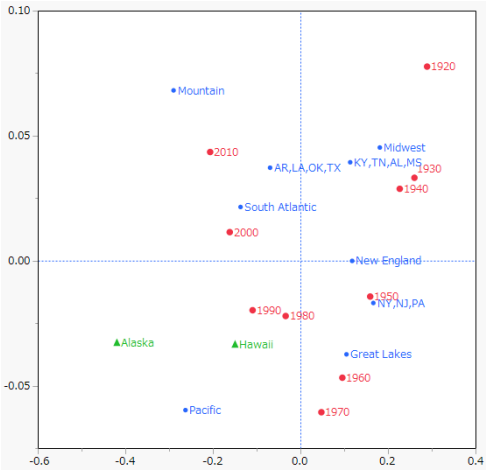
요인 적재 그림
분석 > 다변량 방법 > 요인 분석



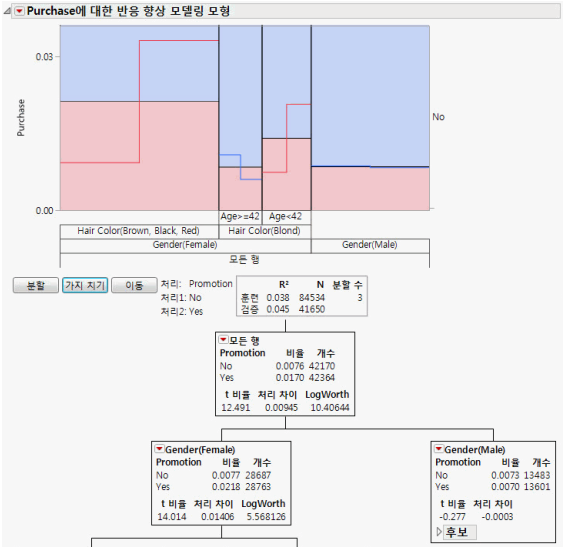
예측 프로파일
분석 > 소비자 조사 > 선택



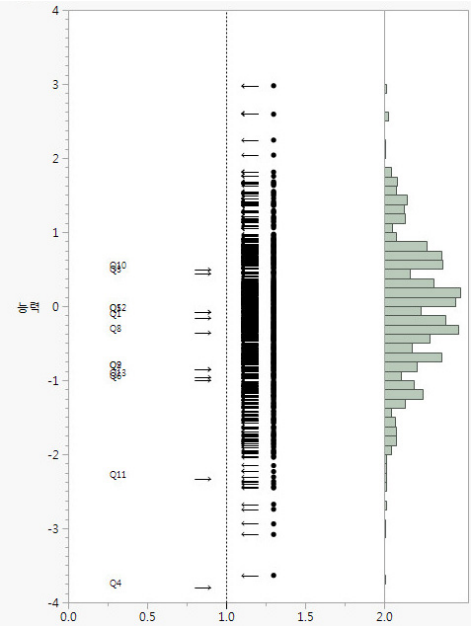
특성 곡선
분석 > 다변량 방법 > 항목 분석



다중 대응 분석
분석 > 다변량 방법 > 다중 대응 분석

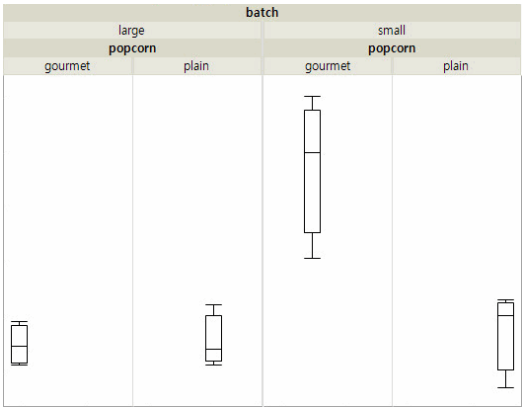


Uplift 모형
분석 > 소비자 조사 > Uplift

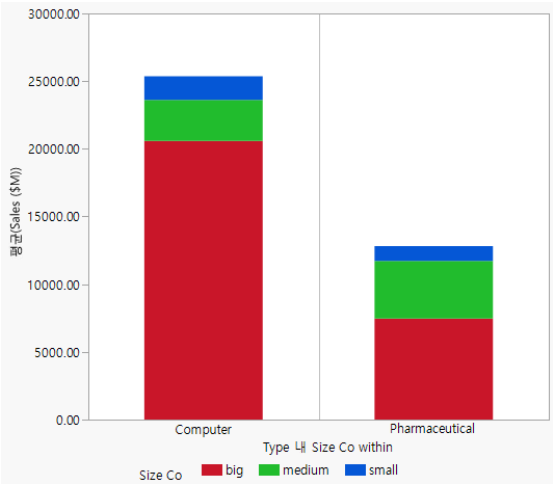


쌍대 그림
분석 > 다변량 방법 > 항목 분석

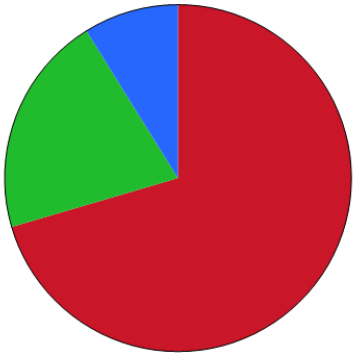
선 그래프
그래프 > 그래프 빌더



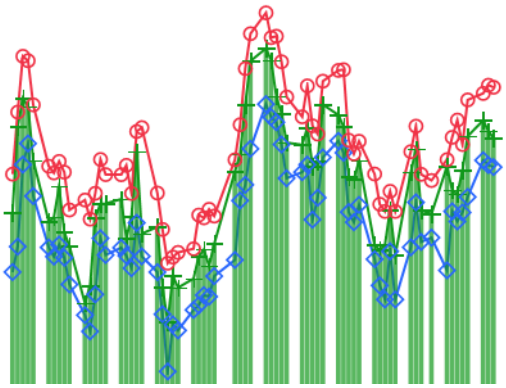
상자 그림
그래프 > 그래프 빌더



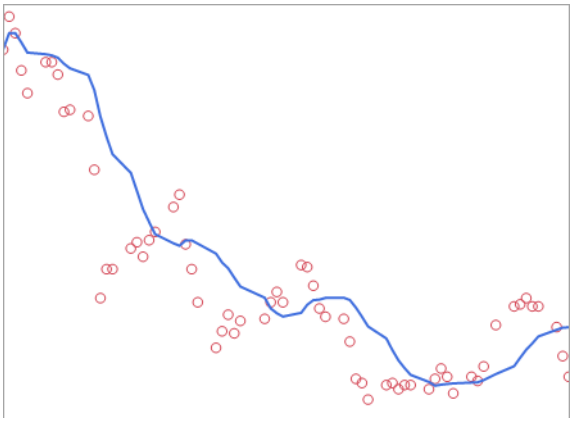
누적 막대 차트
그래프 > 그래프 빌더



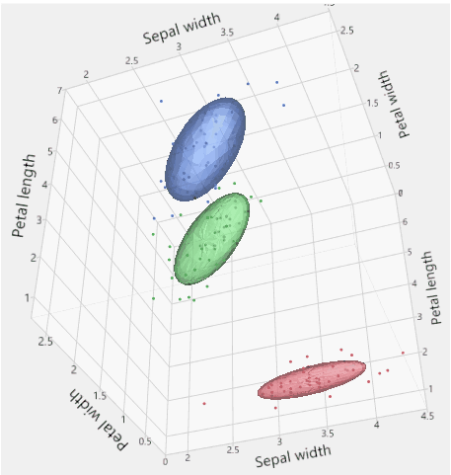
파이 차트
그래프 > 그래프 빌더



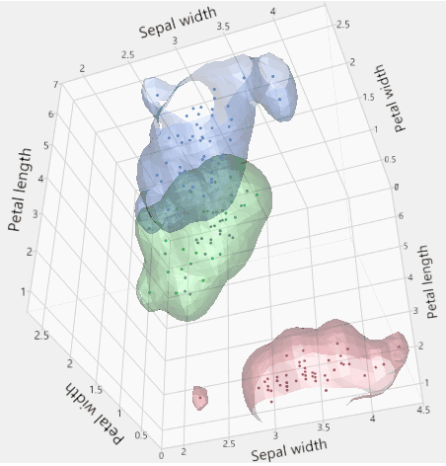
바늘 및 선 차트
그래프 > 그래프 빌더



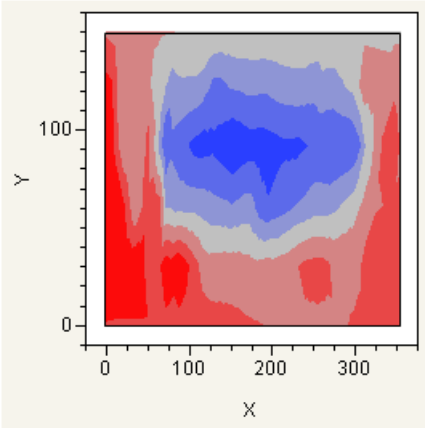
평활기
그래프 > 그래프 빌더



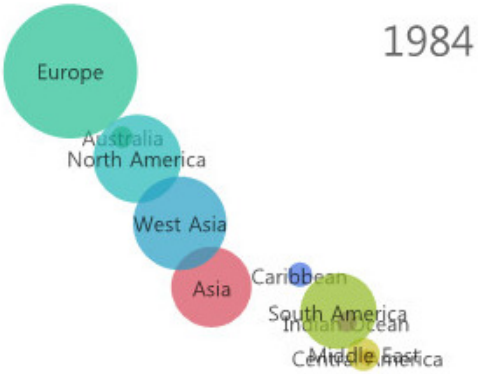
3 차원 산점도
그래프 > 3D 산점도



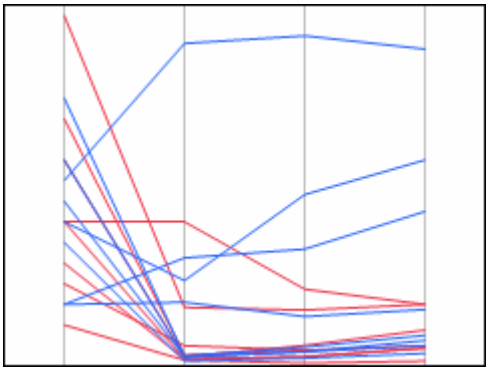
3 차원 산점도
그래프 > 3D 산점도



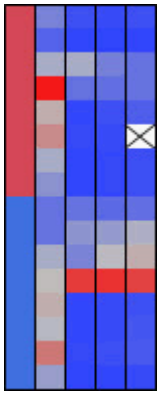
등고선 그림
그래프 > 그래프 빌더



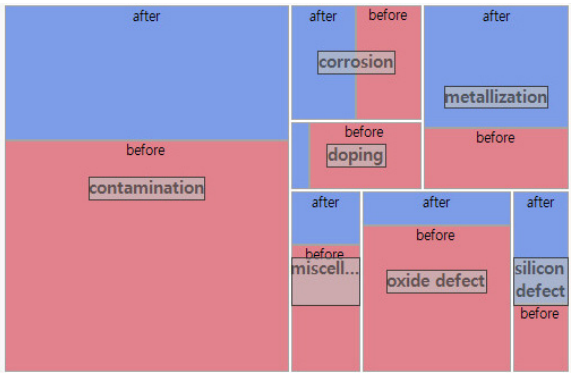
버블 그림
그래프 > 버블 그림



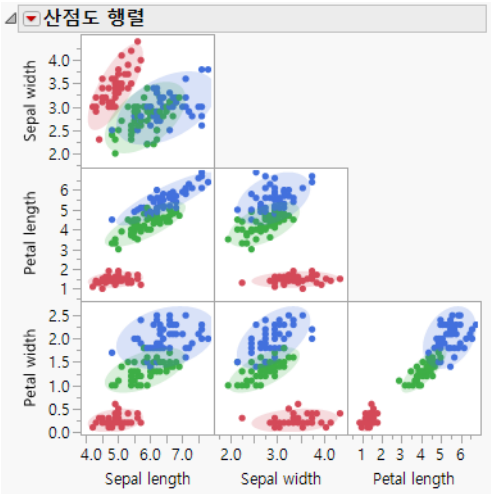
평행 그림
그래프 > 그래프 빌더



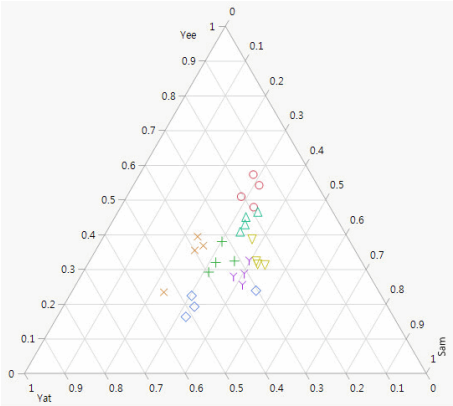
셀 그림
그래프 > 셀 그림



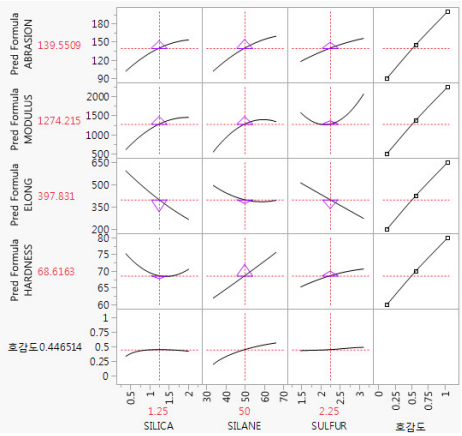
트리맵
그래프 > 그래프 빌더



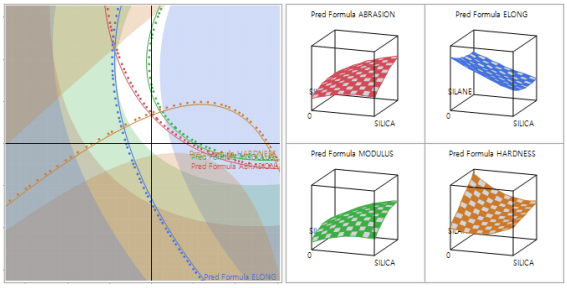
산점도 행렬
그래프 > 산점도 행렬



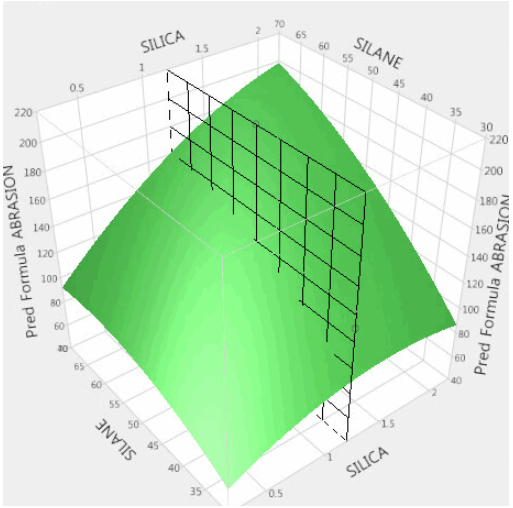
삼원 그림
그래프 > 삼원 그림



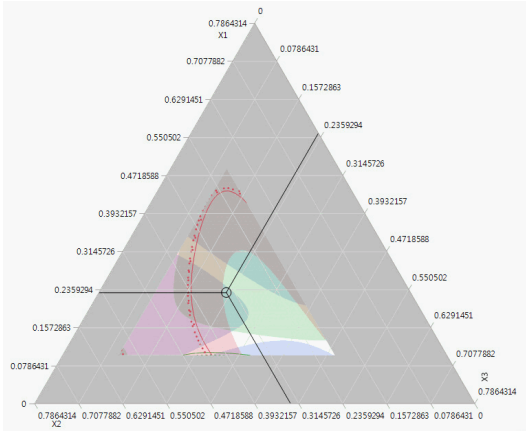
예측 프로파일러
그래프 > 프로파일러



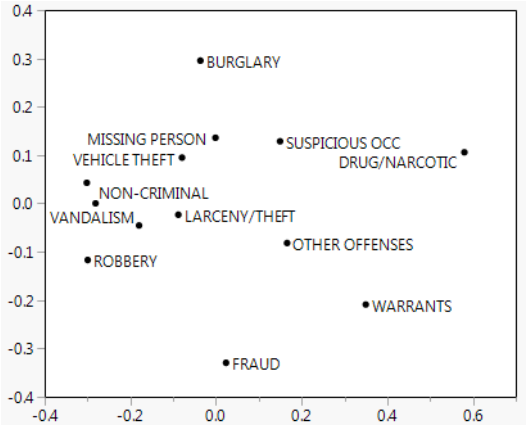
등고선 프로파일러
그래프 > 등고선 프로파일러



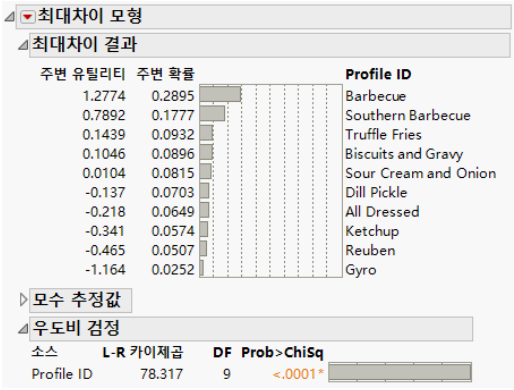
표면 그림
그래프 > 표면 그림



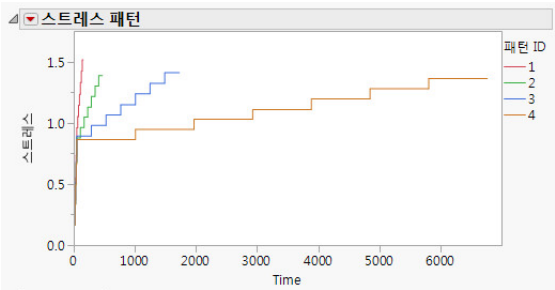
혼합물 프로파일러
그래프 > 혼합물 프로파일러



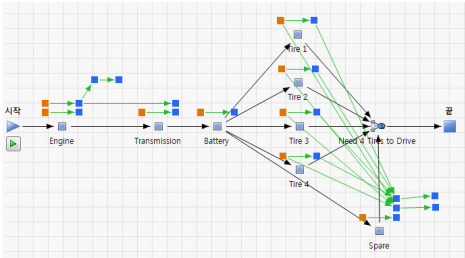
다차원 척도법
분석 > 다변량 방법 > 다차원 척도법



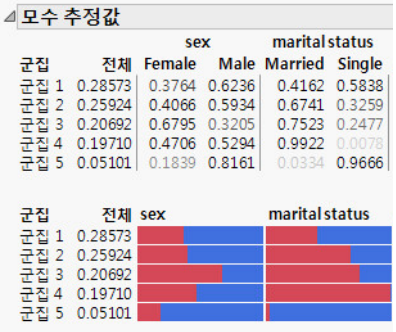
최대차이
분석 > 소비자 조사 > 최대차이



스트레스 패턴 그림
분석 > 신뢰성 및 생존 > 누적 손상



수리 가능 시스템 시뮬레이션
분석 > 신뢰성 및 생존 > 수리 가능 시스템 시뮬레이션

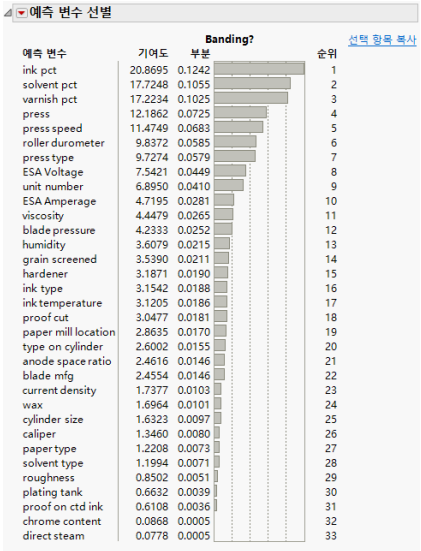


모수 추정값

군집	전체	sex Female	sex Male	marital status Married	marital status Single
군집 1	0.28573	0.3764	0.6236	0.4162	0.5838
군집 2	0.25924	0.4066	0.5934	0.6741	0.3259
군집 3	0.20692	0.6795	0.3205	0.7523	0.2477
군집 4	0.19710	0.4706	0.5294	0.9922	0.0078
군집 5	0.05101	0.1839	0.8161	0.0334	0.9666

군집	전체	sex	marital status
군집 1	0.28573		
군집 2	0.25924		
군집 3	0.20692		
군집 4	0.19710		
군집 5	0.05101		

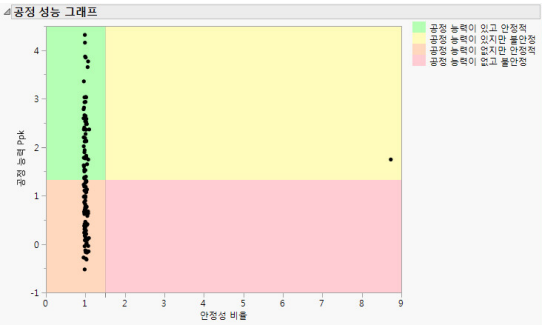
잠재 계층 분석
분석 > 군집화 > 잠재 계층 분석



예측 변수 선별

예측 변수	기여도	부분	순위
ink pct	20.8695	0.1242	1
solvent pct	17.7248	0.1055	2
varnish pct	17.2234	0.1025	3
press	12.1862	0.0725	4
press speed	11.4749	0.0683	5
roller durometer	9.8372	0.0585	6
press type	9.7274	0.0579	7
ESA Voltage	7.5421	0.0449	8
unit number	6.8950	0.0410	9
ESA Amperage	4.7195	0.0281	10
viscosity	4.4479	0.0265	11
blade pressure	4.2333	0.0252	12
humidity	3.6079	0.0215	13
grain screened	3.5390	0.0211	14
hardener	3.1871	0.0190	15
ink type	3.1542	0.0188	16
ink temperature	3.1205	0.0186	17
proof cut	3.0477	0.0181	18
paper mill location	2.8635	0.0170	19
type on cylinder	2.6002	0.0155	20
anode space ratio	2.4616	0.0146	21
blade mfg	2.4554	0.0146	22
current density	1.7377	0.0103	23
wax	1.6964	0.0101	24
cylinder size	1.6323	0.0097	25
caliper	1.3460	0.0080	26
paper type	1.2208	0.0073	27
solvent type	1.1994	0.0071	28
roughness	0.8502	0.0051	29
plating tank	0.6632	0.0039	30
proof on ctd ink	0.6108	0.0036	31
chrome content	0.0868	0.0005	32
direct steam	0.0778	0.0005	33

예측 변수 선별
분석 > 선별 > 예측 변수 선별

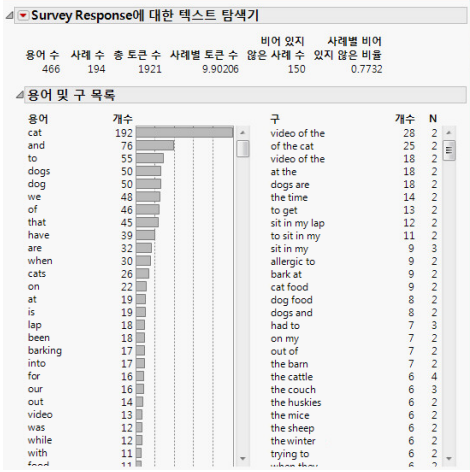


공정 성능 그래프

공정 성능 Ppk
안정성 비율

공정 성능이 있고 안정적
공정 성능이 있지만 불안정적
공정 성능이 없고 불안정

공정 변수 선별
분석 > 선별 > 공정 변수 선별



Survey Response에 대한 텍스트 탐색기

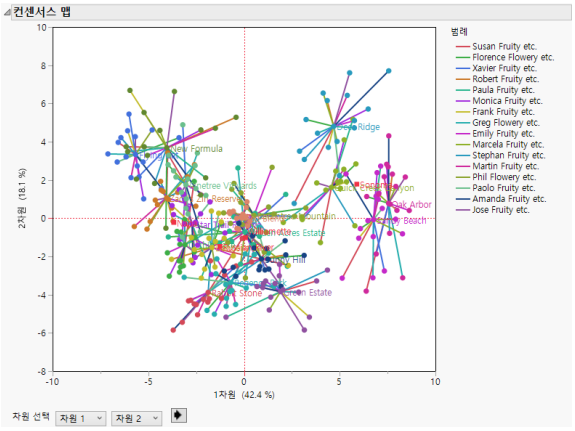
용어 수	사례 수	총 토큰 수	사례별 토큰 수	비어 있지 않은 사례 수	사례별 비어 있지 않은 비율
466	194	1921	9.90206	150	0.7732

용어 및 구 목록

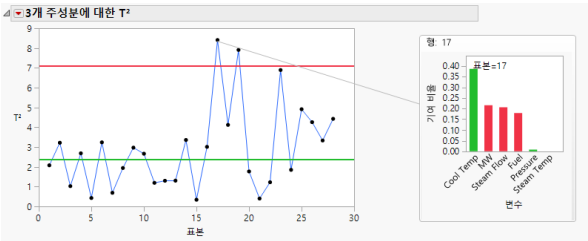
용어	개수
cat	192
and	76
to	55
dogs	50
dog	50
we	48
of	46
that	45
have	39
are	32
when	30
cats	26
on	22
at	19
is	19
lap	18
been	18
barking	17
into	17
for	16
our	16
out	14
video	13
was	12
while	12
with	11

구	개수	N
video of the	28	2
of the cat	25	2
video of the	18	2
at the	18	2
dogs are	18	2
the time	14	2
to get	13	2
sit in my lap	12	2
to sit in my	11	2
sit in my	9	3
allergic to	9	2
bark at	9	2
cat food	9	2
dog food	8	2
dogs and	8	2
had to	7	3
on my	7	2
out of	7	2
the barn	7	2
the cattle	6	4
the couch	6	3
the huskies	6	2
the mice	6	2
the sheep	6	2
the winter	6	2
trying to	6	2

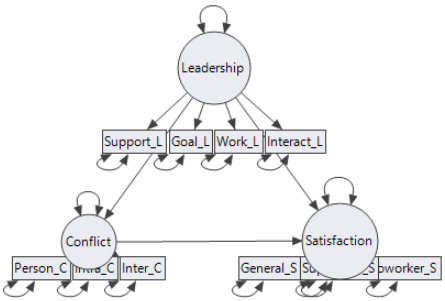
텍스트 탐색기
분석 > 텍스트 탐색기



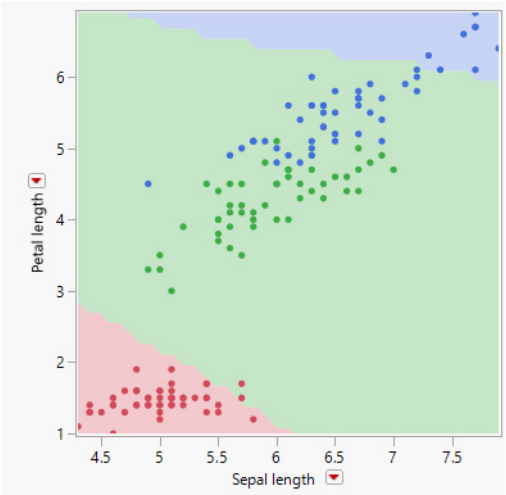
다중 요인 분석
분석 > 소비자 조사 > 다중 요인 분석



모형 기반 다변량 관리도
분석 > 품질 및 공정 > 모형 기반 다변량 관리도



구조 방정식 모형
분석 > 다변량 방법 > 구조 방정식 모형



서포트 벡터 머신
분석 > 예측 모형 > 서포트 벡터 머신

1 장

JMP 알아보기

설명서 및 추가 리소스

설명서 서식 규칙, 각 JMP 문서에 대한 설명, 도움말 시스템, 기타 지원 제공 위치 등 JMP 설명서에 대해 알아봅니다.

목차

JMP 설명서에 사용되는 서식 규칙 3
JMP 도움말 4
JMP 설명서 라이브러리 4
JMP 학습을 위한 추가 리소스 10
JMP 자습서 10
샘플 데이터 테이블 10
통계 및 JSL 용어에 대해 알아보기 11
JMP 팁 및 힌트 알아보기 11
JMP 툴팁 11
JMP 사용자 커뮤니티 11
무료 온라인 Statistical Thinking 교육 과정 12
새로운 사용자를 위한 JMP 입문 키트 12
Statistics Knowledge 포털 12
JMP 교육 과정 12
사용자가 작성한 JMP 설명서 12
JMP 시작하기 창 12
JMP 기술 지원 13

JMP 설명서에 사용되는 서식 규칙

설명서 자료가 가리키는 화면 정보를 쉽게 알아볼 수 있도록 다음과 같은 서식 규칙이 사용됩니다 .

- 샘플 데이터 테이블 이름 , 열 이름 , 경로 이름 , 파일 이름 , 파일 확장자 및 폴더는 Helvetica(또는 sans-serif online) 글꼴로 표시됩니다 .
- 코드는 `Lucida Sans Typewriter`(또는 monospace online) 글꼴로 표시됩니다 .
- 코드 출력은 Lucida Sans Typewriter 기울임꼴 (또는 monospace italic online) 글꼴로 표시되고 앞의 코드보다 더 많은 들여쓰기가 적용됩니다 .
- **Helvetica bold**(또는 bold sans-serif online) 서식은 작업을 수행하기 위해 사용자가 선택하는 다음과 같은 항목을 나타냅니다 .
  - 버튼
  - 체크박스
  - 명령
  - 선택 가능한 목록 이름
  - 메뉴
  - 옵션
  - 탭 이름
  - 텍스트 상자
- 다음 항목은 기울임꼴로 표시됩니다 .
  - 중요하거나 JMP 와 관련된 정의가 있는 단어 또는 구
  - 설명서 제목
  - 변수
- JMP Pro 에만 해당되는 기능에는 JMP Pro 아이콘 JMP PRO 이 표시됩니다 . JMP Pro 기능의 개요는 https://www.jmp.com/software/pro 에서 확인할 수 있습니다 .

참고 : 참고 섹션에는 특수한 정보와 제한 사항이 표시됩니다 .

팁 : 팁 섹션에는 유용한 정보가 표시됩니다 .

JMP 도움말

" 도움말 " 메뉴의 "JMP 도움말 " 에서는 JMP 기능 , 통계적 방법 및 JSL(JMP Scripting Language) 에 대한 정보를 검색할 수 있습니다 . 다음과 같은 몇 가지 방법으로 JMP 도움말을 열 수 있습니다 .

- Windows 에서 **도움말 > JMP 도움말**을 선택하여 JMP 도움말을 검색하고 봅니다 .
- Windows 에서 F1 키를 눌러 기본 브라우저로 도움말 시스템을 엽니다 .
- 데이터 테이블 또는 보고서 창의 특정 부분에 대한 도움말을 확인합니다 . **도구** 메뉴에서 도움말 도구 ? 를 선택한 후 데이터 테이블 또는 보고서 창의 아무 곳이나 클릭하면 해당 영역에 대한 도움말이 표시됩니다 .
- JMP 창 내에서 **도움말** 버튼을 클릭합니다 .

참고 : JMP 도움말은 인터넷이 연결되어 있어야 사용할 수 있습니다. 인터넷이 연결되어 있지 않으면 **도움말 > JMP 설명서 라이브러리**를 선택하여 단일 PDF 파일에서 모든 설명서를 검색할 수 있습니다 . 자세한 내용은 "JMP 설명서 라이브러리 "(4 페이지) 에서 확인하십시오 .

JMP 설명서 라이브러리

도움말 시스템 콘텐츠는 JMP 설명서 라이브러리라는 단일 PDF 파일로도 제공됩니다. 이 파일을 열려면 **도움말 > JMP 설명서 라이브러리**를 선택합니다. JMP 라이브러리에 포함된 각 문서의 개별 PDF 파일을 검색하려면 https://www.jmp.com/documentation 에서 파일을 다운로드하십시오.

다음 표에서는 JMP 라이브러리에 포함된 각 문서의 용도와 내용을 설명합니다 .

문서 제목	문서 용도	문서 내용
JMP 살펴보기	JMP에 익숙하지 않다면 이 설명서부터 시작하십시오.	JMP에 대해 소개하고 데이터 생성 및 분석을 시작하기 위한 정보를 제공합니다. 또한 결과를 공유하는 방법도 알아볼 수 있습니다.
JMP 사용	JMP 데이터 테이블과 기본적인 작업을 수행하는 방법에 대해 알아봅니다.	데이터 가져오기, 열 특성 수정, 데이터 정렬, SAS 연결 등을 비롯하여 JMP 의 모든 영역에 걸친 일반적인 JMP 개념 및 기능을 다룹니다.

문서 제목	문서 용도	문서 내용
기본 분석	이 문서를 사용하여 기본적인 분석을 수행합니다.	"분석" 메뉴의 다음 플랫폼에 대해 설명합니다. • 분포 • X 로 Y 적합 • 테이블 생성 • 텍스트 탐색기 "분석" > "X 로 Y 적합"을 통해 이변량 분석, 일원 ANOVA 및 분할 분석을 수행하는 방법을 다룹니다. 붓스트랩을 사용하여 표집 분포에 근사한 값을 산출하는 방법과 시뮬레이션 플랫폼을 사용하여 모수적 재표집을 수행하는 방법도 포함되어 있습니다.
Essential Graphing	데이터에 이상적인 그래프를 찾습니다.	"그래프" 메뉴의 다음 플랫폼에 대해 설명합니다. • 그래프 빌더 • 3D 산점도 • 등고선 그림 • 버블 그림 • 평행 그림 • 셀 그림 • 산점도 행렬 • 삼원 그림 • 트리맵 • 차트 • 중첩 그림 이 설명서에서는 배경 맵 및 사용자 맵을 생성하는 방법도 다룹니다.
Profilers	반응 표면의 횡단면을 볼 수 있게 해주는 대화식 프로파일링 도구의 사용 방법을 알아봅니다.	"그래프" 메뉴에 나열된 모든 프로파일러를 다룹니다. 랜덤 입력을 사용한 시뮬레이션 실행과 함께 잡음 요인 분석이 포함됩니다.

문서 제목	문서 용도	문서 내용
실험 설계 가이드	실험 설계 방법을 알아보고 적절한 표본 크기를 결정합니다.	"DOE" 메뉴의 모든 항목을 다룹니다.
Fitting Linear Models	모형 적합 플랫폼과 이 플랫폼의 다양한 분석법에 대해 알아봅니다.	"분석" 메뉴의 모형 적합 플랫폼에서 사용할 수 있는 다음 분석법에 대해 설명합니다. • 표준 최소 제곱 • 단계별 • 일반화 회귀 • 혼합 모형 • MANOVA • 로그 선형 분산 • 명목형 로지스틱 • 순서형 로지스틱 • 일반화 선형 모형

문서 제목	문서 용도	문서 내용
Predictive and Specialized Modeling	추가 모델링 기법에 대해 알아봅니다.	"분석" > "예측 모델링" 메뉴의 다음 플랫폼에 대해 설명합니다. • 신경망 • 파티션 • 붓스트랩 포레스트 • 부스티드 트리 • K 최근접 이웃 • Naive Bayes • 서포트 벡터 머신 • 모형 비교 • 모형 선별 • 검증 열 생성 • 계산식 저장소 "분석" > "전문 모델링" 메뉴의 다음 플랫폼에 대해 설명합니다. • 곡선 적합 • 비선형 • 함수 데이터 탐색기 • 가우스 과정 • 시계열 • 매칭 쌍 "분석" > "선별" 메뉴의 다음 플랫폼에 대해 설명합니다. • 모델링 유틸리티 • 반응 변수 선별 • 공정 변수 선별 • 예측 변수 선별 • 연관성 분석 • 공정 기록 탐색기

문서 제목	문서 용도	문서 내용
Multivariate Methods	몇 개의 변수를 동시에 분석하기 위한 기법을 알아봅니다.	"분석" > "다변량 방법" 메뉴의 다음 플랫폼에 대해 설명합니다. • 다변량 • 주성분 • 판별 • 부분 최소 제곱 • 다중 대응 분석 • 구조 방정식 모형 • 요인 분석 • 다차원 척도법 • 항목 분석 "분석" > "군집화" 메뉴의 다음 플랫폼에 대해 설명합니다. • 계층적 군집화 • K 평균 군집화 • 정규 혼합 • 잠재 계층 분석 • 변수 군집화
Quality and Process Methods	공정 평가 및 개선을 위한 도구에 대해 알아봅니다.	"분석" > "품질 및 공정" 메뉴의 다음 플랫폼에 대해 설명합니다. • 관리도 빌더 및 개별 관리도 • 측정 시스템 분석 • 계량형 / 계수형 게이지 차트 • 공정 능력 • 모형 기반 다변량 관리도 • 파레토도 • 다이어그램 • 규격 한계 관리

문서 제목	문서 용도	문서 내용
Reliability and Survival Methods	제품 또는 시스템의 신뢰도 평가 및 향상 방법과 사람 및 제품의 생존 데이터 분석 방법을 알아봅니다.	"분석" > "신뢰성 및 생존" 메뉴의 다음 플랫폼에 대해 설명합니다. • 수명 분포 • 수명 분포 적합 • 누적 손상 • 재발 분석 • 열화 • 파괴 열화 • 신뢰도 예측 • 신뢰도 성장 • 신뢰도 블록 다이어그램 • 수리 가능 시스템 시뮬레이션 • 생존 • 모수 생존 모형 적합 • 비례 위험 모형 적합
Consumer Research	소비자 선호도를 연구하고 해당 정보를 사용하여 보다 나은 제품 및 서비스를 개발하기 위한 방법을 알아봅니다.	"분석" > "소비자 조사" 메뉴의 다음 플랫폼에 대해 설명합니다. • 범주형 • 선택 • 최대차이 • Uplift • 다중 요인 분석
Scripting Guide	강력한 JSL(JMP 스크립트 언어)의 활용 방법을 알아봅니다.	스크립트 작성/디버깅, 데이터 테이블 조작, 표시 상자 생성 및 JMP 응용 프로그램 생성 등의 다양한 주제를 다룹니다.
JSL Syntax Reference	다양한 JSL 함수 및 인수, 그리고 개체 및 표시 상자로 보내는 메시지에 대해 알아봅니다.	JSL 명령의 구문, 예제 및 참고 사항이 포함되어 있습니다.

JMP 학습을 위한 추가 리소스

JMP 도움말 외에도 다음 리소스를 사용하여 JMP에 대해 배울 수 있습니다.

- "JMP 자습서 "
- " 샘플 데이터 테이블 "
- " 통계 및 JSL 용어에 대해 알아보기 "
- "JMP 팁 및 힌트 알아보기 "
- "JMP 툴팁 "
- "JMP 사용자 커뮤니티 "
- " 무료 온라인 Statistical Thinking 교육 과정 "
- " 새로운 사용자를 위한 JMP 입문 키트 "
- "Statistics Knowledge 포털 "
- "JMP 교육 과정 "
- " 사용자가 작성한 JMP 설명서 "
- "JMP 시작하기 창 "

JMP 자습서

도움말 > 자습서를 선택하여 JMP 자습서에 액세스할 수 있습니다. **자습서** 메뉴의 첫 번째 항목은 **자습서 디렉터리**입니다. 이 항목을 클릭하면 모든 자습서가 범주별로 그룹화되어 있는 새 창이 열립니다.

JMP에 익숙하지 않다면 **초보자 자습서**부터 시작하십시오. 이 자습서에서는 JMP 인터페이스를 단계별로 안내하며 JMP를 사용하는 데 필요한 기본 사항을 설명합니다.

나머지 자습서는 실험 설계, 표본 평균과 상수의 비교 같은 JMP의 특정 측면을 이해하는 데 유용합니다.

샘플 데이터 테이블

JMP 설명서 모음에 포함된 모든 예에서는 샘플 데이터를 사용합니다. 샘플 데이터 디렉터리를 열려면 **도움말 > 샘플 데이터 라이브러리**를 선택하십시오.

샘플 데이터 테이블의 사전순 목록을 보거나 범주별로 샘플 데이터를 보려면 **도움말 > 샘플 데이터**를 선택하십시오.

샘플 데이터 테이블은 다음 디렉터리에 설치되어 있습니다 .

Windows: C:\Program Files\SAS\JMP\16\Samples\Data

macOS: \Library\Application Support\JMP\16\Samples\Data

JMP Pro 의 경우에는 JMP 디렉터리가 아니라 JMPPRO 디렉터리에 샘플 데이터가 설치되어 있습니다.

샘플 데이터를 사용한 예를 보려면 **도움말 > 샘플 데이터**를 선택하고 교육 자료 섹션으로 이동하십시오. 교육 자료에 대한 자세한 내용은 https://jmp.com/tools 에서 확인하십시오.

통계 및 JSL 용어에 대해 알아보기

통계 용어에 대한 도움말을 보려면 " 도움말 > 통계 분석 인덱스 " 를 선택합니다 . JSL 스크립트 및 예제에 대한 도움말을 보려면 **도움말 > 스크립트 인덱스**를 선택합니다 .

통계 분석 인덱스 통계 용어에 대한 정의를 제공합니다 .

스크립트 인덱스 JSL 함수 , 개체 및 표시 상자에 대한 정보를 검색할 수 있습니다 . " 스크립트 인덱스 " 에서는 샘플 스크립트를 편집 및 실행하고 명령에 대한 도움말을 볼 수도 있습니다 .

JMP 팁 및 힌트 알아보기

JMP 를 처음 시작할 때는 " 오늘의 유익한 정보" 창이 표시됩니다. 이 창에서는 JMP 를 사용하기 위한 팁을 제공합니다.

"오늘의 유익한 정보" 기능을 해제하려면 **시작할 때 정보 표시** 체크박스를 선택 해제하십시오. 이 창을 다시 보려면 **도움말 > 오늘의 유익한 정보**를 선택하십시오. 또는 " 환경 설정 " 창을 사용하여 이 기능을 해제할 수도 있습니다.

JMP 툴팁

JMP 에서 다음과 같은 항목을 커서로 가리키면 설명 툴팁 (또는 가리키기 라벨) 이 제공됩니다 .

- 메뉴 또는 도구 모음 옵션
- 그래프의 라벨
- 보고서 창의 텍스트 결과 (커서를 원 모양으로 움직이면 표시됨)
- 홈 창의 파일 또는 창
- 스크립트 편집기의 코드

팁 : Windows 의 경우 JMP 환경 설정에서 툴팁을 숨길 수 있습니다 . **파일 > 환경 설정 > 일반**을 선택한 후 **메뉴 팁 표시**를 선택 취소합니다 . 이 옵션은 macOS 에서는 사용할 수 없습니다 .

JMP 사용자 커뮤니티

JMP 사용자 커뮤니티에서는 JMP 에 대해 알아보고 다른 JMP 사용자와 교류하는 데 도움이 되는 다양한 옵션을 제공합니다. 한 페이지 분량의 가이드, 자습서 및 데모로 구성된 학습 라이브러리부터 시작하는 것이 좋습니다. 다양한 JMP 교육 과정에 등록하여 학습을 계속할 수도 있습니다.

그 밖에도 토론 포럼, 샘플 데이터 및 스크립트 파일 교환, 웹 캐스트 및 소셜 네트워킹 그룹을 비롯한 리소스가 있습니다.

웹 사이트의 JMP 리소스에 액세스하려면 **도움말 > JMP 사용자 커뮤니티**를 선택하거나 https://community.jmp.com 을 방문하십시오.

무료 온라인 Statistical Thinking 교육 과정

이 무료 온라인 교육 과정에서는 탐색적 데이터 분석, 품질 관리 방법, 상관 및 회귀 등의 항목에 대한 실용적인 통계적 기술을 배울 수 있습니다. 이 교육 과정은 짧은 비디오와 데모, 연습 등으로 구성되어 있습니다. 자세한 내용은 https://www.jmp.com/statisticalthinking 에서 확인하십시오.

새로운 사용자를 위한 JMP 입문 키트

새로운 사용자를 위한 입문 키트는 JMP 의 기본 사항을 빨리 익힐 수 있도록 돕기 위한 것입니다. 30 개의 짧은 데모 비디오 및 작업을 마치면 좀더 편하게 소프트웨어 사용 방법을 익히고 세계 최대 규모의 JMP 사용자 온라인 커뮤니티와 연결할 수 있습니다. 자세한 내용은 https://www.jmp.com/welcome 에서 확인하십시오.

Statistics Knowledge 포털

Statistics Knowledge 포털에서는 방문자가 확실한 기초를 토대로 통계적 기술을 쌓을 수 있도록 간략한 통계 설명과 함께 명확한 예시 및 그래픽을 제공합니다. 자세한 내용은 https://www.jmp.com/skp 에서 확인하십시오.

JMP 교육 과정

SAS 에서는 숙련된 JMP 전문가 팀의 주도로 다양한 주제에 대한 교육 과정을 제공합니다. 공개 교육, 라이브 웹 교육, 현장 교육 등이 제공되며, 온라인 e-learning 구독을 선택하여 편리한 시간에 학습할 수도 있습니다. 자세한 내용은 https://www.jmp.com/training 에서 확인하십시오.

사용자가 작성한 JMP 설명서

JMP 웹 사이트에서는 JMP 사용자가 작성한 추가 JMP 사용 설명서가 제공됩니다 . 자세한 내용은 https://www.jmp.com/books 에서 확인하십시오 .

JMP 시작하기 창

JMP 또는 데이터 분석에 익숙하지 않다면 먼저 "JMP 시작하기 " 창을 살펴보십시오 . 이 창에는 옵션이 범주별로 설명되어 있으며 버튼을 클릭하여 옵션을 시작할 수 있습니다 . "JMP 시작하기

" 창에는 " 분석 ", " 그래프 ", " 테이블 " 및 " 파일 " 메뉴에 있는 다양한 옵션이 포함됩니다 . 또한 이 창에는 JMP Pro 의 기능 및 플랫폼도 나열됩니다 .

- "JMP 시작하기 " 창을 열려면 **보기** (macOS 의 경우 **창**) **> JMP 시작하기**를 선택합니다 .
- Windows 에서 JMP 를 열 때 자동으로 "JMP 시작하기 " 를 표시하려면 **파일 > 환경 설정 > 일반**을 선택한 후 " 초기 JMP 창 " 목록에서 **JMP 시작하기**를 선택합니다 . macOS 에서는 **JMP > 환경 설정 > 초기 JMP 시작하기 창**을 선택합니다 .

JMP 기술 지원

JMP 기술 지원은 통계학자 또는 SAS 및 JMP 의 교육을 받은 엔지니어가 제공하며 , 이들 중 상당수는 통계 또는 기타 기술 분야의 석사 학위를 갖고 있습니다 .

기술 지원 전화 번호를 포함한 많은 기술 지원 옵션이 https://www.jmp.com/support 에서 제공됩니다.

2 장

JMP 소개

기본 개념

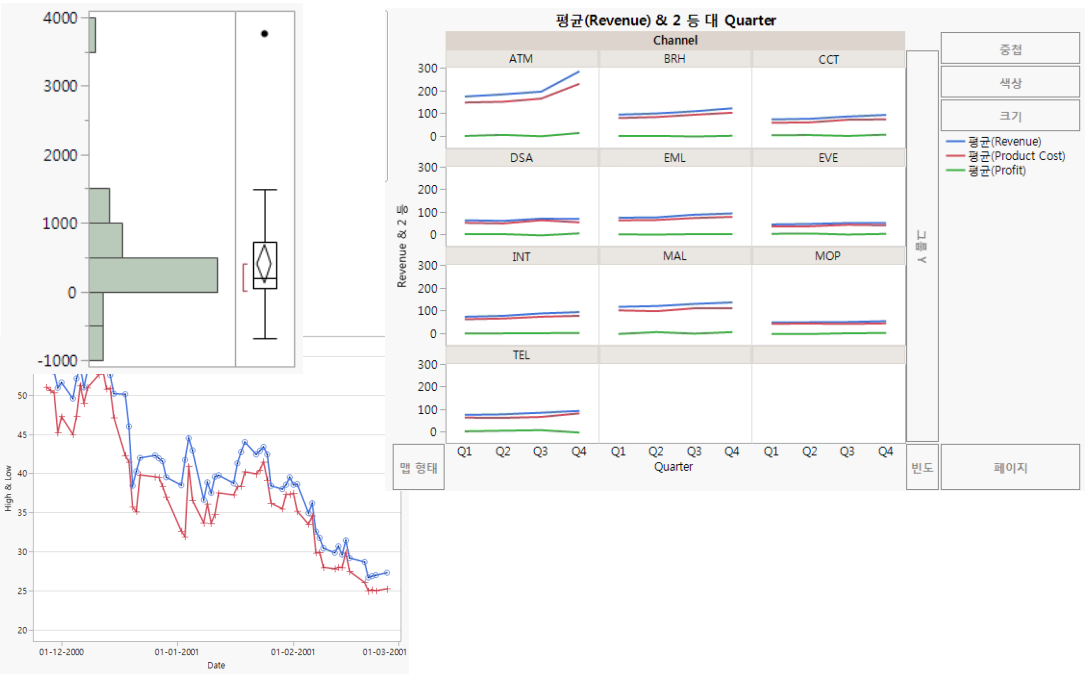
JMP(점프라고 발음)는 강력한 대화식 데이터 시각화 및 통계 분석 도구입니다. JMP에서 데이터 테이블, 그래프, 차트 및 보고서를 기반으로 분석을 수행하고 데이터와 상호 작용하면서 데이터에 대해 자세히 알아볼 수 있습니다.

JMP를 통해 연구자는 광범위한 통계 분석을 수행하고 모형을 생성할 수 있습니다. JMP는 데이터의 경향과 패턴을 신속하게 파악하려는 비즈니스 분석가에게도 유용합니다. JMP를 사용하면 통계 전문가가 아니라도 데이터에서 정보를 얻을 수 있습니다.

예를 들어 JMP를 사용하여 다음을 수행할 수 있습니다.

- 대화식 그래프 및 차트를 생성하여 데이터 탐색 및 관계 발견
- 한 번에 여러 변수에서 변동 패턴 발견
- 많은 양의 데이터 탐색 및 요약
- 미래를 예측할 수 있는 강력한 통계 모형 개발

그림 2.1 JMP 보고서의 예



목차

JMP 필수 개념 47
JMP 시작하기 47
JMP 시작 48
샘플 데이터 사용 49
데이터 테이블 이해 51
JMP 워크플로우 이해 52
1단계: JMP 플랫폼 시작 및 결과 보기 53
2단계: JMP 보고서에서 상자 그림 제거 54
3단계: 추가 JMP 출력 요청 55
4단계: JMP 플랫폼 결과에서의 상호 작용 56
JMP 와 Excel 의 차이점 57
데이터 테이블의 구조 57
JMP의 계산식 58
JMP 분석 및 그래프 생성 59

JMP 필수 개념

JMP 를 사용하기 전에 다음 개념을 잘 알아 두어야 합니다 .

- JMP 데이터 테이블을 사용하여 데이터 입력 , 확인 , 편집 및 조작
- **분석** , **그래프** 또는 **DOE** 메뉴에서 **플랫폼**선택 플랫폼에는 데이터를 분석하고 그래프로 작업하는 데 사용되는 대화식 창이 있습니다 .
- 플랫폼에서는 다음 창을 사용합니다 .
  - **시작 창** - 분석을 설정하고 실행합니다 .
  - **보고서 창** - 분석 결과를 보여 줍니다 .
- 보고서 창에는 일반적으로 다음 항목이 포함됩니다 .
  - 일부 유형의 그래프 (산점도 또는 차트 등)
  - **표시 아이콘** 을 사용하여 표시하거나 숨길 수 있는 **보고서**
  - **빨간색 삼각형 메뉴** 내에 있는 플랫폼 **옵션**

JMP 시작하기

JMP 의 일반적인 워크플로우는 간단합니다 .

1. 데이터를 JMP 로 가져옵니다 .
2. 플랫폼을 선택하고 시작 창을 완료합니다 .
3. 결과를 탐색하고 데이터가 제시하는 정보를 발견합니다 .

이 워크플로우는 "JMP 워크플로우 이해"(52 페이지)에서 자세히 설명합니다.

JMP에서는 일반적으로 그래프를 사용하여 개별 변수 및 변수 간의 관계를 시각화하면서 작업을 시작합니다 . 그래프를 사용하면 이러한 정보를 쉽게 확인하고 더욱 심도 깊은 질문을 파악할 수 있습니다 . 그런 다음 분석 플랫폼을 사용하여 문제를 보다 심층적으로 탐구하고 해결책을 찾습니다 .

- " 데이터 작업 " 장 (61 페이지) 에서는 데이터를 JMP 로 가져오는 방법을 보여 줍니다 .
- " 데이터 시각화 " 장 (91 페이지) 에서는 JMP 에서 제공하는 유용한 그래프 중 일부를 사용하여 데이터를 보다 자세히 분석하는 방법을 보여 줍니다 .
- " 데이터 분석 " 장 (127 페이지) 에서는 일부 분석 플랫폼을 사용하는 방법을 보여 줍니다 .
- " 전체를 보는 시각 " 장 (163 페이지) 에서는 여러 플랫폼에서 분포 , 패턴 및 유사한 값을 분석하는 방법을 보여 줍니다 .

각 장에서는 예를 통해 방법을 안내합니다. 이 장의 다음 섹션에서는 JMP에서 작업하는 데이터 테이블과 일반적인 개념에 대해 설명합니다.

JMP 시작

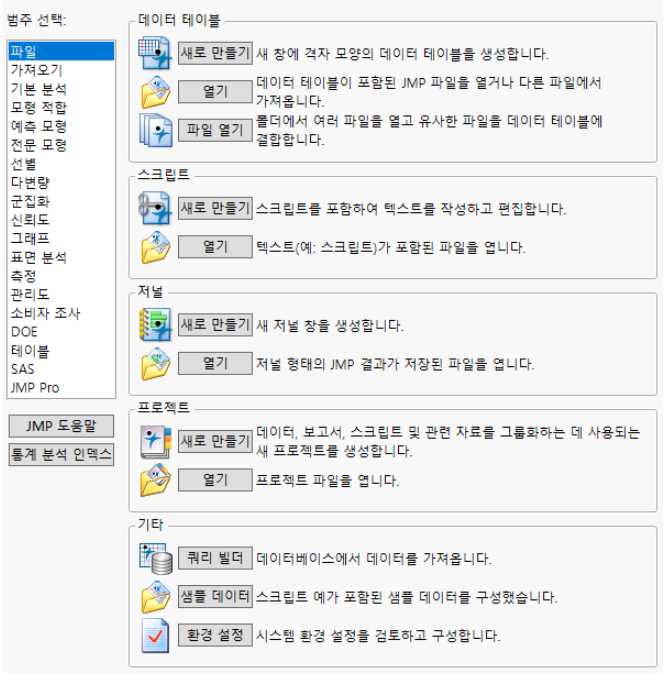
다음 두 가지 방법으로 JMP 를 시작하십시오 .

- 일반적으로 바탕 화면에 있는 JMP 아이콘을 두 번 클릭합니다 . 이렇게 하면 JMP 가 시작되지만 기존 JMP 파일은 열리지 않습니다 .
- 기존 JMP 파일을 두 번 클릭합니다 . 이렇게 하면 JMP 가 시작되고 파일이 열립니다 .

Windows 에서는 JMP 의 첫 화면에 "오늘의 유익한 정보" 창과 홈 창이 포함되고, macOS 에서는 "오늘의 유익한 정보", "JMP 시작하기" 및 홈 창이 처음에 나타납니다.

"JMP 시작하기" 창에서는 범주를 사용하여 작업 및 플랫폼을 분류합니다.

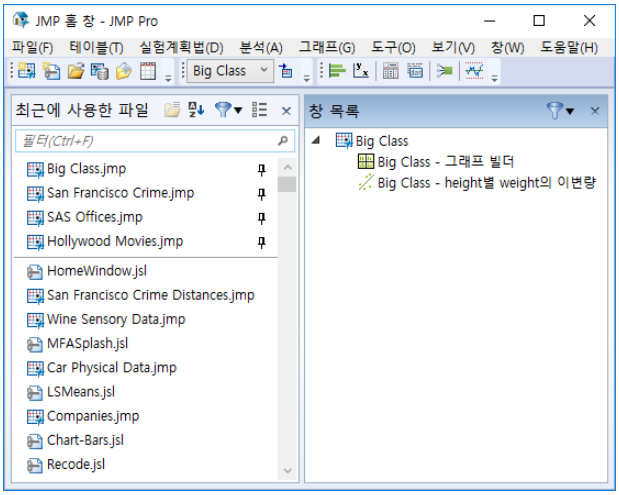
그림 2.2 JMP 시작하기



왼쪽에는 범주 목록이 있습니다. 범주를 클릭하면 해당 범주와 관련된 기능 및 명령을 볼 수 있습니다. "JMP 시작하기" 에는 JMP Pro 의 기능 및 플랫폼도 나열됩니다.

홈 창에서는 JMP 에서 파일을 구성하고 액세스할 수 있도록 도와줍니다.

그림 2.3 Windows 의 홈 창



Windows 에서 홈 창을 열려면 **보기 > 홈 창**을 선택합니다. macOS 에서는 **창 > JMP 홈**을 선택합니다. 홈 창에는 다음에 대한 링크가 있습니다.

- 현재 열려 있는 데이터 테이블 및 보고서 창
- 최근에 연 파일

홈 창에 대한 자세한 내용은 JMP 사용의 에서 확인하십시오.

거의 모든 JMP 창에는 메뉴 표시줄과 도구 모음이 있습니다. 대부분의 JMP 기능은 다음 세 가지 방법으로 찾을 수 있습니다.

- 메뉴 표시줄 사용
- 도구 모음 버튼 사용
- "JMP 시작하기 " 창의 버튼 사용

메뉴 표시줄 및 도구 모음

메뉴 표시줄과 도구 모음은 많은 창에서 숨겨져 있습니다. 표시하려면 창 제목 표시줄 아래에 있는 파란색 막대를 커서로 가리킵니다. "JMP 시작하기" 창, 홈 창 및 모든 데이터 테이블의 메뉴는 항상 표시됩니다.

샘플 데이터 사용

JMP 살펴보기와 기타 JMP 설명서의 예에서는 샘플 데이터 테이블을 사용합니다. Windows에서 샘플 데이터의 기본 위치는 다음과 같습니다.

C:/Program Files/SAS/JMP/16/Samples/Data

C:/Program Files/SAS/JMPPro/16/Samples/Data

샘플 데이터 인덱스는 데이터 테이블을 범주별로 그룹화합니다. 표시 아이콘을 클릭하여 해당 범주의 데이터 테이블 목록을 표시한 후 링크를 클릭하여 데이터 테이블을 엽니다.

macOS 샘플 데이터는 /Library/Application Support/JMP/16/Samples/Data에 설치되어 있습니다.

JMP 샘플 데이터 테이블 열기

1. **도움말** 메뉴에서 **샘플 데이터**를 선택합니다 .
2. **JMP 살펴보기에 사용된 데이터 테이블** 옆에 있는 표시 아이콘을 클릭하여 목록을 엽니다 .
3. 데이터 테이블의 이름을 클릭하여 JMP 살펴보기의 예에서 사용합니다 .

샘플 가져오기 데이터

다른 응용 프로그램의 파일을 사용하여 데이터를 JMP로 가져오는 방법을 알아봅니다.

Windows에서 샘플 가져오기 데이터의 기본 위치는 다음과 같습니다.

C:/Program Files/SAS/JMP/16/Samples/Import Data

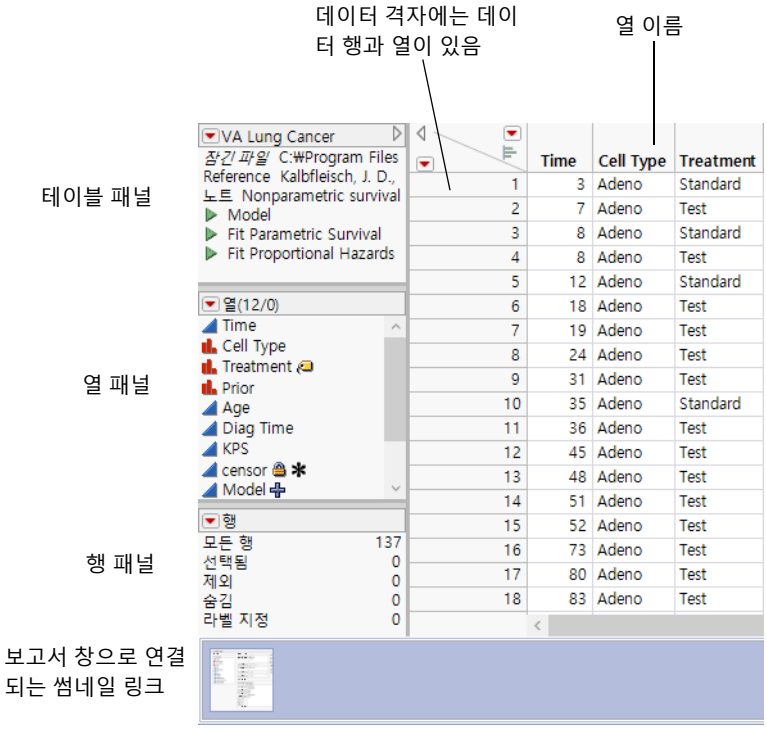
C:/Program Files/SAS/JMPPro/16/Samples/Import Data

데이터 테이블 이해

JMP 데이터 테이블은 행과 열로 구성된 데이터 모음입니다. 데이터 테이블에는 노트, 변수 및 스크립트와 같은 기타 정보가 포함될 수도 있습니다. 이러한 보충 항목은 다음 장에서 설명합니다.

"VA Lung Cancer" 데이터 테이블을 열고 여기에 설명된 데이터 테이블을 살펴보십시오.

그림 2.4 데이터 테이블



데이터 테이블에는 다음과 같은 부분이 있습니다.

데이터 격자 데이터 격자에는 행과 열로 정렬된 데이터가 있습니다. 일반적으로 데이터 격자의 각 행은 관측값이며 열 (변수라고도 함) 은 관측값에 대한 정보를 제공합니다. 그림 2.4 에서 각 행은 시험 대상에 해당하며 12 개의 정보 열이 있습니다. 데이터 격자에는 12 개 열이 모두 표시될 수 없지만 " 열 " 패널에는 모든 열이 표시됩니다. 각 시험 대상에 대한 정보로는 시간, 세포 유형, 치료법 등이 포함됩니다. 각 열에는 머리글 또는 이름이 있습니다. 이 이름은 테이블의 전체 행 수에 포함되지 않습니다.

테이블 패널 테이블 패널에는 테이블 변수 또는 테이블 스크립트가 포함될 수 있습니다. 그림 2.4 에는 자동으로 분석을 다시 생성할 수 있는 **Model** 이라는 저장된 스크립트가 하나 있습니다. 또한 이 테이블에는 데이터에 대한 정보가 들어 있는 "Notes" 라는 변수가 있습니다. 테이블 변수와 테이블 스크립트는 다음 장에서 설명합니다.

열 패널 " 열 " 패널에는 열의 총 수 , 열이 선택되었는지 여부 및 모든 열의 이름순 목록이 표시됩니다 . 괄호 안의 숫자 (12/0) 는 12 개의 열이 있고 선택된 열이 없음을 보여 줍니다 . 각 열 이름 왼쪽에 있는 아이콘은 해당 열의 모델링 유형을 표시합니다 . 모델링 유형에 대해서는 " 데이터 분석 " 장의 " 모델링 유형 이해 "(132 페이지) 에서 설명합니다 . 오른쪽 아이콘은 열에 할당된 속성을 표시합니다 . 이러한 아이콘에 대한 자세한 내용은 " 데이터 작업 " 장의 " 데이터 테이블에서 열 정보 보기 또는 변경 "(75 페이지) 에서 확인하십시오 .

행 패널 " 행 " 패널에는 데이터 테이블의 총 행 수와 선택된 행 , 제외된 행 , 숨겨진 행 또는 라벨이 지정된 행의 개수가 표시됩니다 . 그림 2.4 의 데이터 테이블에는 137 개의 행이 있습니다 .

보고서 창으로 연결되는 썸네일 링크 이 영역에는 데이터 테이블을 기반으로 하는 모든 보고서의 썸네일이 표시됩니다 . 썸네일을 커서로 가리키면 보고서 창의 더 큰 미리보기를 볼 수 있습니다 . 썸네일을 두 번 클릭하면 해당 보고서 창이 맨 앞에 표시됩니다 .

행 및 열 추가, 데이터 입력 및 데이터 편집을 포함한 데이터 격자에서의 상호 작용에 대해서는 " 데이터 작업 " 장 (61 페이지) 에서 설명합니다. 여러 데이터 테이블을 열면 각 테이블이 별도의 창에 나타납니다.

JMP 데이터 테이블과 Excel 스프레드시트의 차이점에 대한 자세한 내용은 "JMP 와 Excel 의 차이점 "(57 페이지) 에서 확인하십시오.

JMP 워크플로우 이해

데이터가 데이터 테이블에 있으면 그래프 또는 그림을 생성하고 분석을 수행할 수 있습니다. 모든 기능은 플랫폼에 있으며 플랫폼은 주로 **분석** 또는 **그래프** 메뉴에 있습니다. 이를 플랫폼이라고 하는 것은 해당 기능이 단순한 정적 결과만 산출하는 것이 아니기 때문입니다. 플랫폼 결과는 보고서 창에 표시되고 대개 대화식이며 데이터 테이블과 연결되고 서로 간에도 연결됩니다.

분석 및 **그래프** 메뉴의 플랫폼은 다양한 분석 기능과 데이터 탐색 도구를 제공합니다.

그래프 또는 분석을 생성하는 일반적인 단계는 다음과 같습니다 .

1. 데이터 테이블을 엽니다 .
2. " 그래프 " 또는 " 분석 " 메뉴에서 플랫폼을 선택합니다 .
3. 플랫폼 시작 창을 완료하여 분석을 설정합니다 .
4. **확인**을 클릭하여 그래프 및 통계 분석이 포함된 보고서 창을 생성합니다 .
5. 보고서 옵션을 사용하여 보고서를 사용자 정의합니다 .
6. 결과를 저장하고 , 내보내고 , 다른 사람과 공유합니다 .

이러한 개념에 대해서는 뒷부분 장에서 자세히 설명합니다.

다음 예에서는 네 단계로 간단한 분석을 수행하고 사용자 정의하는 방법을 보여 줍니다. 이 예에서는 Companies.jmp 파일의 샘플 데이터 테이블을 사용하여 Profits ($M) 변수의 기본 분석을 보여 줍니다.

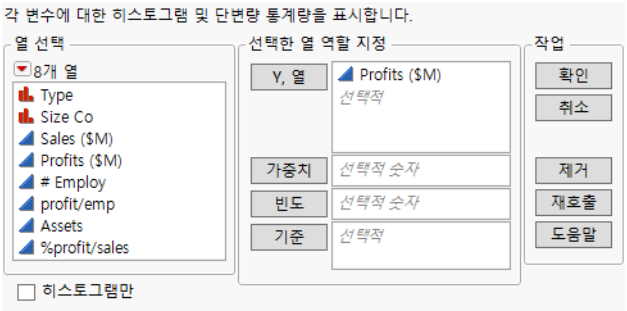
1 단계 : JMP 플랫폼 시작 및 결과 보기

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다 .
2. **분석 > 분포**를 선택하여 분포 시작 창을 엽니다 .
3. " 열 선택 " 상자에서 Profits ($M) 를 선택하고 **Y, 열** 버튼을 클릭합니다 .

   Profits ($M) 변수가 **Y, 열** 역할에 나타납니다 .

   변수를 할당하는 또 다른 방법은 " 열 선택 " 상자에서 열을 클릭하여 역할 상자로 드래그하는 것입니다 .

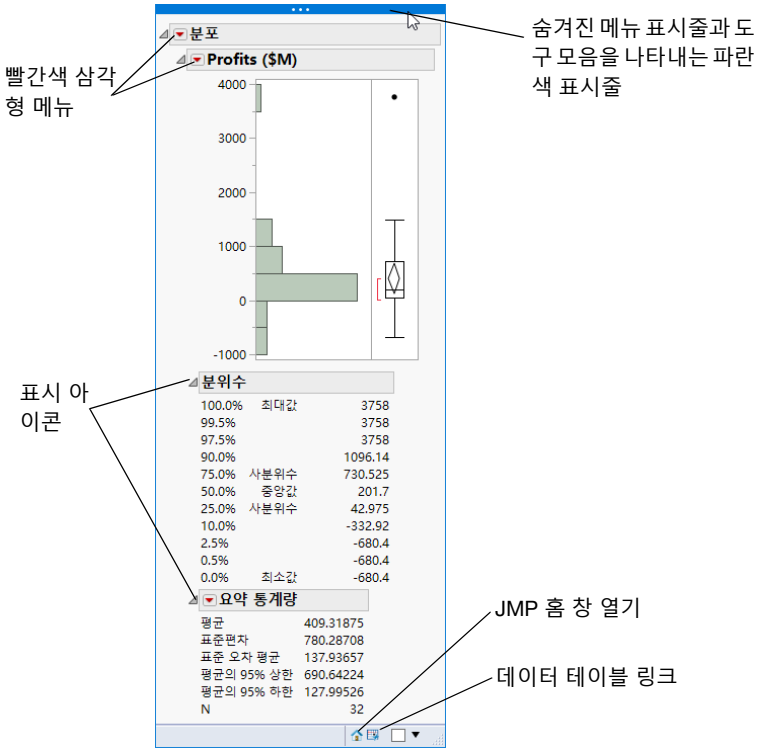
그림 2.5 Profits ($M) 할당



4. **확인**을 클릭합니다 .

   " 분포 " 보고서 창이 나타납니다 .

그림 2.6 Windows 의 분포 보고서 창



이 보고서 창에는 기본 그림 또는 그래프와 예비 분석 보고서가 포함됩니다. 결과는 개요 형식으로 표시되며 표시 아이콘을 클릭하여 보고서를 표시하거나 숨길 수 있습니다.

빨간색 삼각형 메뉴에는 언제든지 추가 그래프 및 분석을 요청할 수 있는 옵션과 명령이 포함되어 있습니다 .

- Windows에서 창 상단의 파란색 막대를 커서로 가리키면 메뉴 표시줄과 도구 모음이 표시됩니다 .
- Windows에서 오른쪽 하단에 있는 데이터 테이블 버튼을 클릭하면 이 보고서를 생성하는 데 사용된 데이터 테이블이 표시됩니다 . macOS 에서는 보고서 창 오른쪽 상단에 있는 **데이터 테이블 표시** 버튼을 클릭합니다 .
- Windows 에서 오른쪽 하단에 있는 **JMP 홈 창** 버튼을 클릭하면 홈 창이 표시됩니다 . macOS 에서는 **창 > JMP 홈**을 선택합니다 .

2 단계 : JMP 보고서에서 상자 그림 제거

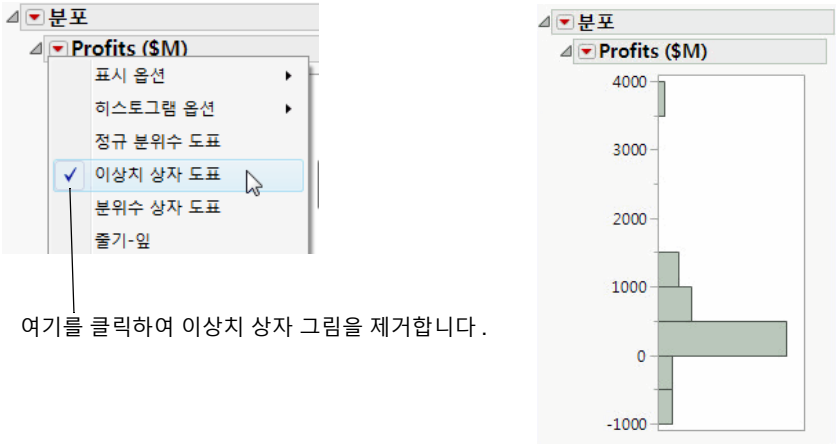
앞서 생성한 " 분포 " 보고서를 계속 사용합니다 .

1. Profits ($M) 옆의 빨간색 삼각형을 클릭하여 보고서 옵션 메뉴를 표시합니다 .

2. **이상치 상자 그림**을 선택 취소하여 이 옵션을 해제합니다 .

   이상치 상자 그림이 보고서 창에서 제거됩니다 .

그림 2.7 이상치 상자 그림 제거



3 단계 : 추가 JMP 출력 요청

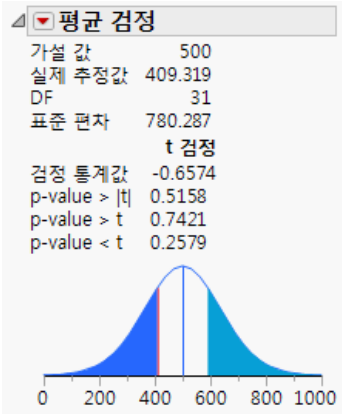
"2 단계 : JMP 보고서에서 상자 그림 제거 " 에서 생성한 같은 보고서 창을 계속 사용합니다 .

1. Profits ($M) 옆의 빨간색 삼각형을 클릭하고 **평균 검정**을 선택합니다 .

   " 평균 검정 " 창이 나타납니다 .
2. **가설 평균 지정** 상자에 "500" 을 입력합니다 .
3. **확인**을 클릭합니다 .

   평균에 대한 검정이 보고서 창에 추가됩니다 .

그림 2.8 평균 검정



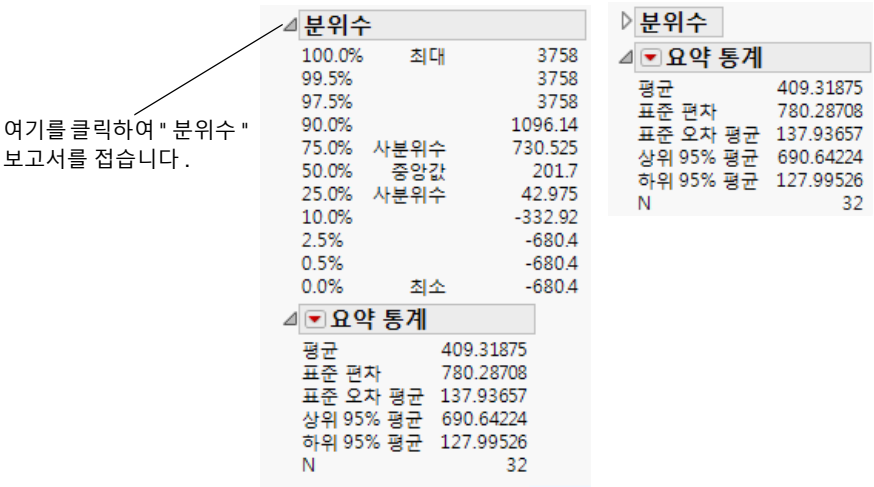
4 단계 : JMP 플랫폼 결과에서의 상호 작용

모든 플랫폼은 예를 들어 다음과 같은 대화식 결과를 생성합니다 .

- 보고서를 표시하거나 숨길 수 있습니다 .
- 용도에 맞게 그래프 및 통계 상세 정보를 추가하거나 제거할 수 있습니다 .
- 플랫폼 결과는 데이터 테이블과 연결되고 서로 간에도 연결됩니다 .

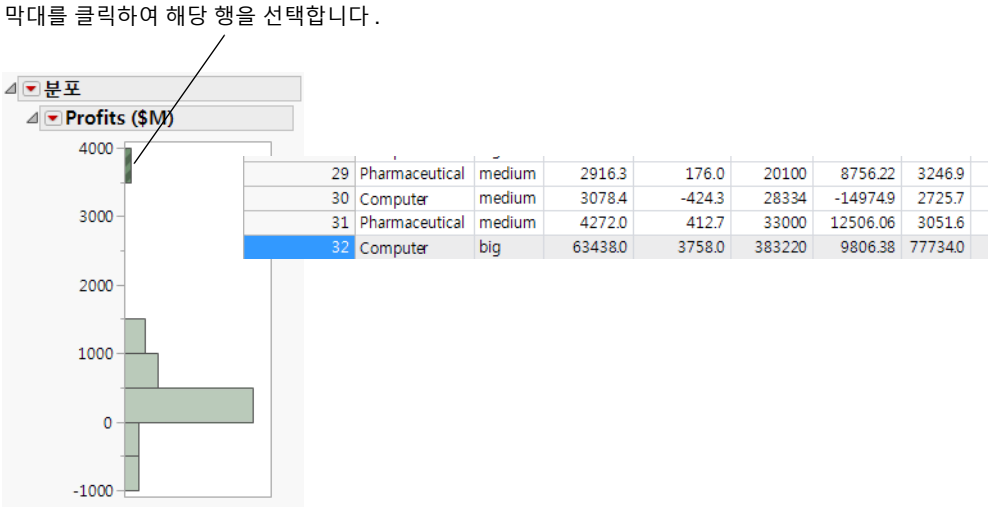
예를 들어 **분위수** 보고서를 닫으려면 **분위수** 옆의 표시 아이콘을 클릭합니다.

그림 2.9 분위수 보고서 닫기



플랫폼 결과는 데이터 테이블과 연결됩니다. 그림 2.10 의 히스토그램에서는 한 회사 그룹이 다른 회사보다 훨씬 높은 수익을 올리고 있음을 보여 줍니다. 이 그룹을 빠르게 식별하려면 해당 히스토그램 막대를 클릭합니다. 데이터 테이블에서 해당 행이 선택됩니다.

그림 2.10 플랫폼 결과와 데이터 테이블 간의 연결



이때 그룹에는 단 하나의 회사만 포함되며 해당 행 하나만 선택됩니다.

JMP 와 Excel 의 차이점

JMP는 데이터 테이블을 사용하는 통계 분석 프로그램입니다. Excel은 스프레드시트 응용 프로그램입니다. 데이터 테이블과 스프레드시트는 구조가 서로 다릅니다.

- " 데이터 테이블의 구조 "
- "JMP 의 계산식 "
- "JMP 분석 및 그래프 생성 "

데이터 테이블의 구조

데이터 테이블에는 고정된 행과 열이 있는 반면 스프레드시트는 셀 기반입니다. 스프레드시트에서는 모든 셀에 데이터, 머리글 또는 계산식을 포함할 수 있습니다. 데이터 테이블에서는 구조에 따라 분석용 데이터가 구성됩니다. 이 구조는 JMP 분석 및 그래프 플랫폼에 사용됩니다.

열 머리글 열 이름이 열 머리글입니다.

열 열에는 데이터가 포함되고 하나의 데이터 유형이 할당됩니다. 기본 열은 숫자 또는 문자 유형입니다. 열에 문자 및 숫자 데이터가 모두 포함되어 있을 때 전체 열의 데이터 유형은 문자이며 숫자는 문자 데이터로 처리됩니다. JMP 에는 이미지 등의 항목을 캡처하기 위한 특수한 열 유형도 있습니다. JMP 에서는 열의 데이터 유형에 따라 분석 옵션과 결과가 결정됩니다. 데이터 유형에 대한 자세한 내용은 " 데이터 분석 " 장의 " 모델링 유형 이해 "(132 페이지) 에서 확인하십시오.

행 행에는 관측값이 포함됩니다 . 행에 관측값이 없으면 해당 셀은 빈 채로 있습니다 . JMP 에서 점은 결측 숫자 값을 나타내고 공백은 결측 문자 값을 나타냅니다 .

JMP 데이터 테이블에 대한 자세한 내용은 "데이터 테이블 이해"(51 페이지)에서 확인하십시오. JMP 열 특성에 대한 자세한 내용은 JMP 사용의 에서 확인하십시오.

JMP 데이터 테이블은 Excel에서처럼 통합 문서 형식으로 구성할 수 없습니다. 각 JMP 데이터 테이블이 개별 파일이 되고 개별 창에 표시됩니다. 여러 테이블을 결합하는 방법은 JMP 사용의 에서 확인하십시오. JMP 테이블 및 출력을 구성하는 방법은 "작업 저장 및 공유" 장의 "스크립트 저장 및 실행"(191 페이지)에서 확인하십시오.

팁 : 단일 분석에 두 개 이상의 테이블에 있는 데이터를 사용하려면 가상 결합 기능을 사용하십시오 . 자세한 내용은 JMP 사용의 에서 확인하십시오 .

JMP 의 계산식

스프레드시트에서는 계산식이 단일 셀에 적용되며 통합 문서의 다른 탭에 있는 셀을 비롯하여 스프레드시트의 모든 셀에 있는 데이터를 활용할 수 있습니다. 데이터 테이블에서는 계산식이 특정 열 전체에 적용되며 데이터 테이블에 있는 다른 열의 데이터를 사용할 수 있습니다. 열의 각 행에는 해당 행의 데이터를 기반으로 동일한 계산이 적용됩니다.

예를 들어 그림 2.11 에 표시된 것처럼 단순 합계를 계산하는 데이터 테이블이 있을 수 있습니다. height + weight 열에는 계산식이 있습니다. 이 계산식은 데이터 테이블의 각 행에 대해 height 값과 weight 값을 더합니다.

그림 2.11 계산식 열이 있는 데이터 테이블

	name	age	sex	height	weight	height + weight
1	KATIE	12	F	59	95	154
2	LOUISE	12	F	61	123	184
3	JANE	12	F	55	74	129
4	JACLYN	12	F	66	145	211
5	LILLIE	12	F	52	64	116
6	TIM	12	M	60	84	144
7	JAMES	12	M	61	128	189

JMP 계산식에 대한 자세한 내용은 JMP 사용의 에서 확인하십시오.

팁 : 기본적인 열 요약 통계량이 필요하면 분포 플랫폼을 사용하십시오 . 자세한 내용은 기본 분석의 에서 확인하십시오 .

JMP 분석 및 그래프 생성

JMP에서는 데이터 분석 시 플랫폼을 사용합니다. 분석을 시작하려면 "분석" 메뉴로 이동합니다. 플랫폼 시작 창에서 분석할 변수를 선택하며, 분석 결과는 데이터 테이블 창이 아닌 별도의 보고서 창에 표시됩니다. 이는 분석 결과가 스프레드시트에 삽입되는 Excel과 다른 점입니다.

그래프 생성 옵션은 "그래프" 메뉴에 있습니다. 그래프 빌더를 사용하면 매우 간편하게 그래프 생성을 시작할 수 있습니다. 그래프 빌더에서 열을 드래그하여 놓고 빠르게 그래프를 생성하여 데이터를 탐색할 수 있습니다. 그래프 빌더에 대한 자세한 내용은 Essential Graphing의 에서 확인하십시오.

3 장

데이터 작업

그래프 생성 및 분석을 위한 데이터 준비

데이터를 그래프로 나타내거나 분석하려면 먼저 데이터 테이블에 적절한 형식으로 데이터가 준비되어야 합니다. 이 장에서는 다음을 포함하여 몇 가지 기본 데이터 관리 작업을 보여 줍니다.

- 새 데이터 테이블 생성
- 기존 데이터 테이블 열기
- 다른 응용 프로그램에서 JMP 로 데이터 가져오기
- 데이터 관리

그림 3.1 데이터 테이블의 예

Companies
잠긴 파일 C:₩Program File
노트 Selected Data on the

열(8/0)
Type
Size Co
Sales ($M)
Profits ($M)
Employ
profit/emp
Assets
%profit/sales

행
모든 행 32
선택됨 0
제외 0
숨김 0
라벨 지정 0

	Type	Size Co	Sales ($M)	Profits ($M)	# Employ	profit/emp	Assets	%profit/sales
1	Computer	small	855.1	31.0	7523	4120.70	615.2	3.63
2	Pharmaceutical	big	5453.5	859.8	40929	21007.11	4851.6	15.77
3	Computer	small	2153.7	153.0	8200	18658.54	2233.7	7.10
4	Pharmaceutical	big	6747.0	1102.2	50816	21690.02	5681.5	16.34
5	Computer	small	5284.0	454.0	12068	37620.15	2743.9	8.59
6	Pharmaceutical	big	9422.0	747.0	54100	13807.76	8497.0	7.93
7	Computer	small	2876.1	333.3	9500	35084.21	2090.4	11.59
8	Computer	small	709.3	41.4	5000	8280.00	468.1	5.84
9	Computer	small	2952.1	-680.4	18000	-378000	1860.7	-23.05
10	Computer	small	784.7	89.0	4708	18903.99	955.8	11.34
11	Computer	small	1324.3	-119.7	13740	-8711.79	1040.2	-9.04
12	Pharmaceutical	medium	4175.6	939.5	28200	33315.60	5848.0	22.50
13	Computer	big	11899.0	829.0	95000	8726.32	10075.0	6.97
14	Computer	small	873.6	79.5	8200	9695.12	808.0	9.10
15	Pharmaceutical	big	9844.0	1082.0	83100	13020.46	7919.0	10.99
16	Pharmaceutical	small	969.2	227.4	3418	66530.13	784.0	23.46
17	Pharmaceutical	medium	6698.4	1495.4	34400	43470.93	6756.7	22.32
18	Computer	big	5956.0	412.0	56000	7357.14	4500.0	6.92
19	Pharmaceutical	big	5903.7	681.1	42100	16178.15	8324.8	11.54
20	Computer	medium	2959.3	252.8	31404	8049.93	5611.1	8.54
21	Pharmaceutical	small	1198.3	86.5	8527	10144.25	1791.7	7.22

목차

JMP 로 데이터 가져오기 63
데이터 테이블에 데이터 복사 및 붙여넣기 63
데이터 테이블로 데이터 가져오기 63
데이터 테이블에 데이터 입력 66
Excel에서 JMP로 데이터 전송 68
데이터 테이블 작업 69
데이터 테이블에서 데이터 편집 70
데이터 테이블에서 값 선택, 선택 취소 및 찾기 72
데이터 테이블에서 열 정보 보기 또는 변경 75
계산식으로 값 계산의 예 77
보고서 데이터 필터링의 예 78
데이터 재구성의 예 80
요약 통계량 보기 80
부분집합 생성의 예 85
예: 데이터 테이블 결합 87
데이터 정렬 88

JMP 로 데이터 가져오기

JMP 에서 제공하는 다양한 방법을 사용하여 JMP 로 데이터를 가져올 수 있습니다.

- 다른 응용 프로그램의 데이터를 복사하여 붙여넣는 방법은 " 데이터 테이블에 데이터 복사 및 붙여넣기 "(63 페이지) 에서 확인하십시오 .
- 다른 응용 프로그램에서 데이터를 가져오는 방법은 " 데이터 테이블로 데이터 가져오기 "(63 페이지) 에서 확인하십시오 .
- 데이터를 직접 데이터 테이블에 입력하는 방법은 " 데이터 테이블에 데이터 입력 "(66 페이지) 에서 확인하십시오 .

데이터베이스에서 JMP 로 데이터를 가져올 수도 있습니다. 자세한 내용은 JMP 사용의 에서 확인하십시오.

이 장에서는 JMP 와 함께 설치되는 샘플 데이터 테이블과 샘플 가져오기 데이터를 사용합니다. 이러한 파일을 찾는 방법은 "JMP 소개" 장의 " 샘플 데이터 사용"(49 페이지) 에서 확인하십시오.

데이터 테이블에 데이터 복사 및 붙여넣기

Microsoft Excel 등의 다른 응용 프로그램이나 텍스트 파일에서 데이터를 복사하여 붙여넣는 방법으로 데이터를 JMP 로 이동할 수 있습니다 .

1. Microsoft Excel 에서 VA Lung Cancer.xlsx 파일을 엽니다 . 이 파일은 샘플 가져오기 데이터 폴더에 있습니다 .
2. 열 이름을 포함하여 모든 행과 열을 선택합니다 . 12 개의 열과 138 개의 행이 있습니다 .
3. 선택한 데이터를 복사합니다 .
4. JMP 에서 **파일 > 새로 만들기 > 데이터 테이블**을 선택하여 빈 테이블을 생성합니다 .
5. **편집 > 열 이름과 함께 붙여넣기**를 선택하여 데이터와 열 머리글을 붙여넣습니다 .

   JMP 에 붙여넣는 데이터에 열 이름이 없는 경우에는 **편집 > 붙여넣기**를 사용할 수 있습니다 .

데이터 테이블로 데이터 가져오기

Microsoft Excel, SAS 등의 다른 응용 프로그램이나 텍스트 파일에서 데이터를 가져오는 방법으로 데이터를 JMP 데이터 테이블로 이동할 수 있습니다 . 데이터를 가져오는 기본적인 단계는 다음과 같습니다 .

1. **파일 > 열기**를 선택합니다 .
2. 파일 위치로 이동합니다 .
3. 파일이 " 데이터 파일 열기 " 창에 표시되지 않으면 **파일 유형** 메뉴에서 올바른 파일 유형을 선택합니다 .
4. **열기**를 클릭합니다 .

예 : Microsoft Excel 파일 가져오기

1. **파일 > 열기**를 선택합니다 .
2. Samples/Import Data 폴더로 이동합니다 .
3. Team Results.xlsx 를 선택합니다 .

   데이터가 시작되는 행과 열에 유의하십시오 . 이 스프레드시트에는 두 개의 워크시트도 포함되어 있습니다 . 이 예에서는 "Ungrouped Team Results" 워크시트를 가져옵니다 .
4. **열기**를 클릭합니다 .

   스프레드시트가 "Excel 가져오기 마법사 " 에서 열리며 가져오기 옵션과 함께 데이터 미리보기가 나타납니다 .

   스프레드시트의 첫 번째 행에 있는 텍스트는 열 머리글입니다 . 하지만 스프레드시트의 3 행에 있는 텍스트를 열 머리글로 변환하려고 합니다 .
5. **열 머리글이 시작되는 행** 옆에 3 을 입력하고 Enter 키를 누릅니다 . 열 머리글이 데이터 미리보기에서 업데이트됩니다 . 데이터의 첫 번째 행에 대한 값이 4 로 업데이트됩니다 .
6. 이 워크시트에 대한 설정만 저장합니다 .
   - 창의 왼쪽 하단에서 **모든 워크시트에 사용**을 선택 취소합니다 .
   - 창의 오른쪽 상단에서 **Ungrouped Team Results** 를 선택합니다 .
7. **가져오기**를 클릭하여 스프레드시트를 지정한 대로 변환합니다 .

Excel 파일을 가져올 때 JMP 는 열 머리글이 있는지 여부와 열 이름이 1 행에 있는지 여부를 예측합니다 . 다음과 같은 경우에는 복사 및 붙여넣기 방법을 사용하는 것이 좋습니다 .

- 열 이름이 1 행 이외의 행에 있는 경우
- 파일에 열 이름이 없고 데이터가 1 행에서 시작하지 않는 경우
- 파일에 열 이름이 있고 데이터가 2 행에서 시작하지 않는 경우

Excel 파일 가져오기에 대한 자세한 내용은 JMP 사용의 "데이터 테이블에 데이터 복사 및 붙여넣기"(63 페이지) 및 에서 확인하십시오.

예 : 텍스트 파일 가져오기

텍스트 파일을 가져오는 한 가지 방법은 JMP에서 데이터 형식을 가정하고 데이터를 데이터 테이블에 배치하도록 하는 것입니다. 이 방법을 사용할 경우 사용자가 환경 설정에서 지정한 설정이 사용됩니다. 텍스트 가져오기 환경 설정을 지정하는 방법에 대한 자세한 내용은 JMP 사용의 에서 확인하십시오.

텍스트 파일을 가져오는 또 다른 방법은 텍스트 미리보기 창에서 데이터를 가져온 후 데이터 테이블이 표시되는 방식을 확인하고 조정하는 것입니다 . 다음 예에서는 텍스트 가져오기 미리보기 창을 사용하는 방법을 보여 줍니다 .

1. **파일 > 열기**를 선택합니다 .
2. Samples/Import Data 폴더로 이동합니다 .

3. Animals_line3.txt 를 선택합니다 .
4. " 열기 " 창 하단에서 **데이터 (미리보기 사용)** 를 선택합니다 .
5. **열기**를 클릭합니다 .

그림 3.2 초기 미리보기 창

이 텍스트 파일은 첫 번째 행에 제목이 있고 세 번째 행에 열 이름이 있으며 네 번째 행에서 데이터가 시작됩니다 . 이 파일을 JMP 에서 직접 열 때는 "Animals Data" 행이 첫 번째 열 이름이 되고 이후의 모든 열 이름과 데이터가 동기화되지 않습니다 . 미리보기 창을 사용하여 파일을 열기 전에 설정을 조정하고 조정 내용이 최종 데이터 테이블에 어떻게 영향을 미치는지 확인할 수 있습니다 .

6. **열 이름이 포함된 파일 . 행 번호** 필드에 3 을 입력합니다 .
7. **데이터가 시작되는 행** 필드에 4 를 입력합니다 .
8. **다음**을 클릭합니다 .

   두 번째 창에서는 가져오기에서 제외할 열을 지정하고 열의 데이터 모형을 변경할 수 있습니다 . 이 예에서는 기본 설정을 사용합니다 .
9. **가져오기**를 클릭합니다 .

새 테이블에는 species, subject, miles 및 season 이라는 열이 있습니다. species 및 season 열은 문자 데이터입니다. subject 및 miles 열은 연속형 숫자 데이터입니다.

팁 : 여러 개의 텍스트 파일을 한 번에 가져와 데이터 테이블을 생성할 수 있습니다 . 자세한 내용은 JMP 사용의 에서 확인하십시오 .

데이터 테이블에 데이터 입력

데이터 테이블에 직접 데이터를 입력할 수 있습니다. 다음 예에서는 몇 개월 동안 수집된 데이터를 데이터 테이블에 입력하는 방법을 보여 줍니다.

시나리오

표 3.1 에서는 새로운 혈압약을 조사한 연구 자료를 보여 줍니다. 각 개인의 혈압을 6 개월 동안 측정했습니다. 대조군 및 위약군과 함께 두 개의 투약 그룹 ("300mg" 및 "450mg") 을 사용했습니다. 다음 데이터는 각 그룹의 평균 혈압을 보여 줍니다.

표 3.1 혈압 데이터

월	Control	Placebo	300mg	450mg
March	165	163	166	168
April	162	159	165	163
May	164	158	161	153
June	162	161	158	151
July	166	158	160	148
August	163	158	157	150

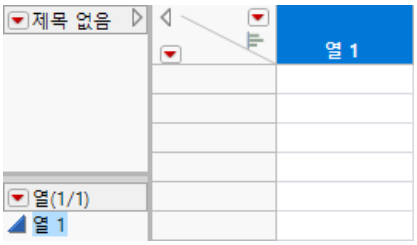
새 데이터 테이블에 데이터 입력

1. **파일 > 새로 만들기 > 데이터 테이블**을 선택하여 빈 데이터 테이블을 생성합니다 .

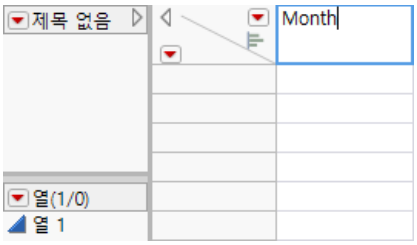
   새 데이터 테이블에는 하나의 열이 있고 행은 없습니다 .
2. 열 이름을 선택하고 이름을 Month 로 변경합니다 .

   참고 : 열 이름을 바꾸려면 열 이름을 두 번 클릭하거나 , 열을 선택하고 Enter 키를 눌러도 됩니다 .

그림 3.3 열 이름 입력



열을 한 번 클릭하여 선택합니다 .

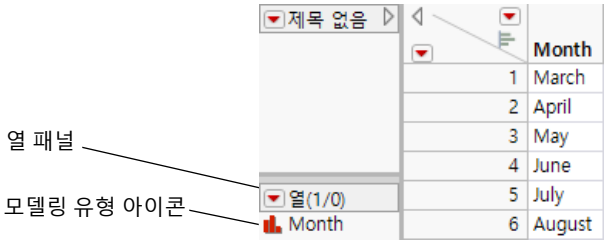


그런 다음 "Month" 를 입력합니다 .

3. **행 > 행 추가**를 선택합니다 .

   " 행 추가 " 창이 나타납니다 .
4. 6 개의 행을 추가하려고 하므로 "6" 을 입력합니다 .
5. **확인**을 클릭합니다 . 6 개의 빈 행이 데이터 테이블에 추가됩니다 .
6. 셀을 클릭하고 입력하는 방법으로 Month 정보를 입력합니다 .

그림 3.4 완료된 Month 열

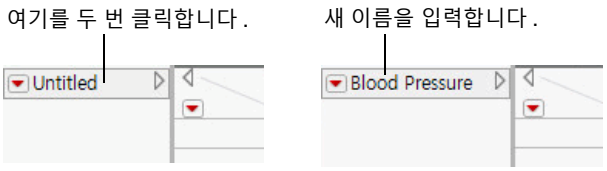


   " 열 " 패널에서 열 이름 왼쪽에 있는 모델링 유형 아이콘을 확인합니다 . 이전에는 연속형이었던 Month 가 이제 명목형임을 반영하여 변경되었습니다 . 그림 3.3 의 "Column 1" 과 그림 3.4 의 "Month" 에 표시된 모델링 유형을 비교해 보십시오 . 이 차이점은 중요하므로 " 데이터 테이블에서 열 정보 보기 또는 변경 "(75 페이지) 에서 자세히 설명합니다 .
7. Control 열을 추가하기 위해 "Month" 열의 오른쪽에 있는 공간을 두 번 클릭합니다 .
8. 이름을 Control 로 변경합니다 .
9. 표 3.1 에 표시된 대로 Control 데이터를 입력합니다 . 데이터 테이블은 이제 6 개의 행과 2 개의 열로 구성되어 있습니다 .
10. 계속해서 표 3.1 에 표시된 대로 열을 추가하고 데이터를 입력하여 6 개의 행과 5 개의 열을 포함하는 최종 데이터 테이블을 생성합니다 .

데이터 테이블 이름 변경

1. 테이블 패널에서 데이터 테이블 이름 (" 제목 없음 ") 을 두 번 클릭합니다 .
2. 새 이름 ("Blood Pressure") 을 입력합니다 .

그림 3.5 데이터 테이블 이름 변경



Excel 에서 JMP 로 데이터 전송

Excel 용 JMP 추가기능을 사용하여 Excel 스프레드시트를 JMP 의 다음 항목으로 전송할 수 있습니다 .

- 데이터 테이블
- 그래프 빌더
- 분포 플랫폼
- X 로 Y 적합 플랫폼
- 모형 적합 플랫폼
- 시계열 플랫폼
- 관리도 플랫폼

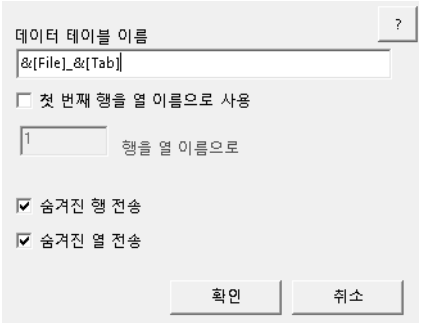
Excel 에서 JMP 추가기능 환경 설정 구성

JMP 추가기능 환경 설정을 구성하려면 다음을 수행하십시오 .

1. Excel 에서 **JMP > 환경 설정**을 선택합니다 .

   "JMP 환경 설정 " 창이 나타납니다 .

그림 3.6 JMP 추가기능 환경 설정



2. 기본 **데이터 테이블 이름** (파일 이름 _ 워크시트 이름) 을 그대로 사용하거나 이름을 입력합니다 .
3. 워크시트의 첫 번째 행에 열 머리글이 있으면 **첫 번째 행을 열 이름으로 사용**을 선택합니다 .

4. 첫 번째 행을 열 머리글로 사용하도록 선택한 경우 사용되는 행 수를 입력합니다 .
5. 워크시트에 있는 숨겨진 행을 JMP 데이터 테이블에 포함하려면 **숨겨진 행 전송**을 선택합니다 .
6. 워크시트에 있는 숨겨진 열을 JMP 데이터 테이블에 포함하려면 **숨겨진 열 전송**을 선택합니다 .
7. **확인**을 클릭하여 환경 설정을 저장합니다 .

JMP 로 이동

Excel 워크시트를 JMP 로 전송하려면 다음을 수행하십시오 .

1. Excel 파일을 엽니다 .
2. 전송할 워크시트를 선택합니다 .
3. **JMP** 를 선택한 후 다음 중에서 JMP 대상을 선택합니다 .
   - 데이터 테이블
   - 그래프 빌더
   - 분포 플랫폼
   - X 로 Y 적합 플랫폼
   - 모형 적합 플랫폼
   - 시계열 플랫폼
   - 관리도 플랫폼

   Excel 워크시트가 JMP에서 데이터 테이블로 열리고 선택한 플랫폼의 시작 창이 나타납니다.

데이터 테이블 작업

이 섹션에서는 데이터 테이블 작업에 대한 기본 개념을 설명합니다 .

- " 데이터 테이블에서 데이터 편집 "
- " 데이터 테이블에서 값 선택 , 선택 취소 및 찾기 "
- " 데이터 테이블에서 열 정보 보기 또는 변경 "
- " 계산식으로 값 계산의 예 "
- " 보고서 데이터 필터링의 예 "

팁 : " 일반 " 환경 설정에서 **자동 저장 타임아웃** 값을 설정하여 열려 있는 데이터 테이블을 지정된 시간 (분) 마다 자동으로 저장하는 것이 좋습니다 . 이 자동 저장 값은 저널 , 스크립트 , 프로젝트 및 보고서에도 적용됩니다 .

데이터 테이블에서 데이터 편집

한 번에 몇 개의 셀 또는 전체 열에 대해 데이터를 입력하거나 변경할 수 있습니다 . 이 섹션에는 다음과 같은 정보가 포함되어 있습니다 .

- " 데이터 테이블 셀의 값 변경 "
- " 재코딩 "
- " 패턴 있는 데이터 생성 "

데이터 테이블 셀의 값 변경

값을 변경하려면 셀을 선택하고 변경 내용을 입력합니다. 셀을 두 번 클릭하여 편집할 수도 있습니다.

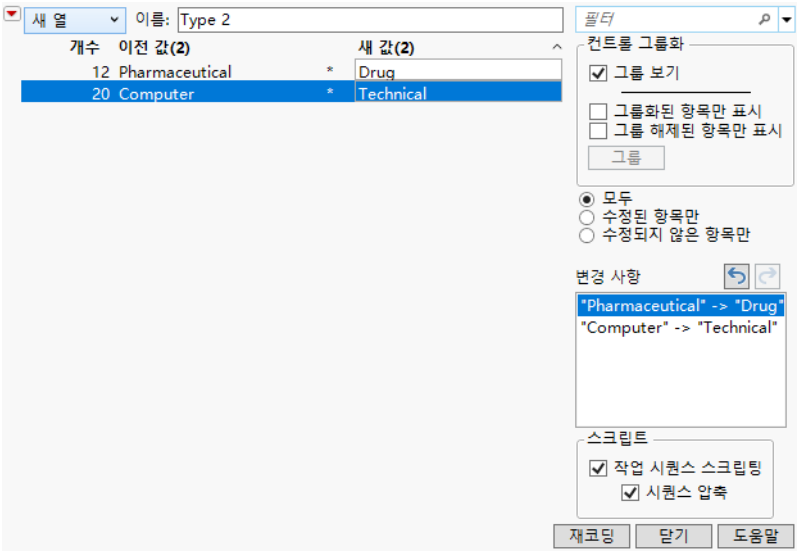
참고 : 셀을 두 번 클릭하는 것은 셀을 선택하는 것과 같지 않습니다 . 한 번 클릭하면 셀이 선택됩니다 . 한 번에 둘 이상의 셀을 선택할 수 있으며 선택한 셀에서 특정 작업을 수행할 수 있습니다 . 두 번 클릭해야 셀을 편집할 수 있습니다 . 행 , 열 및 셀 선택에 대한 자세한 내용은 " 데이터 테이블에서 값 선택 , 선택 취소 및 찾기 "(72 페이지) 에서 확인하십시오 .

재코딩

" 재코딩 " 도구를 사용하여 열의 모든 값을 한 번에 변경할 수 있습니다 . 예를 들어 컴퓨터 회사와 제약 회사의 매출을 비교하는 데 관심이 있다고 가정해 보겠습니다 . 현재 회사 라벨은 "Computer" 와 "Pharmaceutical" 입니다 . 이를 "Technical" 및 "Drug" 로 변경하려고 합니다 . 32 개 행의 모든 데이터를 검토하고 모든 값을 변경하는 것은 지루하고 비효율적이며 오류가 발생하기 쉽습니다 . 더 많은 데이터 행이 있을 때 특히 그렇습니다 . 재코딩이 더 좋은 방법입니다 .

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다 .
2. Type 열의 머리글을 한 번 클릭하여 해당 열을 선택합니다 .
3. **열 > 재코딩**을 선택합니다 .
4. " 재코딩 " 창의 " 새 값 " 열에서 "Computer" 행에 "Technical" 을 입력하고 "Pharmaceutical" 행에 "Drug" 를 입력합니다 .
5. " 새 열 " 목록에서 **현재 위치**를 선택합니다 .
6. **재코딩**을 클릭합니다 .

그림 3.7 재코딩 창



모든 셀이 자동으로 새 값으로 업데이트됩니다.

패턴 있는 데이터 생성

"채우기" 옵션을 사용하여 열을 패턴 있는 데이터로 채울 수 있습니다. "채우기" 옵션은 데이터 테이블이 커서 각 행에 값을 입력하는 것이 번거로울 때 특히 유용합니다.

예 : 패턴으로 열 채우기

1. 새 열을 추가합니다 .
2. 첫 번째 셀에는 "1" 을 , 두 번째 셀에는 "2" 를 , 세 번째 셀에는 "3" 을 입력합니다 .
3. 세 개의 셀을 선택하고 선택한 셀의 아무 곳이나 마우스 오른쪽 버튼으로 클릭하여 메뉴를 표시합니다 .
4. **채우기 > 테이블 끝까지 시퀀스 반복**을 선택합니다 .

   열의 나머지가 해당 시퀀스로 채워집니다 (1, 2, 3, 1, 2, 3, ...).

시퀀스를 반복하지 않고 패턴을 계속하려면 (1, 2, 3, 4, 5, 6, ...) **테이블 끝까지 시퀀스 계속**을 선택합니다. 이 명령을 사용하여 1, 1, 1, 2, 2, 2, 3, 3, 3, ... 같은 패턴을 생성할 수도 있습니다.

"채우기" 옵션은 간단한 산술 및 기하 시퀀스를 인식할 수 있습니다. 문자 데이터의 경우 "채우기" 옵션은 단순히 값을 반복합니다.

데이터 테이블에서 값 선택 , 선택 취소 및 찾기

데이터 테이블 내에서 행 , 열 또는 셀을 선택할 수 있습니다 . 예를 들어 기존 데이터 테이블의 부분집합을 생성하려면 먼저 테이블에서 부분집합을 만들려는 부분을 선택해야 합니다 . 또한 행을 선택하면 그래프에서 데이터 점이 쉽게 눈에 띄게 할 수 있습니다 . 수동으로 행과 열을 클릭하는 방법으로 선택하거나 , 특정 검색 기준을 충족하는 행을 선택하십시오 . 이 섹션에는 다음과 같은 정보가 포함되어 있습니다 .

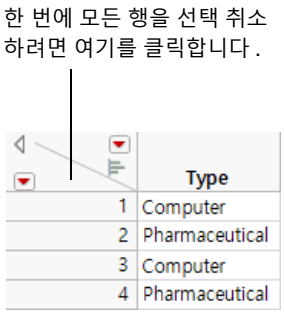
- " 행 선택 및 선택 취소 "(72 페이지)
- " 열 선택 및 선택 취소 "(73 페이지)
- " 셀 선택 및 선택 취소 "(74 페이지)
- " 값 검색 "(74 페이지)

행 선택 및 선택 취소

표 3.2 행 선택 및 선택 취소

작업	작업
한 번에 하나씩 행 선택	행 번호를 클릭합니다.
인접한 여러 행 선택	클릭하고 드래그하여 행 번호를 선택합니다. 또는 시작 행을 선택하고 Shift 키를 누른 후 마지막 행 번호를 클릭합니다.
인접하지 않은 여러 행 선택	첫 번째 행을 선택하고 Ctrl 키를 누른 후 다른 행 번호를 클릭합니다.
한 번에 하나씩 행 선택 취소	Ctrl 키를 누르고 행 번호를 클릭합니다.
모든 행 선택 취소	테이블의 왼쪽 상단에서 아래쪽 삼각형 공간을 클릭합니다(그림 3.8).

그림 3.8 행 선택 취소

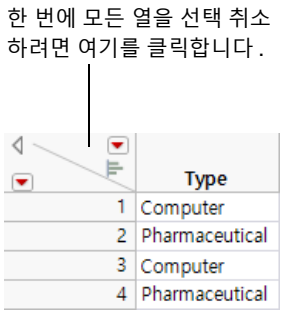


열 선택 및 선택 취소

표 3.3 열 선택 및 선택 취소

작업	작업
한 번에 하나씩 열 선택	열 머리글을 클릭합니다.
인접한 여러 열 선택	클릭하고 드래그하여 여러 열 머리글을 선택합니다. 또는 시작 열을 선택하고 Shift 키를 누른 후 마지막 머리글을 클릭합니다.
인접하지 않은 여러 열 선택	첫 번째 열을 선택하고 Ctrl 키를 누른 후 다른 열 머리글을 클릭합니다.
한 번에 하나씩 열 선택 취소	Ctrl 키를 누르고 열 머리글을 클릭합니다.
모든 열 선택 취소	테이블의 왼쪽 상단에서 위쪽 삼각형 공간을 클릭합니다(그림 3.9).

그림 3.9 열 선택 취소



셀 선택 및 선택 취소

표 3.4 셀 선택 및 선택 취소

작업	작업
한 번에 하나씩 셀 선택	각 셀을 개별적으로 클릭합니다.
인접한 여러 셀 선택	클릭하고 드래그하여 여러 셀을 선택합니다. 또는 시작 셀을 선택하고 Shift 키를 누른 후 마지막 셀을 클릭합니다.
인접하지 않은 여러 셀 선택	첫 번째 셀을 선택하고 Ctrl 키를 누른 후 다른 셀을 클릭합니다.
셀 선택 취소	테이블의 왼쪽 상단에서 위쪽 및 아래쪽 삼각형 공간을 클릭합니다.

값 검색

수천 또는 수만 개의 행이 있는 데이터 테이블에서는 테이블을 스크롤하여 특정 셀을 찾는 것이 어려울 수 있습니다. 특정 정보를 찾으려면 검색 기능을 사용하십시오. 데이터가 검색 기준과 매칭되면 해당 셀이 선택되고 창에 해당 셀이 표시되도록 데이터 격자가 스크롤됩니다. 예를 들어 Companies.jmp 데이터 테이블에는 총 매출이 11,899 달러인 회사에 대한 정보가 있습니다. 해당 셀을 찾으려면 검색 기능을 사용하십시오.

예 : 값 검색

1. **편집 > 검색 > 찾기**를 선택하여 검색 창을 시작합니다.
2. **찾을 내용** 상자에 "11899" 를 입력합니다.

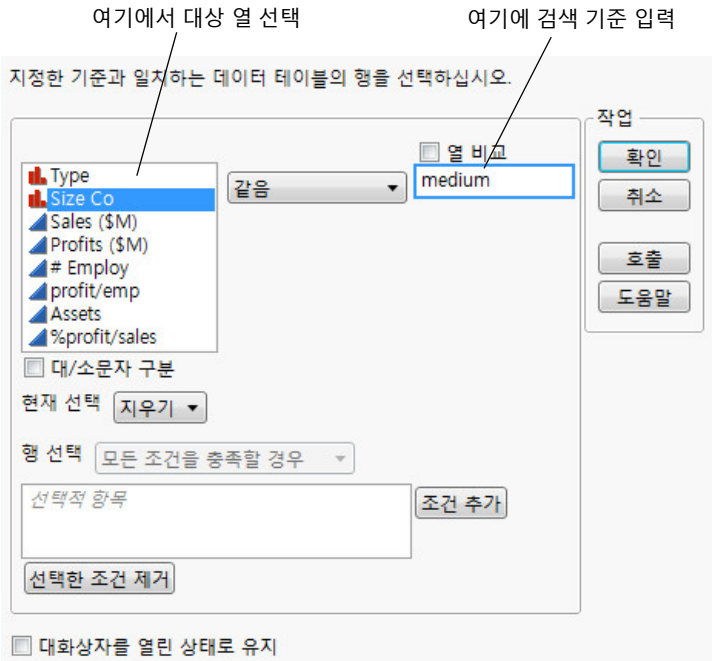
3. **찾기**를 클릭합니다 . 11,899 가 포함된 첫 번째 셀이 검색되고 선택됩니다 .

여러 셀이 검색 기준을 충족하면 **찾기**를 다시 클릭하여 검색어와 매칭되는 다음 셀을 찾습니다. 기준과 매칭되는 여러 행을 한 번에 검색할 수도 있습니다.

예 : 중간 규모 회사에 해당하는 모든 행 선택

1. **행 > 행 선택 > 선택 조건**을 선택하여 **행 선택하기** 창을 엽니다 .
2. 왼쪽의 열 목록 상자에서 Size Co 를 선택합니다 .
3. 오른쪽의 텍스트 상자에 "medium" 을 입력합니다 .
4. **확인**을 클릭합니다 .

그림 3.10 행 선택하기 창



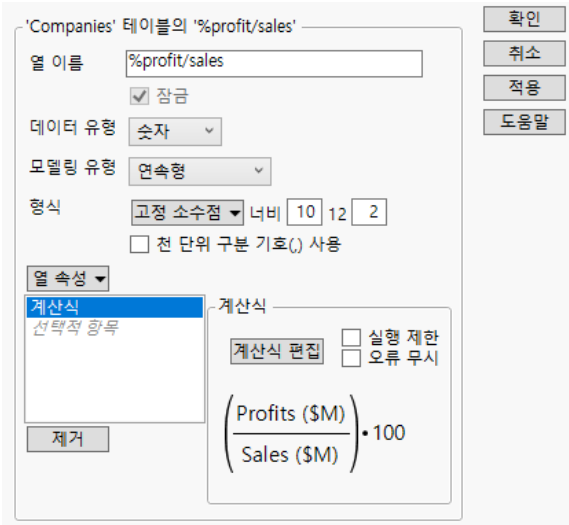
Size Co가 "medium"인 행이 모두 선택됩니다. 7개가 있습니다.

데이터 테이블에서 열 정보 보기 또는 변경

데이터 테이블 열에 대한 정보는 열의 데이터로 제한되지 않습니다. 데이터 유형, 모델링 유형, 형식 및 계산식도 설정할 수 있습니다.

열 특성을 보거나 변경하려면 열 머리글을 두 번 클릭합니다. 또는 열 머리글을 마우스 오른쪽 버튼으로 클릭하고 **열 정보**를 선택합니다. "열 정보" 창이 나타납니다.

그림 3.11 열 정보 창



열 이름 열 이름을 입력하거나 변경합니다. 두 열의 이름이 같을 수 없습니다.

데이터 유형 다음 데이터 유형 중 하나를 선택합니다.

숫자 열 값을 숫자로 지정합니다.

문자 열 값을 문자 또는 기호와 같이 숫자가 아닌 값으로 지정합니다.

행 상태 열 값을 행 상태로 지정합니다. 이에 대해서는 자세한 설명이 필요합니다. 자세한 내용은 JMP 사용의 에서 확인하십시오.

모델링 유형 모델링 유형은 값이 분석에서 사용되는 방식을 정의합니다. 다음 모델링 유형 중 하나를 선택합니다.

연속형 숫자 유형에만 해당됩니다.

순서형 숫자 또는 문자 값으로, 순서가 있는 범주입니다.

명목형 숫자 또는 문자 값이지만 순서가 없습니다.

형식 숫자 값의 형식을 선택합니다. 문자 데이터에는 이 옵션을 사용할 수 없습니다. 다음은 가장 일반적인 형식 중 몇 가지입니다.

최적 최적의 표시 형식이 자동으로 선택됩니다.

고정 소수점 표시되는 소수 자릿수를 지정합니다.

날짜 날짜 값의 구문을 지정합니다.

시간 시간 값의 구문을 지정합니다.

통화 통화 값에 사용되는 통화 유형과 소수 자릿수를 지정합니다.

열 특성 계산식, 노트 및 값 순서와 같은 특수한 열 특성을 설정합니다. 자세한 내용은 JMP 사용의 에서 확인하십시오.

잠금 열의 값을 변경할 수 없도록 열을 잠급니다.

계산식으로 값 계산의 예

계산식 편집기를 사용하여 계산된 값이 포함된 열을 생성할 수 있습니다.

시나리오

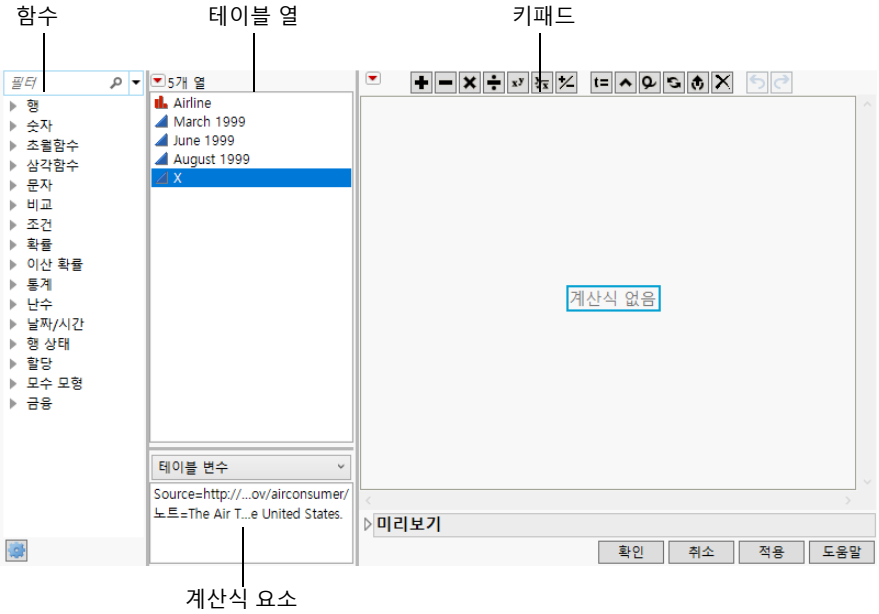
샘플 데이터 테이블 On-Time Arrivals.jmp는 여러 항공사의 정시 도착률을 나타냅니다. 수집된 데이터는 1999년 3월, 6월, 8월의 데이터입니다.

계산식 생성

각 항공사의 평균 정시 도착률을 포함하는 새 열을 생성하려고 한다고 가정해 보겠습니다.

1. 새 열을 추가합니다.
2. 새 열의 열 머리글을 마우스 오른쪽 버튼으로 클릭하고 **계산식**을 선택합니다. 계산식 편집기 창이 나타납니다.

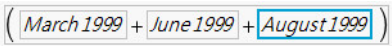
그림 3.12 계산식 편집기



각 항공사의 평균 정시 도착률에 대한 계산식을 생성합니다 .

3. 열 목록에서 March 1999 를 선택합니다 .
4. 키패드에서 ✚ 버튼을 클릭합니다 .
5. June 1999 를 선택한 후 다시 ✚ 기호를 클릭합니다 .
6. August 1999 를 선택합니다 .

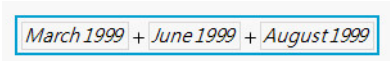
그림 3.13 모든 월의 합계



현재는 "August 1999" 만 선택되어 있습니다 (파란색 상자로 둘러싸임).

7. 전체 계산식을 둘러싼 상자를 클릭합니다 .

그림 3.14 전체 계산식이 선택됩니다 .



8. ➗ 버튼을 클릭합니다 .
9. 분모 상자에 "3" 을 입력한 다음 계산식의 바깥쪽 공백을 클릭합니다 .

그림 3.15 완료된 계산식

10. **확인**을 클릭합니다 .

    새 열에 평균이 포함됩니다 .

계산식 편집기에는 많은 기본 제공 산술 함수와 통계 함수가 있습니다. 예를 들어 평균 정시 도착률을 계산하는 또 다른 방법은 통계 함수 목록에서 **Mean** 함수를 사용하는 것입니다. 계산식 편집기의 모든 함수에 대한 자세한 내용은 JMP 사용의 에서 확인하십시오.

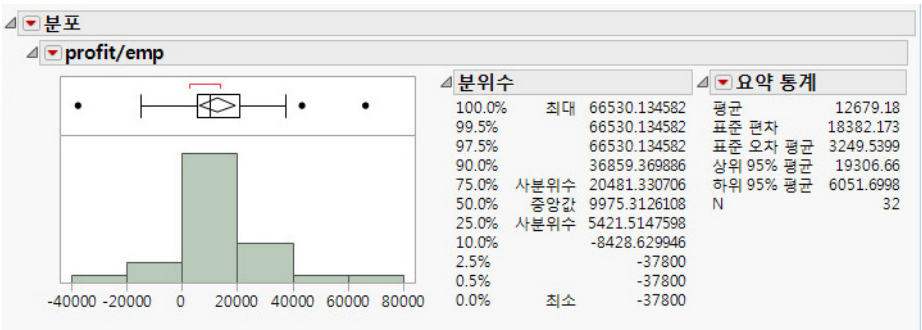
보고서 데이터 필터링의 예

데이터 필터를 사용하여 복잡한 데이터 부분집합을 대화식으로 선택하거나 , 이러한 부분집합을 그림에서 숨기거나 , 분석에서 제외할 수 있습니다 . 한 예로 컴퓨터 회사와 제약 회사의 직원당 수익을 살펴보겠습니다 .

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다 .
2. **분석** > **분포**를 선택합니다 .
3. profit/emp 를 선택하고 **Y, 열**을 클릭합니다 .

4. **확인**을 클릭합니다 .
5. "profit/emp" 옆의 빨간색 삼각형을 클릭하고 **표시 옵션 > 가로 레이아웃**을 선택합니다 .

그림 3.16 profit/emp 의 분포



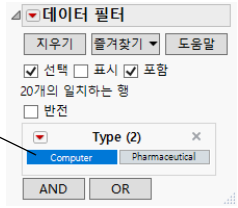
6. " 분포 " 의 빨간색 삼각형에서 **다시 실행 > 자동 재계산**을 선택하여 자동 재계산 기능을 설정합니다 .

   이 옵션을 설정하면 점을 숨기거나 제외하는 등의 변경 작업을 수행할 때마다 보고서 창이 자동으로 업데이트됩니다 .
7. 데이터 테이블에서 **행** > **데이터 필터**를 선택합니다 .
8. Type 을 선택하고 **추가**를 클릭합니다 .
9. " 선택 " 체크박스가 선택되어 있는지 확인합니다 .
10. 분포 결과에서 제약 회사를 제외하고 컴퓨터 회사만 포함하려면 데이터 필터 창에서 **Computer** 상자를 클릭합니다 .

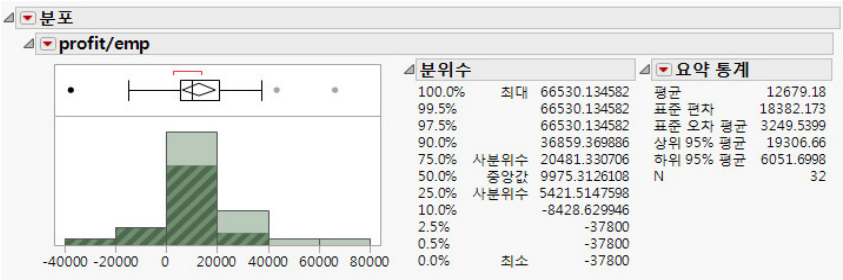
    분포 결과가 업데이트되어 컴퓨터 회사만 포함됩니다 .

그림 3.17 컴퓨터 회사에 대한 필터

분포 결과에 컴퓨터 회사만 포함하려면 "Computer" 상자를 클릭합니다.



그래프 및 통계량 보고서에 선택된 행만 자동으로 반영됩니다.



반대로 제약 회사만 포함되도록 분포 결과를 변경하려면 "데이터 필터" 창에서 **Pharmaceutical** 상자를 클릭합니다.

데이터 재구성의 예

테이블 메뉴의 명령과 **분석** 메뉴의 " 테이블 생성 " 명령은 그래프 생성 및 분석에 필요한 형식으로 데이터 테이블을 요약하고 조작합니다. 이 섹션에서는 다음 5 가지 명령에 대해 설명합니다.

요약 데이터를 설명하는 요약 통계량을 포함하는 테이블을 생성합니다.

테이블 생성 요약 통계량을 생성하기 위한 드래그앤드롭 작업 공간을 제공합니다.

부분집합 데이터의 부분집합을 포함하는 테이블을 생성합니다.

결합 두 데이터 테이블의 데이터를 하나의 새 데이터 테이블로 결합합니다.

정렬 하나 이상의 열을 기준으로 데이터를 정렬합니다.

이러한 명령과 기타 "테이블" 메뉴 명령에 대한 자세한 내용은 JMP 사용의 에서 확인하십시오.

요약 통계량 보기

합 및 평균과 같은 요약 통계량은 데이터에 대한 유용한 정보를 즉시 제공합니다. 예를 들어 32개 회사의 개별 연간 수익만 봐서는 소규모, 중간 규모, 대규모 기업의 수익을 비교하기가 어렵습니다. 요약에서는 이러한 정보를 즉시 보여 줍니다.

요약 테이블을 생성하려면 **요약** 또는 **테이블 생성** 명령을 사용합니다. **요약** 명령을 사용하면 새 데이터 테이블이 생성됩니다. 다른 데이터 테이블과 마찬가지로 요약 테이블에서도 분석을 수

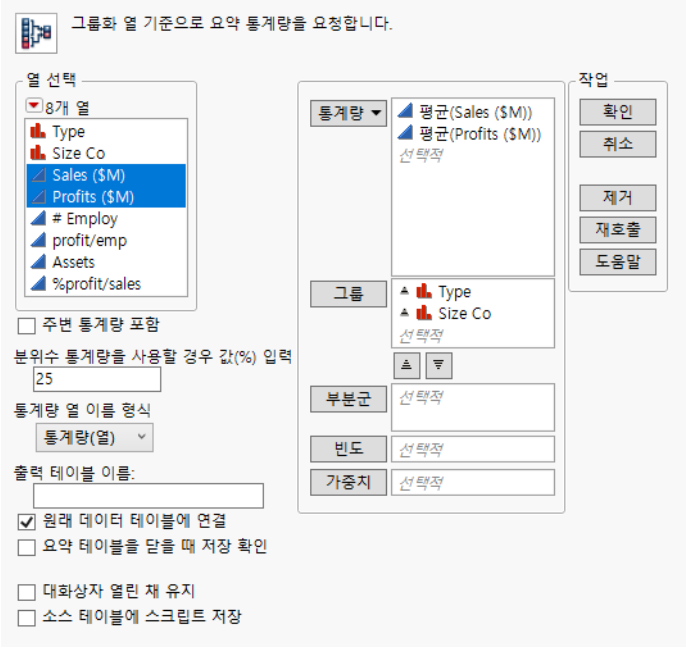
행하고 그래프를 생성할 수 있습니다. **테이블 생성** 명령을 사용하면 요약 데이터 테이블이 있는 보고서 창이 생성됩니다. "테이블 생성" 보고서에서 테이블을 생성할 수도 있습니다.

요약 테이블 예제

요약 테이블에는 그룹화 변수의 각 수준에 대한 통계량이 포함됩니다. 예를 들어 컴퓨터 및 제약 회사의 재무 데이터를 살펴보십시오. 회사 유형 및 규모의 각 조합에 대해 매출 평균 및 수익 평균을 계산한다고 가정해 보겠습니다.

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다.
2. **테이블 > 요약**을 선택합니다.
3. Type 및 Size Co 를 선택하고 **그룹**을 클릭합니다.
4. Sales ($M) 및 Profits ($M) 를 선택하고 **통계량 > 평균**을 클릭합니다.

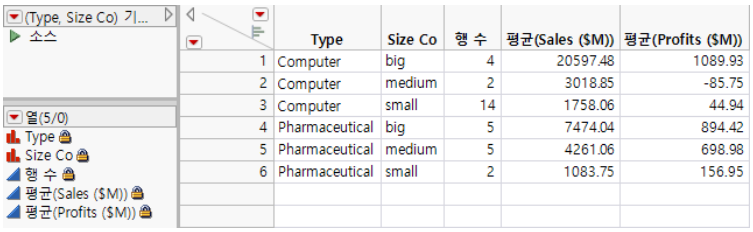
그림 3.18 완료된 요약 창



5. **확인**을 클릭합니다.

Type 및 Size Co 의 각 조합에 대해 Sales ($M) 평균과 Profit ($M) 평균이 계산됩니다.

그림 3.19 요약 테이블



	Type	Size Co	행 수	평균(Sales ($M))	평균(Profits ($M))
1	Computer	big	4	20597.48	1089.93
2	Computer	medium	2	3018.85	-85.75
3	Computer	small	14	1758.06	44.94
4	Pharmaceutical	big	5	7474.04	894.42
5	Pharmaceutical	medium	5	4261.06	698.98
6	Pharmaceutical	small	2	1083.75	156.95

요약 테이블의 내용은 다음과 같습니다 .

- 각 그룹화 변수 (이 예의 경우 Type 및 Size Co) 에 대한 열이 있습니다 .
- **행 수** 열에는 그룹화 변수의 각 조합에 해당하는 원래 테이블의 행 수가 표시됩니다 . 예를 들어 원래 데이터 테이블에는 소규모 컴퓨터 회사에 해당하는 14 개 행이 있습니다 .
- 요청된 각 요약 통계량에 대한 열이 있습니다 . 이 예에는 Sales ($M) 의 평균 열과 Profits ($M) 의 평균 열이 있습니다 .

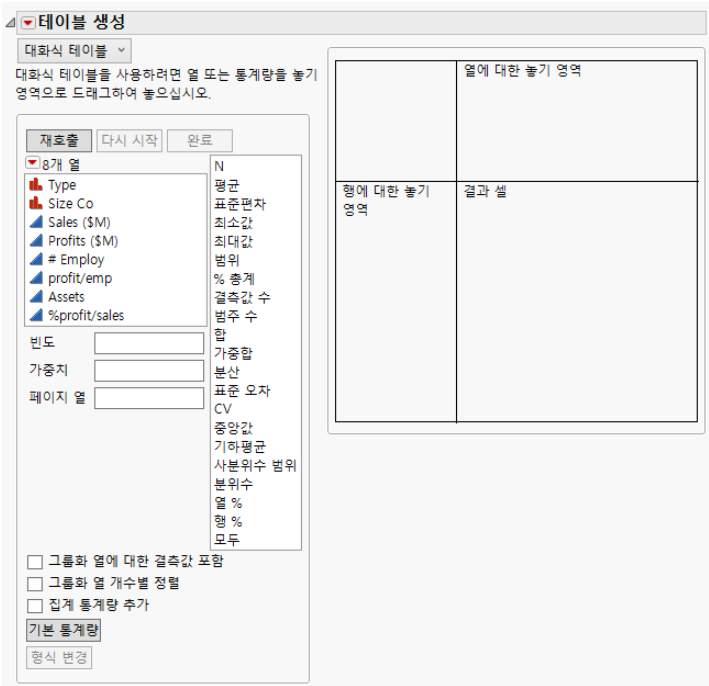
요약 테이블은 소스 테이블에 연결됩니다. 요약 테이블에서 행을 선택하면 소스 테이블에서도 해당 행이 선택됩니다.

테이블 생성 예제

" 테이블 생성 " 명령을 사용하여 열을 작업 공간으로 드래그하면 그룹화 변수의 각 조합에 대한 요약 통계량을 생성할 수 있습니다 . 이 예에서는 방금 " 요약 " 명령을 사용하여 생성한 것과 동일한 요약 정보를 " 테이블 생성 " 명령으로 생성하는 방법을 보여 줍니다 .

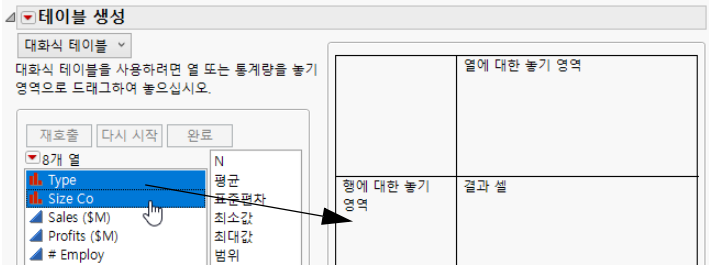
1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다 .
2. **분석** > **테이블 생성**을 선택합니다 .

그림 3.20 테이블 생성 작업 공간



3. Type 과 Size Co 를 모두 선택합니다 .
4. 선택한 항목을 **행에 대한 놓기 영역**으로 드래그하여 놓습니다 .

그림 3.21 열을 행 영역으로 드래그



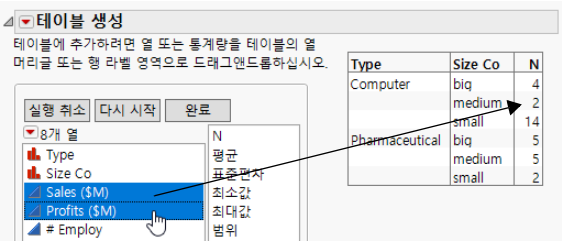
5. 머리글을 마우스 오른쪽 버튼으로 클릭하고 **그룹화 열 내포**를 선택합니다 .
   초기 테이블은 그룹당 행 수를 보여 줍니다 .

그림 3.22 초기 테이블

Type	Size Co	개수
Computer	big	4
	medium	2
	small	14
Pharmaceutical	big	5
	medium	5
	small	2

6. Sales ($M) 및 Profits ($M) 를 모두 선택한 후 테이블의 **개수**로 드래그하여 놓습니다 .

그림 3.23 매출 및 수익 추가



이제 테이블에는 그룹당 Sales ($M) 합과 Profits ($M) 합이 표시됩니다 .

그림 3.24 합 테이블

		Sales ($M)	Profits ($M)
Type	Size Co	합계	합계
Computer	big	82389.9	4359.7
	medium	6037.7	-171.5
	small	24612.8	629.1
Pharmaceutical	big	37370.2	4472.1
	medium	21305.3	3494.9
	small	2167.5	313.9

7. 마지막 단계는 합을 평균으로 변경하는 것입니다 . 둘 중 하나의 **합**을 마우스 오른쪽 버튼으로 클릭하고 **통계량** > **평균**을 선택합니다 .

그림 3.25 최종 테이블

		Sales ($M)	Profits ($M)
Type	Size Co	평균	평균
Computer	big	20597.48	1089.9
	medium	3018.85	-85.75
	small	1758.06	44.94
Pharmaceutical	big	7474.04	894.42
	medium	4261.06	698.98
	small	1083.75	156.95

이 평균은 요약 명령을 사용하여 구한 것과 동일합니다. 그림 3.25 와 그림 3.19 를 비교하십시오.

부분집합 생성의 예

데이터 테이블의 일부만 자세히 보기 위해 부분집합을 생성할 수 있습니다. 예를 들어 대규모, 중간 규모, 소규모 컴퓨터 회사와 제약 회사의 매출과 수익을 이미 비교했다고 가정해 보겠습니다. 이제 중간 규모 회사의 매출과 수익만 살펴보려고 합니다.

부분집합 생성 과정은 두 단계로 이루어집니다. 먼저 대상 데이터를 선택한 후 해당 데이터를 새 테이블로 추출합니다.

부분집합 명령으로 부분집합 생성

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다 .

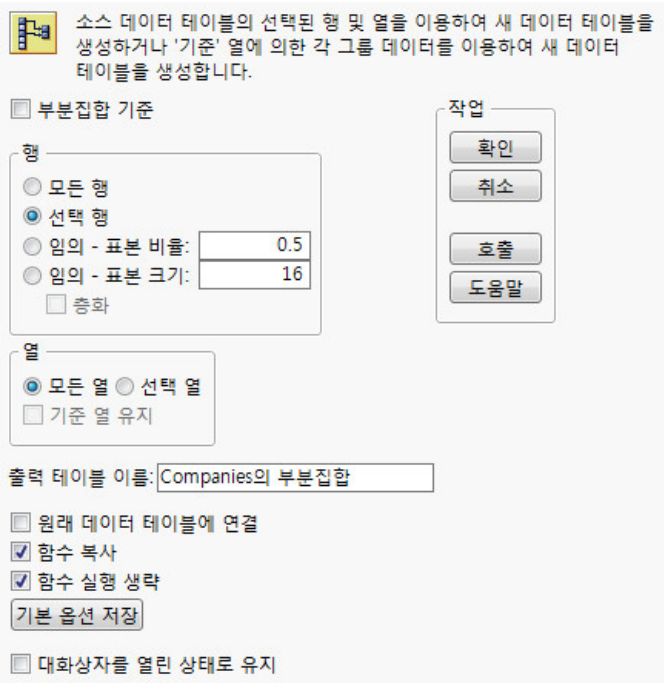
부분집합을 생성할 행 및 열 선택

2. **행 > 행 선택 > 선택 조건**을 선택합니다 .
3. 왼쪽의 열 목록 상자에서 Size Co 를 선택합니다 .
4. 텍스트 입력 상자에 "medium" 을 입력합니다 .
5. **확인**을 클릭합니다 .
6. Ctrl 키를 누르고 Type, Sales ($M) 및 Profits ($M) 열을 선택합니다 .

부분집합 테이블 생성

7. **테이블 > 부분집합**을 선택하여 " 부분집합 " 창을 시작합니다 .

그림 3.26 부분집합 창



8. 선택한 열로만 부분집합을 생성하려면 **선택 열**을 선택합니다. 추가 옵션을 선택하여 부분집합 테이블을 추가로 사용자 정의할 수도 있습니다.
9. **확인**을 클릭합니다.

결과 부분집합 데이터 테이블에는 7개의 행과 3개의 열이 있습니다. "부분집합" 명령에 대한 자세한 내용은 JMP 사용의 에서 확인하십시오.

분포 플랫폼으로 부분집합 생성

부분집합을 생성하는 또 다른 방법은 플랫폼 결과와 데이터 테이블 간의 연결을 사용하는 것입니다.

예 : 분포 명령을 사용하여 부분집합 생성

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다.
2. **분석 > 분포**를 선택합니다.
3. Type 을 선택하고 **Y, 열**을 클릭합니다.
4. **확인**을 클릭합니다.
5. "Computer" 를 나타내는 히스토그램 막대를 두 번 클릭하여 컴퓨터 회사의 부분집합 테이블을 생성합니다.

주의 : 이 방법을 사용하면 **연결된** 부분집합 테이블이 생성됩니다 . 즉 , 부분집합 테이블의 데이터를 변경하면 소스 테이블에서도 해당 값이 변경됩니다 .

예 : 데이터 테이블 결합

"결합" 옵션을 사용하여 여러 데이터 테이블의 정보를 단일 데이터 테이블에 결합할 수 있습니다. 예를 들어 팝콘 산출량에 대한 실험 결과가 포함된 데이터 테이블이 있다고 가정해 보겠습니다. 다른 데이터 테이블에는 팝콘 산출량에 대한 두 번째 실험 결과가 있습니다. 두 실험을 비교하거나, 두 결과 모두를 사용하여 실험을 분석하려면 동일한 테이블에 데이터가 있어야 합니다. 또한 실험 데이터는 동일한 순서로 데이터 테이블에 입력되지 않았습니다. 한 열은 이름이 다르며 두 번째 실험은 불완전합니다. 따라서 한 테이블에서 다른 테이블로 복사하여 붙여넣을 수 없습니다.

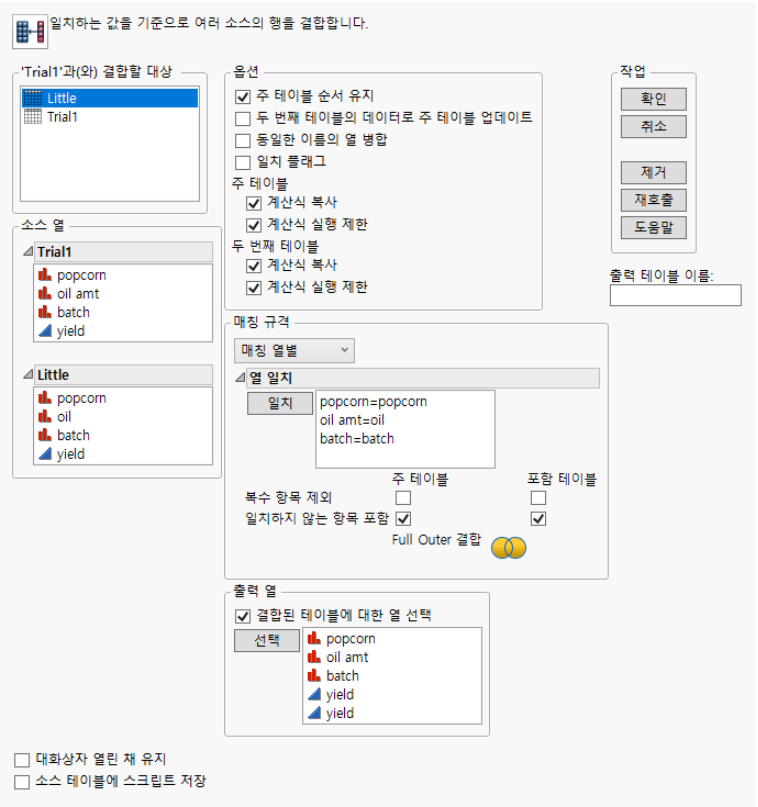
예 : 두 개의 데이터 테이블 결합

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Trial1.jmp 및 Little.jmp 를 엽니다 .
2. Trial1.jmp 데이터 테이블을 클릭하여 활성화합니다 .
3. **테이블 > 결합**을 선택합니다 .
4. **'Trial1' 과 (와) 결합할 대상** 상자에서 Little 을 선택합니다 .
5. **매칭 규격** 메뉴에서 **매칭 열별**이 선택되어 있지 않으면 이 항목을 선택합니다 .
6. **소스 열** 상자 안의 두 상자 모두에서 popcorn 을 선택하고 **매칭**을 클릭합니다 .
7. 같은 방법으로 두 상자의 batch 와 batch 를 매칭하고 oil amt 와 oil 을 매칭합니다 .

   매칭 열의 이름이 같을 필요는 없습니다 .
8. 두 테이블 모두에서 **매칭되지 않는 항목 포함**을 선택합니다 .

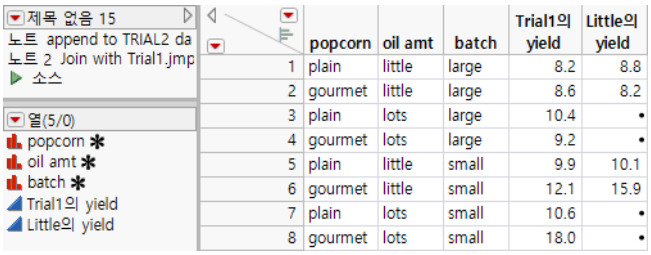
   한 실험은 부분적이기 때문에 결측 데이터가 있는 행을 비롯해서 모든 행을 포함하고자 합니다.
9. 열 중복을 피하기 위해 **결합된 테이블의 열 선택** 옵션을 선택합니다 .
10. Trial1 에서 4 개의 열을 모두 선택하고 **선택**을 클릭합니다 .
11. Little 에서 yield 만 선택하고 **선택**을 클릭합니다 .

그림 3.27 완료된 결합 창



12. **확인**을 클릭합니다 .

그림 3.28 결합된 테이블



	popcorn	oil amt	batch	Trial1의 yield	Little의 yield
1	plain	little	large	8.2	8.8
2	gourmet	little	large	8.6	8.2
3	plain	lots	large	10.4	•
4	gourmet	lots	large	9.2	•
5	plain	little	small	9.9	10.1
6	gourmet	little	small	12.1	15.9
7	plain	lots	small	10.6	•
8	gourmet	lots	small	18.0	•

데이터 정렬

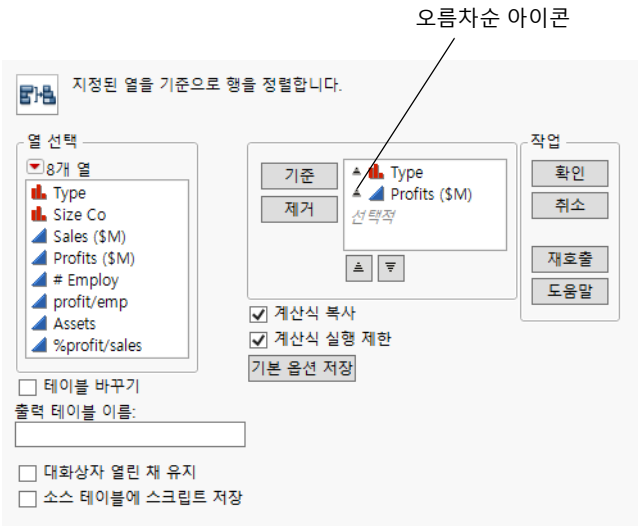
" 정렬 " 명령을 사용하여 데이터 테이블에 있는 하나 이상의 열을 기준으로 데이터 테이블을 정렬할 수 있습니다 . 예를 들어 컴퓨터 및 제약 회사의 재무 데이터를 살펴보십시오 . 처음에는 Type,

그 다음에는 Profits ($M) 를 기준으로 하여 데이터 테이블을 정렬한다고 가정해 보겠습니다 . 또한 각 Type 내에서 Profits ($M) 를 내림차순으로 표시하려고 합니다 .

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다 .
2. **테이블** > **정렬**을 선택합니다 .
3. Type 을 선택하고 **기준**을 클릭하여 Type 을 정렬 변수로 할당합니다 .
4. Profits ($M) 를 선택하고 **기준**을 클릭합니다 .

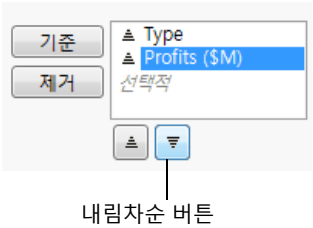
현재 두 변수는 모두 오름차순으로 정렬되도록 설정되어 있습니다. 그림 3.29 에서 변수 옆에 있는 오름차순 아이콘을 확인하십시오.

그림 3.29 오름차순 정렬 아이콘



5. Profits ($M) 를 내림차순으로 정렬하도록 변경하기 위해 Profits ($M) 를 선택하고 내림차순 버튼을 클릭합니다 .

그림 3.30 수익을 내림차순으로 변경



"Profits ($M)" 옆의 아이콘이 내림차순으로 변경됩니다 .

6. **테이블 바꾸기** 체크박스를 선택합니다 .

테이블 바꾸기를 선택하면 정렬된 값으로 새 테이블이 생성되지 않고 원래 데이터 테이블이 정렬됩니다. 원래 데이터 테이블에서 생성된 보고서 창이 열려 있으면 이 옵션을 사용할 수 없습니다. 관련 보고서 창이 열려 있는 상태에서 데이터 테이블을 정렬하면 데이터 중 일부가 보고서 창, 특히 그래프에 표시되는 방식이 변경될 수 있습니다.

7. **확인**을 클릭합니다.

이제 데이터 테이블이 유형 기준 사전순으로 정렬되고, 유형 내에서는 수익 합계 기준 내림차순으로 정렬됩니다.

4 장

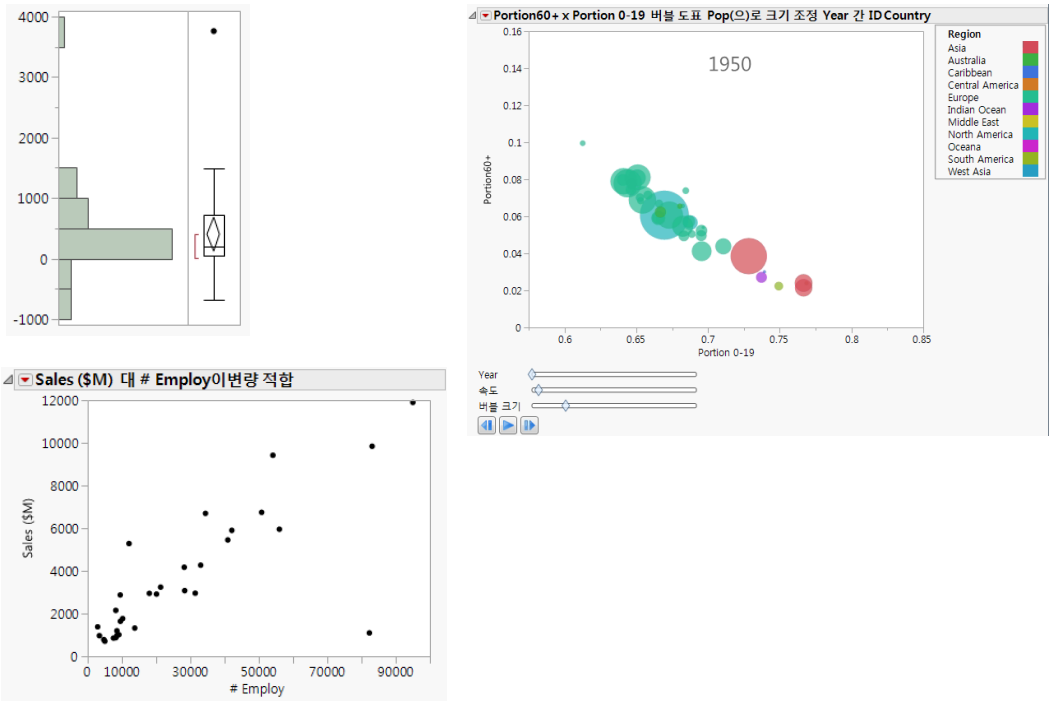
데이터 시각화

일반적인 그래프

데이터 시각화는 중요한 첫 번째 단계입니다. 이 장에서 설명하는 그래프는 데이터에 대한 중요한 상세 정보를 발견하는 데 도움이 됩니다. 예를 들어 히스토그램은 데이터의 모양과 범위를 보여 주고 비정상적인 데이터 점을 찾는 데 도움을 줍니다.

이 장에서는 JMP에서 데이터를 시각화하고 탐색하는 데 사용되는 가장 일반적인 그래프 및 그림 중 몇 가지를 설명합니다. 이 장에서는 JMP의 그래픽 도구 및 플랫폼을 소개합니다. JMP를 사용하여 단일 변수의 분포나 여러 변수 간의 관계를 시각화할 수 있습니다.

그림 4.1 JMP 를 사용하여 데이터 시각화



목차

단변량 그래프의 단일 변수 분석 93
연속형 변수에 히스토그램 사용 93
범주형 변수에 막대 차트 사용 96
다중 변수 비교 98
산점도를 사용하여 여러 변수 비교 99
산점도 행렬을 사용하여 여러 변수 비교 102
병렬 상자 그림을 사용하여 여러 변수 비교 105
그래프 빌더를 사용하여 여러 변수 비교 107
버블 그림을 사용하여 여러 변수 비교 113
중첩 그림을 사용하여 여러 변수 비교 118
계량형 차트를 사용하여 여러 변수 비교 122

단변량 그래프의 단일 변수 분석

단일 변수 그래프, 즉 **단변량** 그래프를 사용하여 한 번에 하나의 변수를 자세히 살펴볼 수 있습니다. 데이터를 살펴보기 시작할 때는 각 변수 간의 교호작용을 확인하기 전에 각 변수에 대해 알아두는 것이 중요합니다. 단변량 그래프를 사용하면 각 변수를 개별적으로 시각화할 수 있습니다.

이 섹션에서는 단일 변수의 분포를 보여 주는 두 가지 그래프를 다룹니다 .

- " 연속형 변수에 히스토그램 사용 "(93 페이지)
- " 범주형 변수에 막대 차트 사용 "(96 페이지)

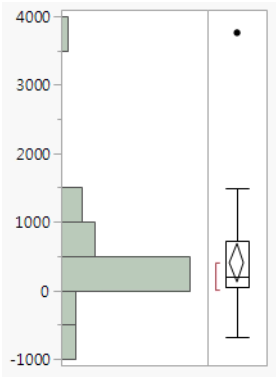
분포 플랫폼을 사용하여 두 그래프를 모두 생성할 수 있습니다. 분포 플랫폼은 각 변수에 대한 그래픽 설명과 기술 통계량을 생성합니다.

연속형 변수에 히스토그램 사용

히스토그램은 연속형 변수의 분포를 파악하는 데 가장 유용한 그래픽 도구 중 하나입니다 . 데이터에서 다음 정보를 확인하려면 히스토그램을 사용합니다 .

- 평균 값 및 변동
- 극단값

그림 4.2 히스토그램의 예



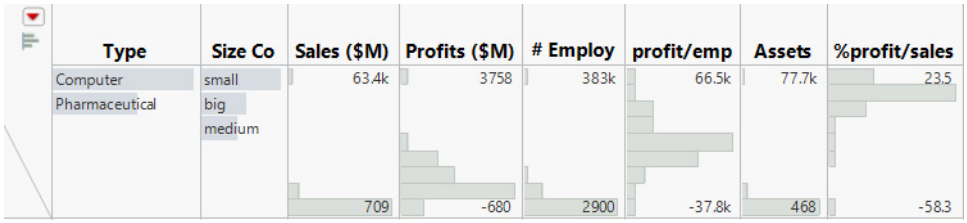
즉석 히스토그램

열 머리글에 있는 히스토그램 아이콘을 클릭하면 곧바로 히스토그램을 볼 수 있습니다. 이 히스토그램은 열 머리글 아래에 표시됩니다.

그림 4.3 즉석 히스토그램

	Type	Size Co	Sales ($M)	Profits ($M)	# Employ	profit/emp	Assets	%profit/sales
1	Computer	small	855.1	31.0	7523	4120.70	615.2	3.63
2	Pharmaceutical	big	5453.5	859.8	40929	21007.11	4851.6	15.77
3	Computer	small	2153.7	153.0	8200	18658.54	2233.7	7.10

히스토그램 아이콘을 클릭합니다 .



시나리오

이 예에서는 특정 회사 그룹의 수익 데이터가 포함된 Companies.jmp 데이터 테이블을 사용합니다. 재무 분석가가 다음과 같은 질문에 대한 답을 구하려고 합니다 .

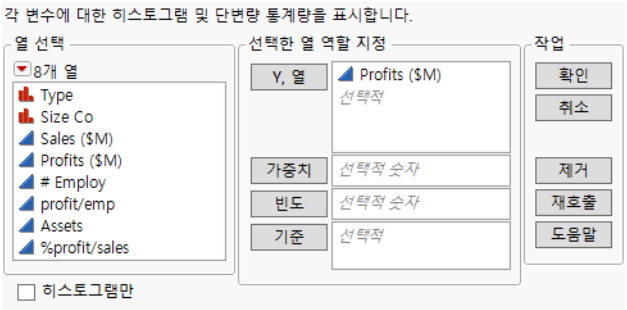
- 일반적으로 각 회사는 얼마나 많은 수익을 올립니까 ?
- 평균 수익은 얼마입니까 ?
- 다른 회사에 비해 극단적으로 높거나 낮은 수익을 올리는 회사가 있습니까 ?

이러한 질문에 답하려면 Profits ($M) 의 히스토그램을 사용합니다.

히스토그램 생성

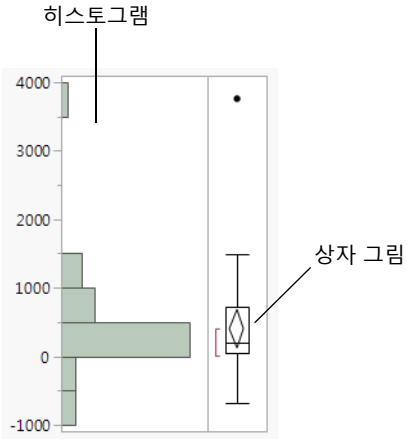
1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다 .
2. **분석** > **분포**를 선택합니다 .
3. Profits ($M) 를 선택하고 **Y, 열**을 클릭합니다 .

그림 4.4 Profits ($M) 의 분포 창



4. **확인**을 클릭합니다 .

그림 4.5 Profits ($M) 의 히스토그램



히스토그램 해석

이 히스토그램에서는 다음을 알 수 있습니다 .

- 대부분의 회사는 수익이 -1,000 달러 ~ 1,500 달러 사이입니다 .

  하나를 제외한 모든 막대가 이 범위에 있습니다 . 또한 수익이 0 달러 ~ 500 달러 범위에 있는 회사가 가장 많습니다 . 해당 범위를 나타내는 막대가 다른 모든 막대보다 훨씬 깁니다 .
- 평균 수익은 500 달러보다 약간 적습니다 .

  상자 그림에 있는 마름모의 중심은 평균 값을 나타냅니다 . 여기서는 평균이 500 달러보다 약간 낮습니다 .
- 한 회사는 다른 회사들보다 훨씬 더 높은 수익을 올리고 있어 **이상치**로 간주될 수 있습니다 . 이상치는 다른 데이터 점의 일반적인 패턴과 동떨어져 있는 데이터 점입니다 .

이 이상치는 히스토그램의 상단에 하나의 매우 짧은 막대로 표시됩니다 . 이 작은 막대는 소규모 그룹 (여기서는 단일 회사) 을 나타내며 나머지 히스토그램 막대와 상당히 동떨어져 있습니다 .

이 보고서에는 히스토그램 외에도 다음이 포함됩니다 .

- 데이터의 또 다른 그래픽 요약인 상자 그림 . 상자 그림에 대한 자세한 내용은 Essential Graphing 의 에서 확인하십시오 .
- **분위수** 및 **요약 통계량** 보고서 이러한 보고서에 대해서는 "데이터 분석" 장의 "분포 분석"(136 페이지) 에서 설명합니다 .

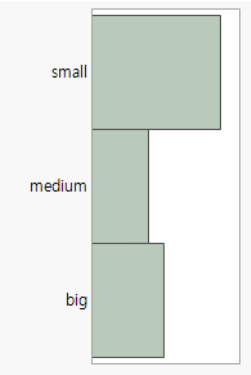
히스토그램에서의 상호 작용

JMP에서 데이터 테이블과 보고서는 모두 연결되어 있습니다. 히스토그램 막대를 클릭하면 데이터 테이블에서 해당 행이 선택됩니다.

범주형 변수에 막대 차트 사용

막대 차트를 사용하여 범주형 변수의 분포를 시각화할 수 있습니다. 막대 차트와 히스토그램은 둘 다 변수의 각 수준에 해당하는 막대가 있기 때문에 모양이 비슷합니다. 막대 차트는 변수의 모든 수준에 대한 막대를 표시하지만 히스토그램은 변수의 값 범위를 표시합니다.

그림 4.6 막대 차트의 예



시나리오

이 예에서는 특정 회사 그룹의 규모 및 유형과 관련된 데이터가 포함된 Companies.jmp 데이터 테이블을 사용합니다 .

재무 분석가가 다음과 같은 질문에 대한 답을 구하려고 합니다.

- 가장 일반적인 회사 유형은 무엇입니까 ?
- 가장 일반적인 회사 규모는 어떠합니까 ?

이러한 질문에 답하기 위해 Type 및 Size Co 의 막대 차트를 사용합니다.

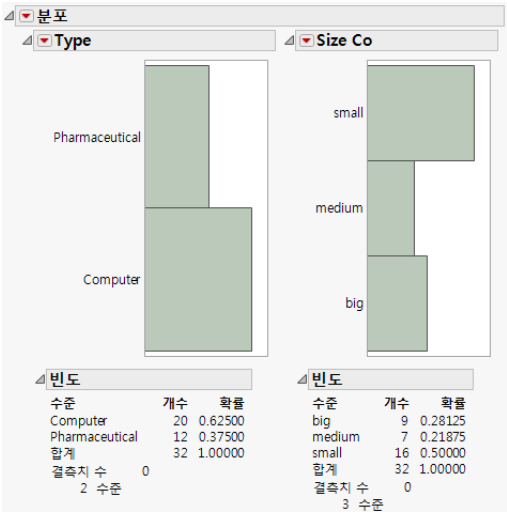
막대 차트 생성

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다.
2. **분석** > **분포**를 선택합니다.
3. Type 및 Size Co 를 선택하고 **Y, 열**을 클릭합니다.
4. **확인**을 클릭합니다.

그림 4.7 Type 및 Size Co 의 막대 차트

막대 차트

요약 정보



막대 차트 해석

이 막대 차트에서는 다음을 알 수 있습니다.

- 제약 회사보다 컴퓨터 회사가 많습니다.

  컴퓨터 회사를 나타내는 막대가 제약 회사를 나타내는 막대보다 큽니다.
- 가장 일반적인 회사 규모는 소규모입니다.

  소규모 회사를 나타내는 막대가 중간 규모 및 대규모 회사를 나타내는 막대보다 큽니다.

추가적인 요약 출력을 통해 상세한 빈도가 제공됩니다. 이 보고서에 대해서는 "데이터 분석" 장의 "범주형 변수의 분포"(139 페이지)에서 설명합니다.

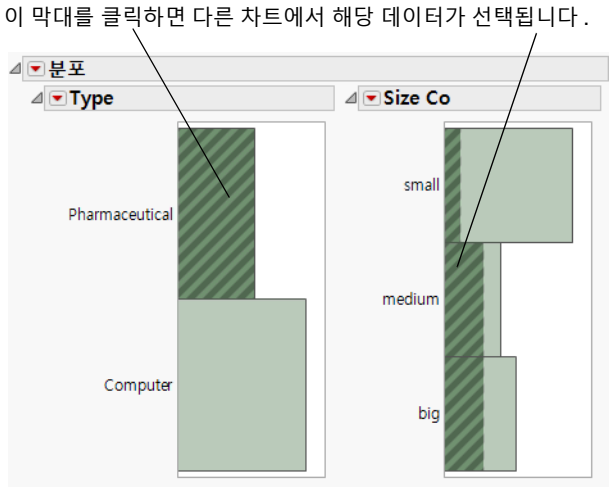
막대 차트에서의 상호 작용

히스토그램과 마찬가지로 개별 막대를 클릭하면 데이터 테이블의 해당 행이 강조 표시됩니다. 그래프가 두 개 이상 생성된 경우 한 막대 차트의 막대를 클릭하면 다른 막대 차트에서 그에 해당하는 막대가 강조 표시됩니다.

예를 들어 제약 회사의 회사 규모 분포를 보려고 한다고 가정해 보겠습니다. Type 막대 차트에서 제약 회사 막대를 클릭하면 Size Co 막대 차트에서 제약 회사가 강조 표시됩니다. 그림 4.8에서는 이 데이터 테이블에 있는 대부분의 회사가 소규모이지만 대부분의 제약 회사는 중간 규모 또는 대규모임을 보여 줍니다.

또한 데이터 테이블에서도 해당 행이 선택됩니다.

그림 4.8 막대 클릭



다중 변수 비교

다중 변수 그래프를 사용하여 둘 이상의 변수 사이에 존재하는 관계 및 패턴을 시각화할 수 있습니다. 이 섹션에서는 다음 그래프를 다룹니다.

표 4.1 다중 변수 그래프

"산점도를 사용하여 여러 변수 비교"(99 페이지)	산점도를 사용하면 두 개의 연속형 변수를 비교할 수 있습니다.
"산점도 행렬을 사용하여 여러 변수 비교"(102 페이지)	산점도 행렬을 사용하면 여러 쌍의 연속형 변수를 비교할 수 있습니다.
"병렬 상자 그림을 사용하여 여러 변수 비교"(105 페이지)	병렬 상자 그림을 사용하면 하나의 연속형 변수와 하나의 범주형 변수를 비교할 수 있습니다.
"계량형 차트를 사용하여 여러 변수 비교"(122 페이지)	계량형 차트를 사용하면 하나의 연속형 Y 변수를 하나 이상의 범주형 X 변수와 비교할 수 있습니다. 계량형 차트는 몇 가지 범주형 X 변수 간의 평균 및 변동 차이를 보여 줍니다.

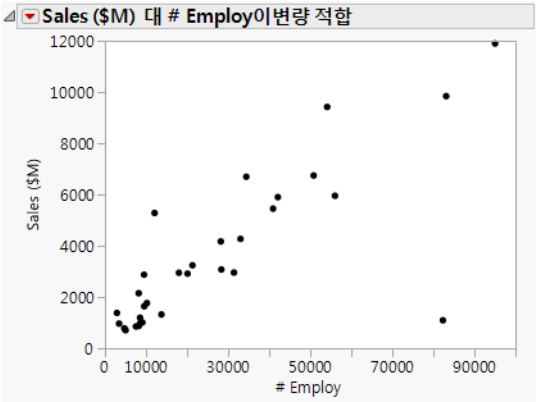
표 4.1 다중 변수 그래프 (계속)

"그래프 빌더를 사용하여 여러 변수 비교"(107 페이지)	그래프 빌더를 사용하면 대화식으로 그래프를 생성 및 변경할 수 있습니다.
"중첩 그림을 사용하여 여러 변수 비교"(118 페이지)	중첩 그림을 사용하면 Y 축에 있는 하나 이상의 변수를 X 축의 다른 변수와 비교할 수 있습니다. 중첩 그림에서는 두 개 이상의 변수가 시간에 따라 어떻게 변하는지 비교할 수 있기 때문에 X 변수가 시간 변수일 때 특히 유용합니다.
"버블 그림을 사용하여 여러 변수 비교"(113 페이지)	버블 그림은 색상 및 버블 크기를 사용하여 한 번에 최대 5개의 변수를 표시하는 특수한 산점도입니다. 변수 중 하나가 시간 변수이면 그림에 애니메이션을 적용하여 다른 변수가 시간에 따라 변화하는 것을 볼 수 있습니다.

산점도를 사용하여 여러 변수 비교

산점도는 모든 다중 변수 그래프 중에서 가장 단순합니다. 두 연속형 변수 간의 관계를 확인하고 두 연속형 변수 사이에 **상관관계**가 있는지 확인하려면 산점도를 사용합니다. 상관관계는 두 변수가 얼마나 밀접하게 관련되어 있는지를 나타냅니다. 상관관계가 높은 두 변수가 있으면 한 변수가 다른 변수에 영향을 미칠 수 있습니다. 또는 두 변수가 모두 비슷한 방식으로 다른 변수의 영향을 받을 수 있습니다.

그림 4.9 산점도의 예



시나리오

이 예에서는 특정 회사 그룹의 매출액과 직원 수가 포함된 Companies.jmp 데이터 테이블을 사용합니다 .

재무 분석가가 다음과 같은 질문에 대한 답을 구하려고 합니다.

- 매출과 직원 수 사이에 어떤 관계가 있습니까 ?

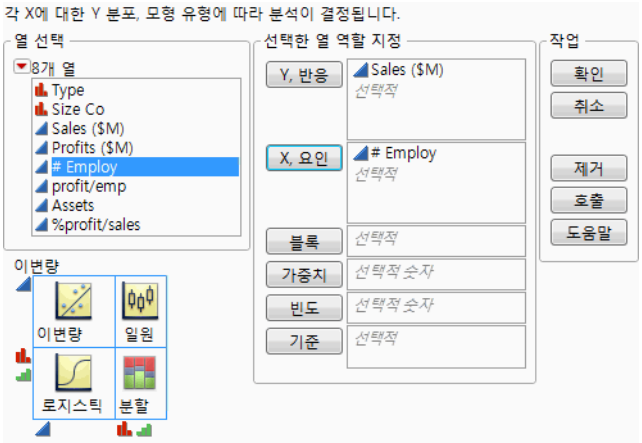
- 직원 수에 따라 매출액이 증가합니까 ?
- 직원 수를 토대로 평균 매출을 예측할 수 있습니까 ?

이러한 질문에 답하려면 Sales ($M) 대 # Employ 산점도를 사용하십시오.

산점도 생성

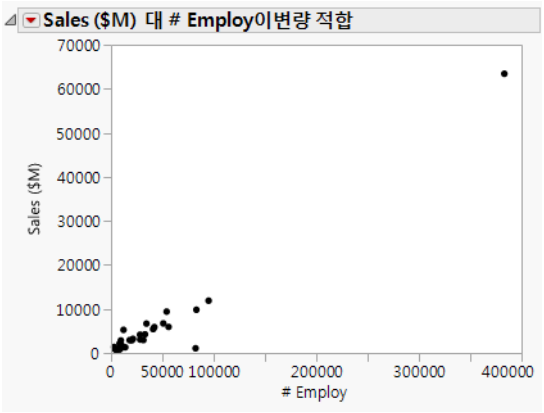
1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다 .
2. **분석 > X 로 Y 적합**을 선택합니다 .
3. Sales ($M) 를 선택하고 **Y, 반응**을 클릭합니다 .
4. # Employ 를 선택하고 **X, 요인**을 클릭합니다 .

그림 4.10 X 로 Y 적합 창



5. **확인**을 클릭합니다 .

그림 4.11 Sales ($M) 대 # Employ 산점도



산점도 해석

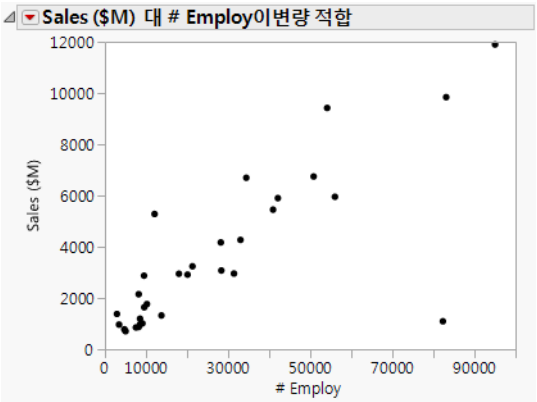
한 회사는 직원 수가 매우 많고 매출이 상당히 높습니다 . 이 회사는 그림 오른쪽 상단에 단일 점으로 표시됩니다 . 이 데이터 점과 나머지 모든 점 사이에는 상당한 거리가 있어 나머지 회사 간의 관계를 시각화하기가 어렵습니다 . 다음 단계에 따라 그림에서 해당 점을 제거하고 그림을 다시 생성하십시오 .

1. 해당 점을 클릭하여 선택합니다 .
2. **행 > 숨기고 제외하기**를 선택합니다 . 해당 데이터 점이 숨겨지고 더 이상 계산에 포함되지 않습니다 .

   참고 : 숨기기와 제외 사이에는 중요한 차이가 있습니다 . 점을 숨기면 해당 점이 모든 그래프에서 제거되지만 통계 계산에는 계속 사용됩니다 . 점을 제외하면 해당 점이 통계 계산에서 제거되지만 그래프에서는 제거되지 않습니다 . 점을 숨기고 제외하면 모든 계산과 모든 그래프에서 해당 점이 제거됩니다 .

3. 이상치를 제외하고 그림을 다시 생성하려면 " 이변량 " 의 빨간색 삼각형을 클릭하고 **다시 실행 > 분석 다시 실행**을 선택합니다 . 원래 보고서 창은 닫아도 됩니다 .

그림 4.12 이상치가 제거된 산점도



업데이트된 산점도에서는 다음을 알 수 있습니다 .

- 매출과 직원 수 사이에는 관계가 있습니다 .

  데이터 점에 눈에 띄는 패턴이 있습니다 . 데이터 점은 그래프 전체에 무작위로 흩어져 있지 않습니다 . 대부분의 데이터 점에 근접하는 주대각선을 그릴 수 있습니다 .

- 직원 수에 따라 매출이 늘어나며 이 관계는 선형입니다 .

  주대각선을 그린다면 왼쪽 하단에서 오른쪽 상단으로 향하는 기울기를 보일 것입니다. 이 기울기는 직원 수가 증가함에 따라 (하단 축의 왼쪽에서 오른쪽으로) 매출 또한 증가함 (왼쪽 축의 아래쪽에서 위쪽으로) 을 보여 줍니다 . 대부분의 데이터 점이 직선 근처에 위치하므로 선형 관계를 나타냅니다 . 데이터 점에 근접하는 선을 곡선으로만 그릴 수 있는 경우라도 점의 패턴 때문에 관계가 유지됩니다 . 그러나 이 관계는 선형적이지 않습니다 .

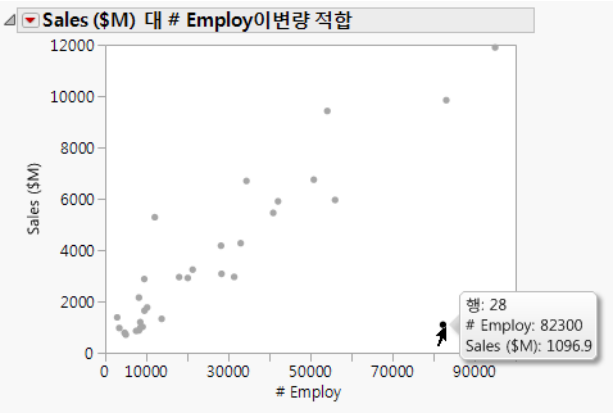
- 직원 수를 토대로 평균 매출을 예측할 수 있습니다.

  이 산점도는 일반적으로 직원 수가 증가할수록 매출도 증가한다는 것을 보여 줍니다. 따라서 회사의 직원 수만 알고 있으면 회사의 매출을 예측할 수 있습니다. 이러한 예측점은 가상의 선 위에 있을 것입니다. 예측은 정확하지는 않겠지만 실제 매출과 비슷할 것입니다.

산점도에서의 상호 작용

산점도는 다른 JMP 그래픽과 마찬가지로 대화식입니다. 오른쪽 하단의 점을 마우스 커서로 가리키면 행 번호와 x 및 y 값이 표시됩니다.

그림 4.13 커서로 점 가리키기



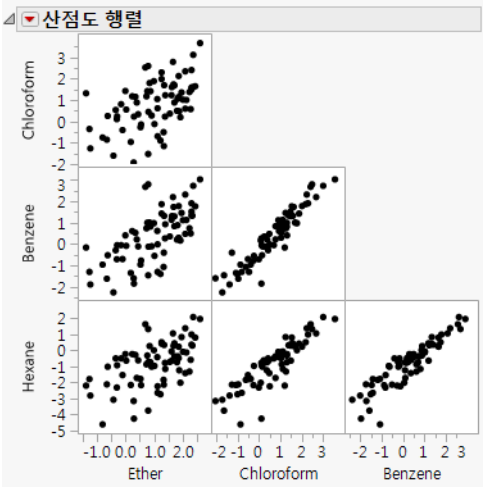
점을 클릭하면 데이터 테이블에서 해당 행이 강조 표시됩니다. 여러 점을 선택하려면 다음 중 하나를 수행하십시오.

- 점 주변에 커서를 놓고 클릭하여 드래그합니다. 그러면 사각형 영역의 점이 선택됩니다.
- 올가미 도구를 선택한 후 여러 점 주변을 클릭하여 드래그합니다 . 그러면 모양이 불규칙한 영역이 선택됩니다.

산점도 행렬을 사용하여 여러 변수 비교

산점도 행렬은 격자(또는 행렬)로 구성된 산점도 모음입니다. 각 산점도는 한 쌍의 변수 사이에 있는 관계를 보여 줍니다.

그림 4.14 산점도 행렬의 예



시나리오

이 예에서는 72가지 용질의 용해도 측정 데이터가 들어 있는 Solubility.jmp 데이터 테이블을 사용합니다 .

실험실 기술자가 다음과 같은 질문에 대한 답을 구하려고 합니다.

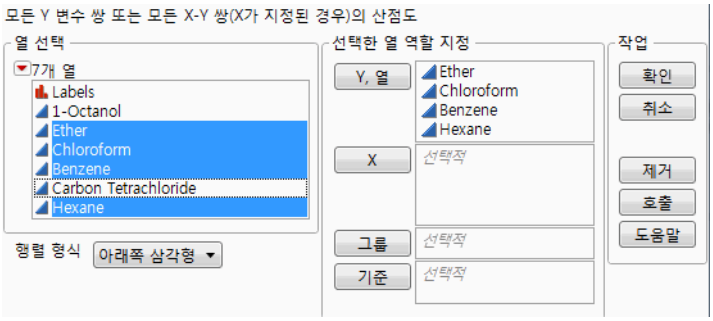
- 관련성이 있는 화학 물질 쌍이 있습니까 ? 가능한 6 개의 쌍이 있습니다 .
- 어떤 쌍의 관련성이 가장 큽니까 ?

이 질문에 답하려면 네 가지 용매의 산점도 행렬을 사용하십시오.

산점도 행렬 생성

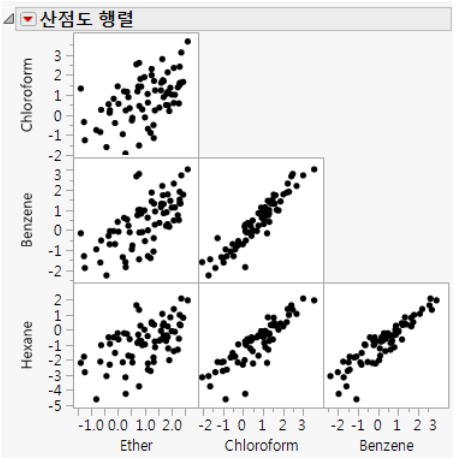
1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Solubility.jmp 를 엽니다 .
2. **그래프** > **산점도 행렬**을 선택합니다 .
3. Ether, Chloroform, Benzene 및 Hexane 을 선택하고 **Y, 열**을 클릭합니다 .

그림 4.15 산점도 행렬 창



4. **확인**을 클릭합니다.

그림 4.16 산점도 행렬



산점도 행렬 해석

이 산점도 행렬에서는 다음을 알 수 있습니다.

- 6 개의 변수 쌍은 모두 양의 상관관계에 있습니다.

  한 변수가 증가하면 다른 변수도 증가합니다.
- Benzene 과 Chloroform 사이에 가장 강한 관계가 있는 것으로 나타납니다.

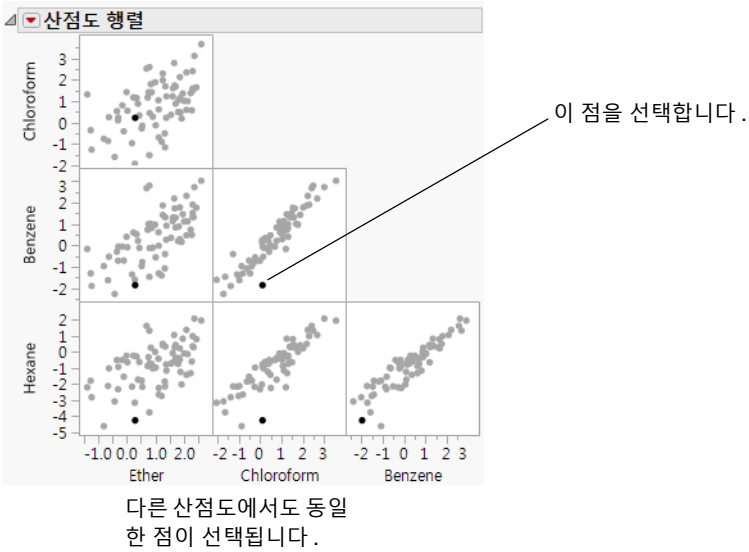
  Benzene 과 Chloroform 의 산점도에 있는 데이터 점은 가상의 선을 따라 가장 조밀하게 밀집되어 있습니다.

산점도 행렬에서의 상호 작용

한 산점도에서 점을 선택하면 다른 모든 산점도에서도 해당 점이 선택됩니다.

예를 들어 Benzene 대 Chloroform 산점도에서 한 점을 선택하면 다른 다섯 개의 산점도에서 동일한 점이 선택됩니다.

그림 4.17 선택된 점

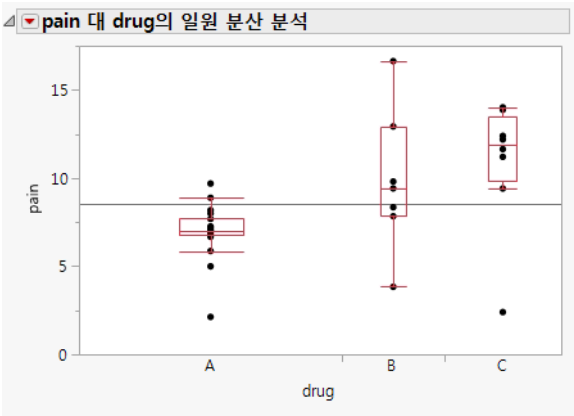


병렬 상자 그림을 사용하여 여러 변수 비교

병렬 상자 그림은 다음을 보여 줍니다.

- 연속형 변수 하나와 범주형 변수 하나 사이의 관계
- 범주형 변수의 전체 수준에서 나타나는 연속형 변수의 차이

그림 4.18 병렬 상자 그림의 예



시나리오

이 예에서는 세 가지 약물의 투약 시 환자 통증 측정 데이터가 포함된 Analgesics.jmp 데이터 테이블을 사용합니다.

연구원이 다음과 같은 질문에 대한 답을 구하려고 합니다 .

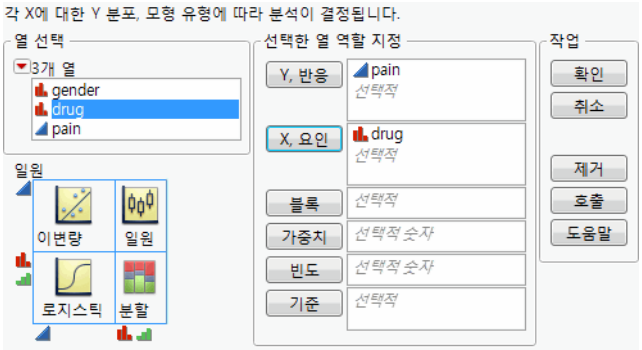
- 약물의 평균적인 통증 조절 수준 간에 차이가 있습니까 ?
- 약물마다 통증 조절 수준의 **변동성**이 다릅니까 ? 변동성이 높은 약물은 변동성이 낮은 약물보다 신뢰도가 떨어집니다 .

이러한 질문에 답하려면 통증 수준과 약물 범주에 대한 병렬 상자 그림을 사용하십시오.

병렬 상자 그림 생성

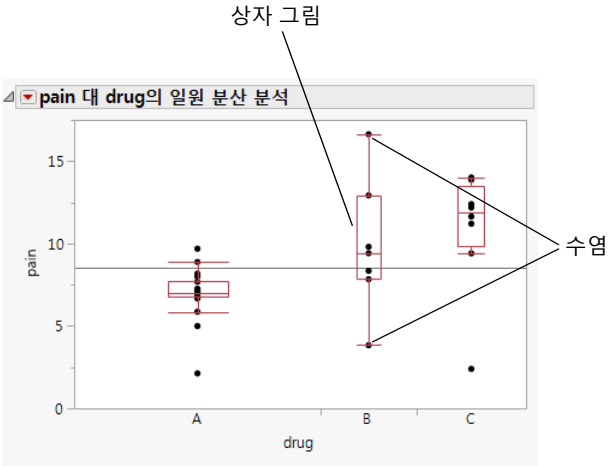
1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Analgesics.jmp 를 엽니다 .
2. **분석** > **X 로 Y 적합**을 선택합니다 .
3. pain 을 선택하고 **Y, 반응**을 클릭합니다 .
4. drug 를 선택하고 **X, 요인**을 클릭합니다 .

그림 4.19 X 로 Y 적합 창



5. **확인**을 클릭합니다 .
6. "pain 대 drug 의 일원 분석 " 옆의 빨간색 삼각형을 클릭하고 **표시 옵션 > 상자 그림**을 선택합니다 .

그림 4.20 병렬 상자 그림



병렬 상자 그림 해석

상자 그림은 다음과 같은 원리에 따라 만들어졌습니다.

- 상자 안의 선은 중앙값을 나타냅니다.
- 데이터의 가운데 절반은 상자 안에 있습니다.
- 대부분의 데이터는 양쪽 수염 끝 사이에 있습니다.
- 수염을 벗어나는 데이터 점은 이상치일 수 있습니다.

그림 4.20 의 상자 그림에서는 다음을 알 수 있습니다.

- 약물 A 의 상자 그림이 다른 약물보다 통증 수준이 낮기 때문에 약물 A 환자의 통증이 덜하다고 믿을 만한 근거가 됩니다.
- 약물 B 는 약물 A 및 C 보다 상자 그림이 더 길기 때문에 변동성이 높은 것으로 보입니다.

약물 C의 한 점은 약물 C의 다른 점보다 훨씬 낮습니다. 이 점을 커서로 가리키면 데이터 테이블의 26 행에 해당하는 것을 알 수 있습니다. 이 점은 약물 그룹 A 또는 B의 데이터와 더 유사해 보이므로 26 행의 정보는 조사할 만합니다. 데이터를 기록할 때 잘못 적었을 수도 있습니다.

그래프 빌더를 사용하여 여러 변수 비교

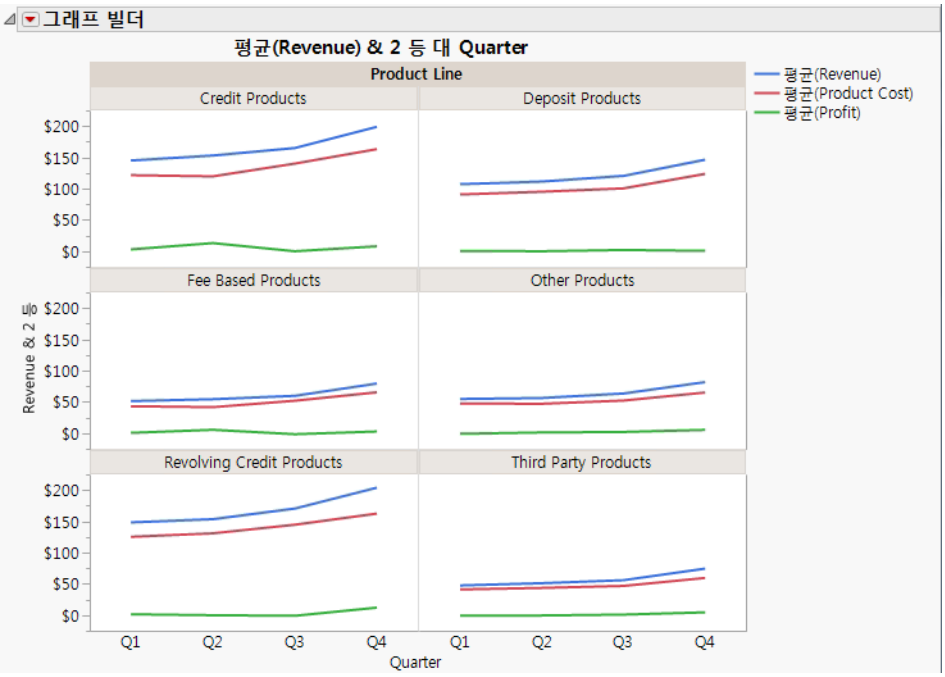
그래프 빌더를 사용하여 대화식으로 그래프를 생성하고 수정할 수 있습니다. JMP 의 그래프는 대부분 플랫폼을 시작하고 변수를 지정하는 방법으로 생성됩니다. 다른 유형의 그래프를 생성하려면 "그래프" 메뉴에서 특정 플랫폼을 시작합니다. 하지만 그래프 빌더를 사용하면 언제든지 변수를 변경하고 그래프 유형을 변경할 수 있습니다.

그래프 빌더를 사용하여 다음 작업을 수행할 수 있습니다.

- 변수를 그래프 안팎으로 드래그하여 놓는 방법으로 변경합니다.

- 몇 번의 마우스 클릭으로 다른 유형의 그래프를 생성합니다 .
- 그래프를 가로 또는 세로로 분할합니다 .

그림 4.21 그래프 빌더로 생성된 그래프의 예



참고 : 여기에서는 그래프 빌더의 일부 기능만 다룹니다 . 자세한 내용은 Essential Graphing 의 에서 확인하십시오 .

시나리오

이 예에서는 여러 제품 라인의 수익 데이터가 포함된 Profit by Product.jmp 데이터 테이블을 사용합니다.

비즈니스 분석가가 다음과 같은 질문에 대한 답을 구하려고 합니다 .

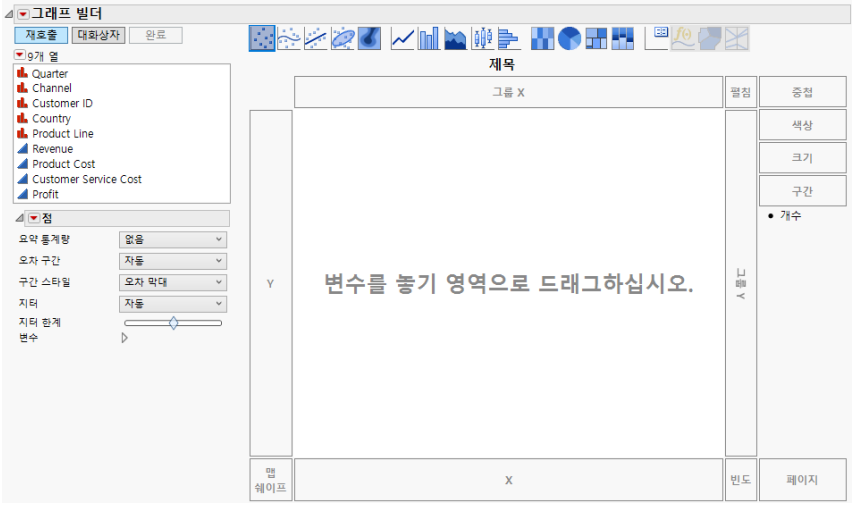
- 제품 라인 간의 수익성이 어떻게 다릅니까 ?

이 질문에 답하려면 다양한 제품 라인의 매출, 제품 비용 및 수익 데이터를 표시하는 선 그림을 사용하십시오.

그래프 생성

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Profit by Product.jmp 를 엽니다 .
2. **그래프** > **그래프 빌더**를 선택합니다 .

그림 4.22 그래프 빌더 작업 공간

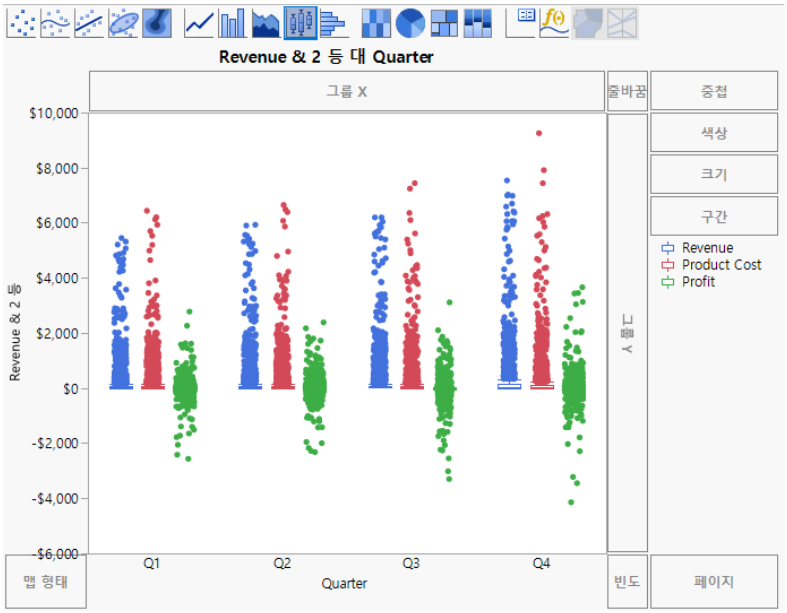


3. Quarter 를 클릭한 후 X 영역으로 드래그하여 놓음으로써 Quarter 를 X 변수로 할당합니다 .
4. Revenue, Product Cost 및 Profit 을 클릭한 후 Y 영역으로 드래그하여 놓음으로써 이 세 변수를 모두 Y 변수로 할당합니다 .

   이제 X 및 Y 영역이 축이 되었습니다 .

참고 : 변수를 클릭한 후 영역을 클릭하여 할당할 수도 있습니다 . 그러나 영역이 축이 된 후에는 변수와 축을 클릭하는 대신 추가 변수를 축으로 드래그하여 놓으십시오 .

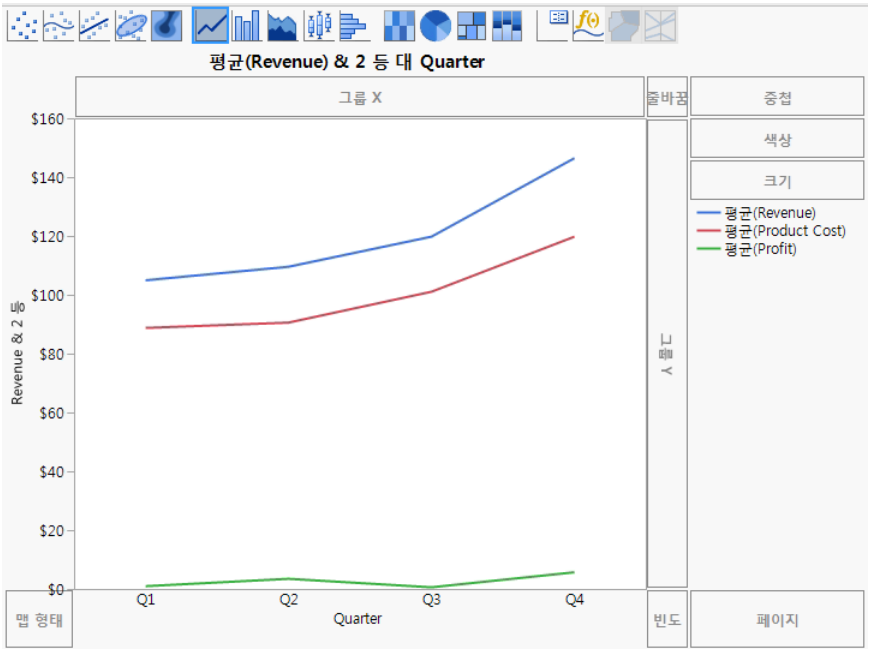
그림 4.23 Y 및 X 변수 추가 후



그래프 빌더는 사용하고 있는 변수에 따라 병렬 상자 그림을 표시합니다.

5. 상자 그림을 선 그림으로 변경하려면 선 아이콘 을 클릭합니다.

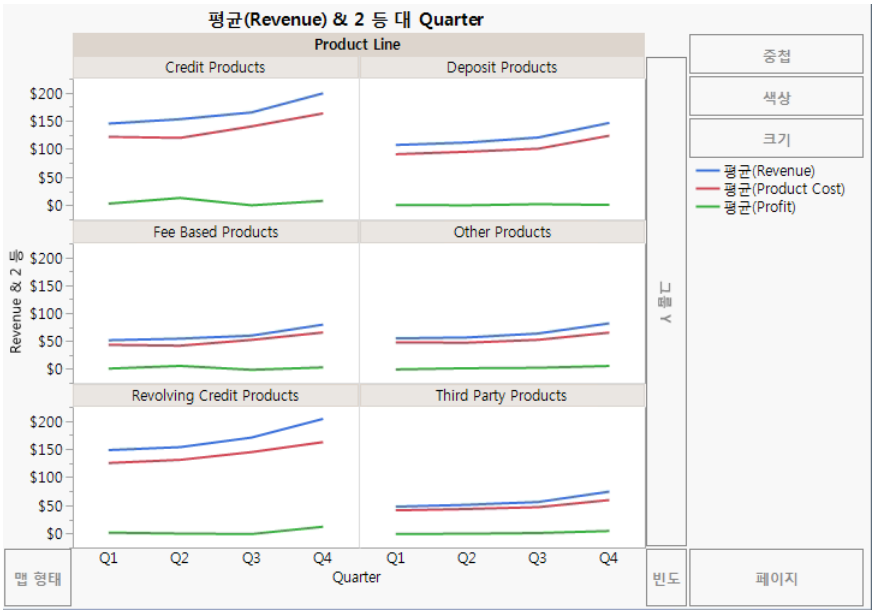
그림 4.24 선 그림



6. 각 제품에 대해 개별 차트를 생성하려면 Product Line 을 클릭한 후 **펼침** 영역으로 드래그하여 놓습니다.

   각 제품에 대해 개별 선 그림이 생성됩니다.

그림 4.25 최종 선 그림



그래프 해석

그림 4.25 에서는 제품 라인별로 분류된 매출 , 비용 및 수익을 보여 줍니다 . 비즈니스 분석가는 제품 라인 간의 수익성 차이를 확인하는 데 관심이 있었습니다 . 그림 4.25 의 선 그림을 통해 다음과 같은 몇 가지 사항을 확인할 수 있습니다 .

- 신용 상품 , 예금 상품 및 회전 신용 상품이 수수료 기반 상품 , 타사 상품 및 기타 상품보다 많은 매출을 올립니다 .
- 그러나 모든 제품 라인의 수익은 비슷합니다 .

이 데이터 테이블에는 판매 채널에 대한 데이터도 포함되어 있습니다 . 비즈니스 분석가는 매출 , 제품 비용 및 수익이 서로 다른 판매 채널 간에 어떻게 다른지 확인하고자 합니다 .

1. 그래프에서 Product Line 을 제거하려면 그래프의 제목 (Product Line) 을 클릭하고 그래프 빌더 안의 빈 영역으로 드래그하여 놓습니다 .
2. Channel을 펼침 변수로 추가하려면 Channel을 클릭하고 **펼침** 영역으로 드래그하여 놓습니다.

그림 4.26 판매 채널을 보여주는 선 그림

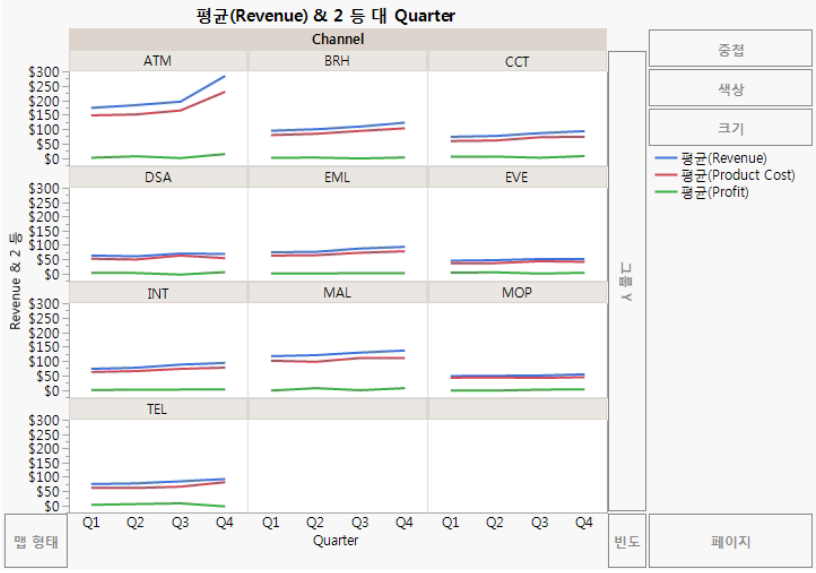
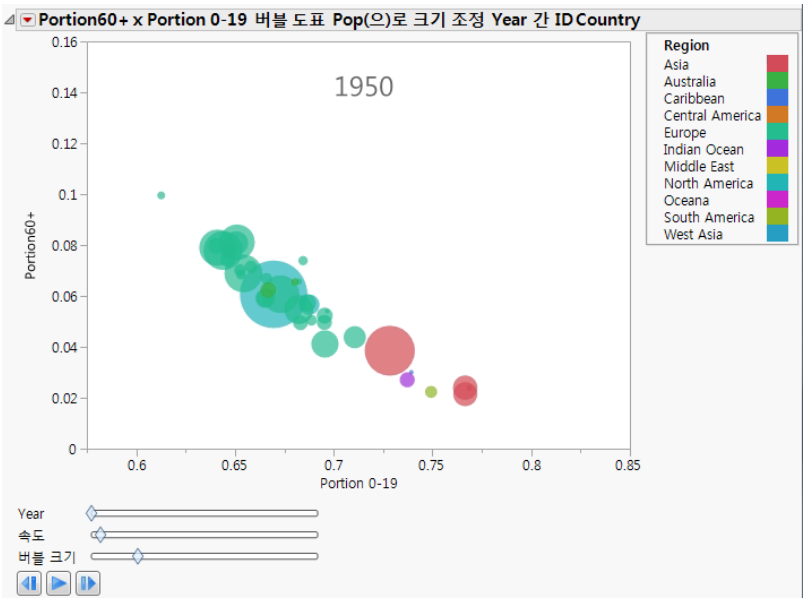


그림 4.26에서는 ATM의 매출 및 제품 비용이 가장 높으며 가장 빠르게 증가하고 있음을 알 수 있습니다.

버블 그림을 사용하여 여러 변수 비교

버블 그림은 점을 버블로 나타내는 산점도입니다. 버블의 크기와 색상을 변경하고 시간 경과에 따라 애니메이션도 적용할 수 있습니다. 버블 그림은 최대 5개의 차원(x 위치, y 위치, 크기, 색상 및 시간)을 나타낼 수 있기 때문에 인상적인 시각적 표현을 생성하며 이를 통해 데이터를 손쉽게 탐색할 수 있습니다.

그림 4.27 버블 그림의 예



시나리오

이 예에서는 1950 년부터 2004 년까지 116 개 국가 또는 지역의 인구 통계가 포함된 PopAgeGroup.jmp 데이터 테이블을 사용합니다. 총 인구 수는 연령대별로 분류되며, 모든 국가의 데이터가 매년 있지는 않습니다.

사회학자가 다음과 같은 질문에 대한 답을 구하려고 합니다.

- 세계 인구의 연령대가 바뀌고 있습니까?

이 질문에 답하려면 인구 중 가장 높은 연령대(60 세 이상)와 가장 낮은 연령대(20 세 미만) 간의 관계를 살펴보아야 합니다. 버블 그림을 사용하여 이 관계가 시간에 따라 어떻게 변하는지 확인하십시오.

버블 그림 생성

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 PopAgeGroup.jmp 를 엽니다.
2. **그래프** > **버블 그림**을 선택합니다.
3. Portion60+ 를 선택하고 **Y** 를 클릭합니다.

   이 열은 버블 그림의 *Y* 변수에 해당합니다.
4. Portion 0-19 를 선택하고 **X** 를 클릭합니다.

   이 열은 버블 그림의 *X* 변수에 해당합니다.
5. Country 를 선택하고 **ID** 를 클릭합니다.

"ID" 변수의 고유 수준이 각각 그림에 버블로 표시됩니다 .

6. Year 를 선택하고 **시간**을 클릭합니다 .

   버블 그림에 애니메이션이 적용될 때 이 열에 따라 시간 인덱싱이 제어됩니다 .

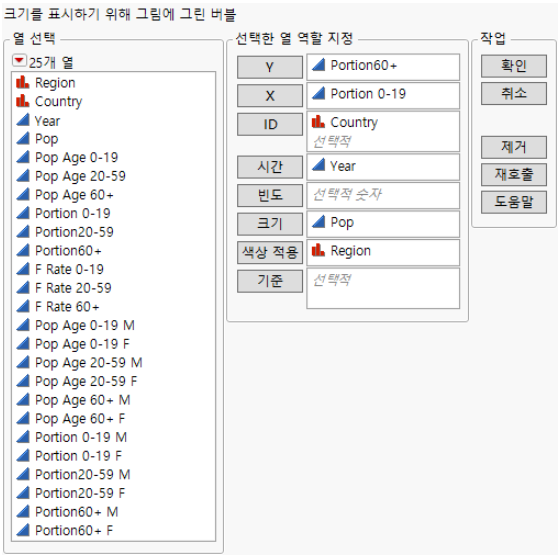
7. Pop 을 선택하고 **크기**를 클릭합니다 .

   이 열에 따라 버블의 크기가 제어됩니다 .

8. Region 을 선택하고 **색상 적용**을 클릭합니다 .

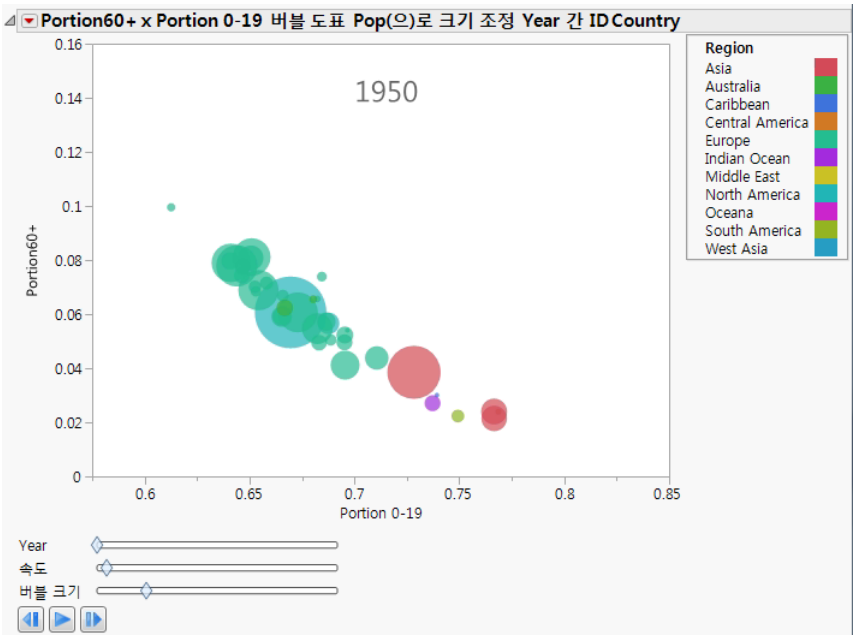
   " 색상 적용 " 변수의 각 수준에 고유한 색상이 할당됩니다 . 따라서 이 예에서 같은 지역에 위치한 국가의 모든 버블은 색상이 동일합니다 . 그림 4.29 에 나타나는 버블 색상은 JMP 기본 색상입니다 .

그림 4.28 버블 그림 시작 창



9. **확인**을 클릭합니다 .

그림 4.29 초기 버블 그림



버블 그림 해석

시간 변수 (이 예의 경우 연도) 가 1950 년부터 시작되기 때문에 초기 버블 그림은 1950 년의 데이터를 보여 줍니다 . 재생 / 일시 중지 버튼을 클릭하여 모든 연도를 순환하도록 버블 그림에 애니메이션을 적용하십시오. 각각의 연이은 버블 그림은 해당 연도의 데이터를 보여 줍니다. 각 연도의 데이터에 따라 다음이 결정됩니다 .

- *X* 및 *Y* 좌표
- 버블 크기
- 버블 색상
- 버블 집계

참고 : 버블 그림에서 여러 행의 정보가 집계되는 방식에 대한 자세한 내용은 Essential Graphing 의 에서 확인하십시오 .

1950 년의 버블 그림에서는 국가의 20 세 미만 인구 비율이 높으면 60 세 이상 인구 비율이 낮다는 것을 보여 줍니다.

버블 그림의 전체 연도 범위에 애니메이션을 적용하려면 재생 / 일시 중지 버튼을 클릭하십시오 . 시간이 지나면서 Portion 0-19 의 비율이 감소하고 Portion60+ 의 비율은 증가합니다 .

애니메이션을 재생하며 , 클릭 후에는 일시 중지 버튼으로 바뀝니다 .

애니메이션을 일시 중지합니다 .

애니메이션을 시간 단위 하나만큼 뒤로 수동 제어합니다 .

애니메이션을 시간 단위 하나만큼 앞으로 수동 제어합니다 .

Year 시간 인덱스를 수동으로 변경합니다 .

속도 애니메이션의 속도를 제어합니다 .

버블 크기 상대 크기를 유지하면서 버블의 절대 크기를 제어합니다 .

사회학자는 세계 인구의 연령이 어떻게 변화하고 있는지 알고 싶어했습니다. 이 버블 그림은 세계 인구가 점점 고령화되고 있음을 보여 줍니다.

버블 그림에서의 상호 작용

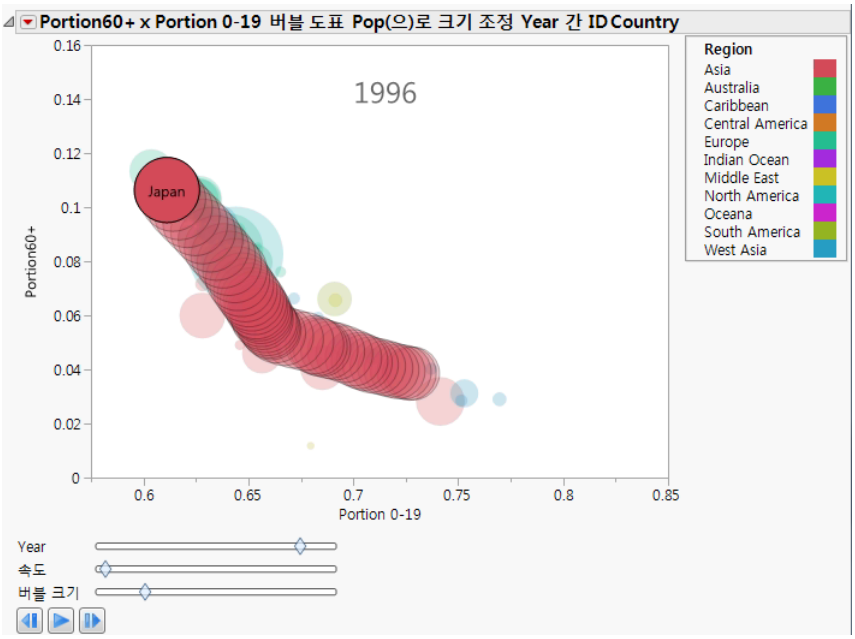
버블을 클릭하여 선택하면 시간 경과에 따른 버블의 추세를 확인할 수 있습니다. 예를 들어 1950 년 그림에서 중간에 있는 큰 버블은 일본의 데이터입니다.

수년에 걸친 일본의 인구 변화 패턴을 보려면 다음을 수행하십시오 .

1. 일본 버블의 가운데를 클릭하여 선택합니다 .
2. " 버블 그림 " 의 빨간색 삼각형을 클릭하고 **이동 경로 버블 > 선택됨**을 선택합니다 .
3. 재생 버튼을 클릭합니다 .

애니메이션이 시간이 지남에 따라 진행되면서 일본 버블은 그 기록을 보여 주는 자취를 남깁니다.

그림 4.30 일본의 인구 변동 기록



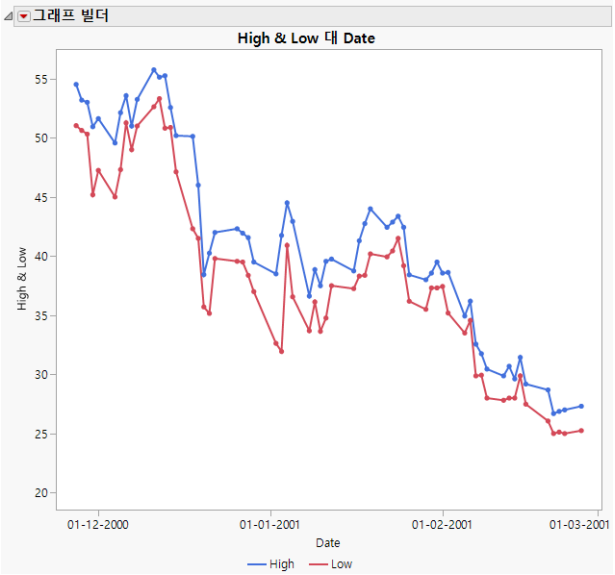
일본 버블에 초점을 맞추면 시간 경과에 따라 다음 정보를 확인할 수 있습니다 .

- 19 세 이하 인구 비율이 감소했습니다 .
- 60 세 이상 인구 비율이 증가했습니다 .

중첩 그림을 사용하여 여러 변수 비교

중첩 그림은 산점도와 마찬가지로 두 개 이상의 변수 사이에 존재하는 관계를 보여 줍니다. 그러나 변수 중 하나가 시간 변수인 경우에는 산점도보다 중첩 그림이 시간에 따른 추세를 더 잘 나타냅니다.

그림 4.31 중첩 그림의 예



참고 : 시간에 따른 데이터 그림을 생성하려는 경우 버블 그림 , 관리도 및 계량형 차트를 사용할 수도 있습니다 . 그래프 빌더와 버블 그림에 대한 자세한 내용은 Essential Graphing 의 에서 확인하십시오 . 관리도와 계량형 차트에 대한 자세한 내용은 Quality and Process Methods 의 및 에서 확인하십시오 .

시나리오

이 예에서는 3 개월 동안의 주가에 대한 데이터가 포함된 Stock Prices.jmp 데이터 테이블을 사용합니다.

잠재 투자자가 다음과 같은 질문에 대한 답을 구하려고 합니다 .

- 지난 3 개월 동안 주식 종가에 변화가 있습니까 ?

  이 질문에 답하려면 시간 경과에 따른 주식 종가의 중첩 그림을 사용해야 합니다 .
- 주식의 고가와 저가는 서로 어떤 관련이 있습니까 ?

  이 질문에 답하려면 시간 경과에 따른 주식의 고가 및 저가를 나타내는 또 다른 중첩 그림을 사용해야 합니다 .

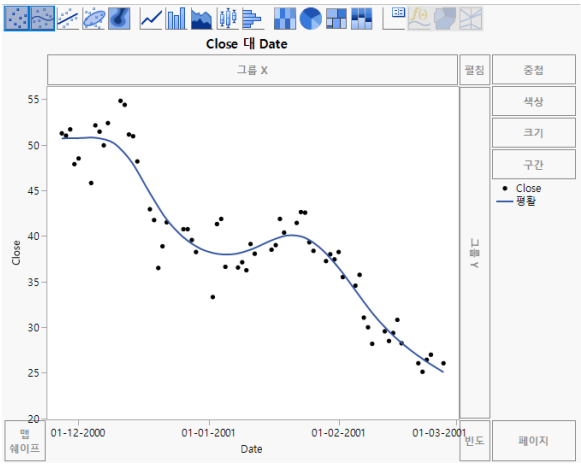
첫 번째 질문에 답하기 위해 첫 번째 중첩 그림을 생성한 후 두 번째 질문에 답하기 위해 두 번째 중첩 그림을 생성합니다.

시간 경과에 따른 주가의 중첩 그림 생성

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Stock Prices.jmp 를 엽니다 .

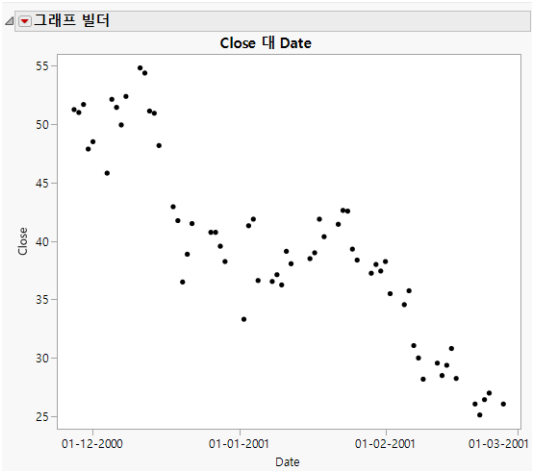
2. **그래프 > 그래프 빌더**를 선택합니다 .
3. Close 를 선택하고 **Y** 를 클릭합니다 .
4. Date 를 선택하고 **X** 를 클릭합니다 .

그림 4.32 평활기를 사용한 중첩 그림



5. Ctrl 키를 누르고 그래프 위의 평활기 아이콘 을 클릭하여 평활기 선을 제거합니다 .

그림 4.33 시간에 따른 종가의 중첩 그림

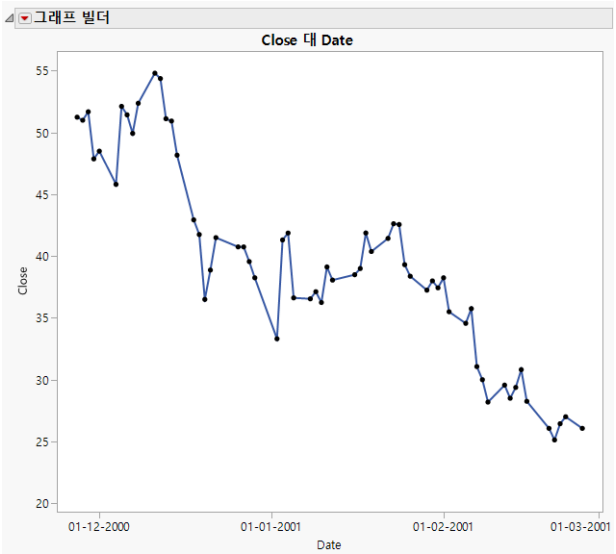


중첩 그림 해석 및 상호 작용

이 중첩 그림에서는 주식 종가가 지난 몇 개월 동안 하락하고 있음을 보여 줍니다. 추세를 더 명확하게 보려면 점을 연결하십시오.

1. Shift 키를 누르고 그래프 위의 선 아이콘 을 클릭합니다.

그림 4.34 연결된 점



이로써 잠재 투자자는 지난 3 개월 동안 주가가 오르내렸지만 전반적으로는 하락 추세였다는 것을 알 수 있습니다.

주식의 고가 및 저가에 대한 중첩 그림 생성

중첩 그림을 사용하여 둘 이상의 Y 변수에 대한 그림을 생성할 수 있습니다. 예를 들어 고가와 저가를 모두 동일한 그림에서 보고 싶다고 가정해 보겠습니다.

1. " 시간 경과에 따른 주가의 중첩 그림 생성 "(119 페이지) 의 단계를 따르되, 이번에는 **Y** 역할에 High 와 Low 를 할당합니다.
2. " 중첩 그림 해석 및 상호 작용 "(121 페이지) 에 표시된 대로 점을 연결하고 격자선을 추가합니다.

그림 4.35 두 개의 Y 변수

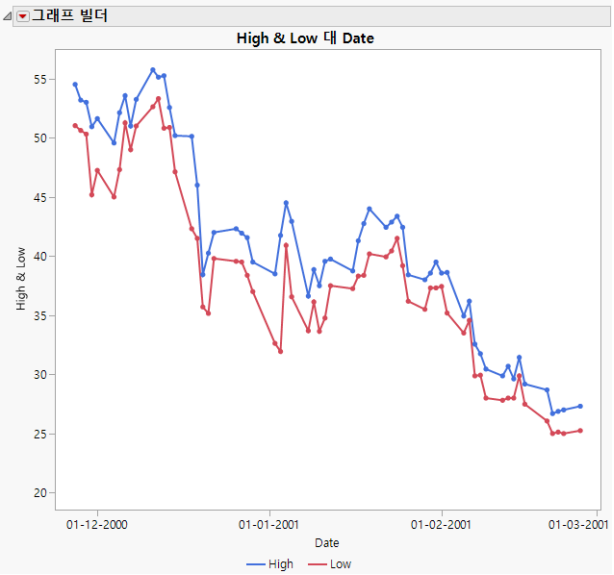


그림 하단의 범례는 그래프에서 High 및 Low 변수에 사용된 색상과 표식을 보여 줍니다. 이 중첩 그림은 High 가격과 Low 가격이 서로 매우 밀접하게 얽혀 있음을 보여 줍니다.

질문에 대한 답

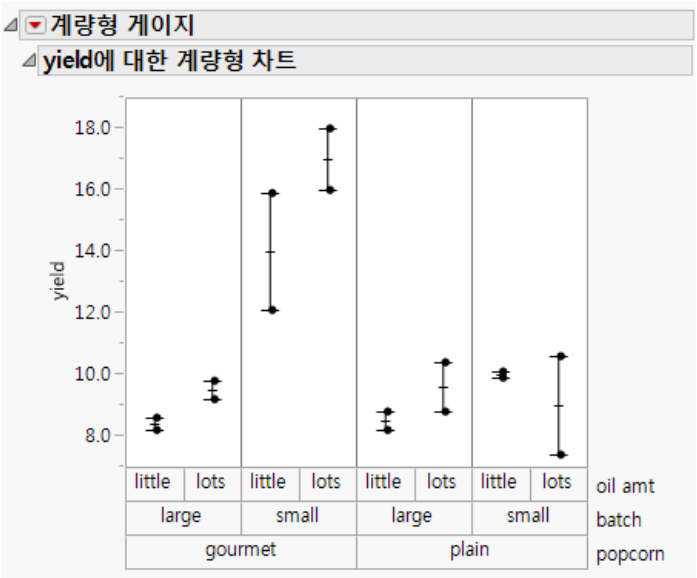
두 중첩 그림은 이 예의 시작 부분에서 제기했던 두 가지 질문에 대한 답을 제공합니다.

- 첫 번째 그림은 이 주식의 가격이 같은 수준으로 유지되지 않고 하락해 왔음을 보여 줍니다.
- 두 번째 그림은 이 주식의 고가와 저가 사이에 그다지 차이가 없다는 것을 보여 줍니다. 즉, 일간 주가 변동이 그리 크지 않습니다.

계량형 차트를 사용하여 여러 변수 비교

지금까지 설명한 그래프에서는 단일 X 변수만 지정했습니다. 계량형 차트를 사용하면 여러 개의 X 변수를 지정하고 모든 변수의 평균 및 변동성 차이를 한 번에 볼 수 있습니다.

그림 4.36 계량형 차트의 예



시나리오

이 예에서는 팝콘 제조업체의 데이터가 포함된 Popcorn.jmp 데이터 테이블을 사용합니다. 이 테이블에는 팝콘 스타일, 배치 크기 및 기름 사용량의 각 조합에 따른 산출량(주어진 낟알 측정값에 대한 팝콘 부피)을 측정한 결과가 있습니다.

팝콘 제조업체에서 다음과 같은 질문에 대한 답을 구하려고 합니다.

- 어떤 요인의 조합에서 팝콘 산출량이 가장 높습니까?

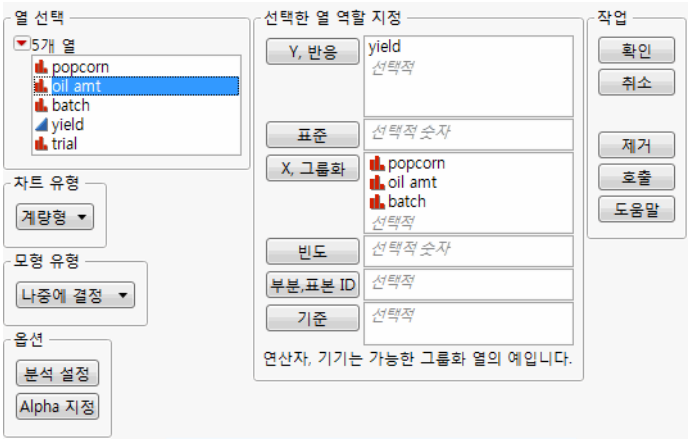
이 질문에 답하려면 산출량과 스타일, 배치 크기 및 기름 양의 관계를 보여 주는 계량형 차트를 사용해야 합니다.

계량형 차트 생성

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Popcorn.jmp 를 엽니다.
2. **분석 > 품질 및 공정 > 계량형 / 계수형 게이지 차트**를 선택합니다.
3. yield 를 선택하고 **Y, 반응**을 클릭합니다.
4. popcorn 을 선택하고 **X, 그룹화**를 클릭합니다.
5. batch 를 선택하고 **X, 그룹화**를 클릭합니다.
6. oil amt 을 선택하고 **X, 그룹화**를 클릭합니다.

참고 : 이 창에서의 순서에 따라 계량형 차트의 내포 순서가 결정되므로 변수를 **X, 그룹화** 역할에 할당하는 순서가 중요합니다.

그림 4.37 계량형 차트 창

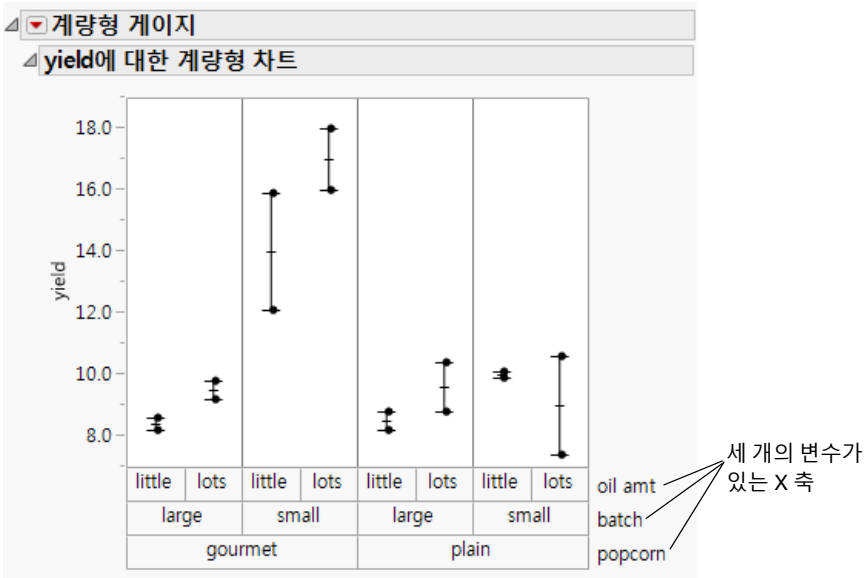


7. **확인**을 클릭합니다.

   위쪽 차트는 계량형 차트로, 세 변수의 조합에 따라 분류된 산출량을 보여 줍니다. 아래쪽 차트는 세 변수의 각 조합에 대한 표준편차를 보여 줍니다. 아래쪽 차트는 산출량을 표시하지 않으므로 숨겨 둡니다.

8. " 계량형 게이지 " 의 빨간색 삼각형을 클릭하고 **S 관리도**를 선택 취소합니다.

그림 4.38 결과 창



계량형 차트 해석

산출량의 계량형 차트는 작은 고메이 배치의 산출량이 가장 높음을 나타냅니다 .

팝콘 제조업체는 보다 구체적으로 다음과 같은 추가 질문을 제기할 수 있습니다. 산출량이 높은 이유가 이 배치가 작기 때문입니까, 이 배치가 고메이이기 때문입니까?

계량형 차트는 다음을 보여 줍니다 .

- 작은 플레인 배치의 산출량이 낮습니다 .
- 큰 고메이 배치의 산출량이 낮습니다 .

이 정보를 감안할 때 팝콘 제조업체는 작은 크기와 고메이 스타일의 조합에서 배치의 산출량이 높다고 결론 내릴 수 있습니다. 단일 변수만 허용하는 차트로는 이 결론에 도달하는 것이 불가능했을 것입니다.

5 장

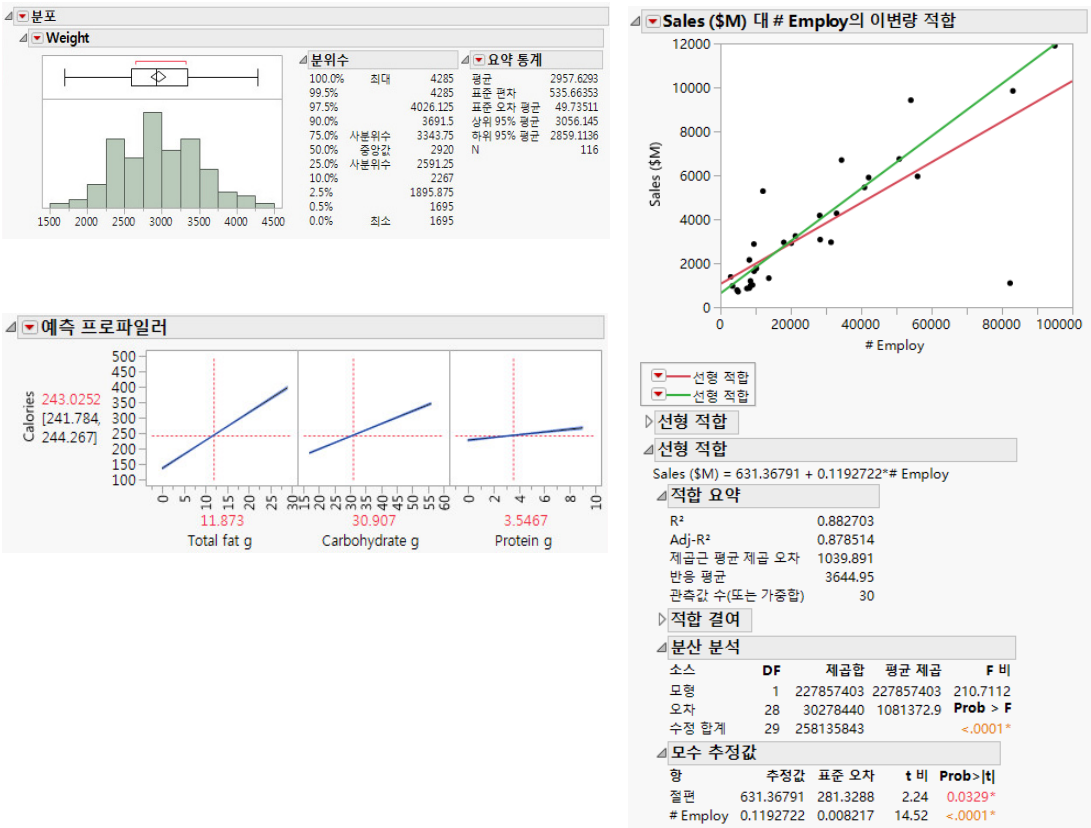
데이터 분석

분포 , 관계 및 모형

JMP에서 데이터를 분석하면 정보를 기반으로 의사 결정을 내릴 수 있습니다. 데이터 분석에는 대개 다음과 같은 작업이 포함됩니다.

- 분포 조사
- 관계 발견
- 가설 검정
- 모형 구축

그림 5.1 분석 예



목차

내용 소개 129
데이터 그래프 생성의 중요성 129
모델링 유형 이해 132
예: 모델링 유형 결과 보기 133
모델링 유형 변경 134
분포 분석 136
연속형 변수의 분포 137
범주형 변수의 분포 139
관계 분석 142
하나의 예측 변수가 있는 회귀 사용 143
한 변수에 대한 평균 비교 148
비율 비교 151
여러 변수의 평균 비교 153
다중 예측 변수가 있는 회귀 사용 157

내용 소개

데이터를 분석하기 전에 기본 개념에 대한 다음 정보를 검토하십시오 .

- " 데이터 그래프 생성의 중요성 "(129 페이지)
- " 모델링 유형 이해 "(132 페이지)

이 장의 나머지 부분에서는 JMP 에서 몇 가지 기본 분석 방법을 사용하는 방법을 소개합니다 .

- " 분포 분석 "(136 페이지)
- " 관계 분석 "(142 페이지)

고급 예측 모형 및 분석 방법에 대한 설명은 다음 JMP 설명서에서 확인하십시오 .

- Fitting Linear Models
- Multivariate Methods
- Predictive and Specialized Modeling
- Consumer Research
- Reliability and Survival Methods
- Quality and Process Methods

데이터 그래프 생성의 중요성

데이터를 그래프로 생성하거나 시각화하는 것은 모든 데이터 분석에 중요하며 항상 통계 검정 또는 모형 구축 전에 수행되어야 합니다 . 데이터 시각화가 데이터 분석 프로세스의 초기 단계에 이루어져야 하는 이유를 분명하게 파악하려면 다음 예를 참조하십시오 .

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Anscombe.jmp(F. J. Anscombe(1973), American Statistician, 27, 17-21) 를 엽니다 .

   이 데이터는 4 쌍의 *X* 및 *Y* 변수로 구성되어 있습니다 .

2. 테이블 패널에서 **The Quartet** 스크립트 옆의 녹색 삼각형을 클릭합니다 .

   이 스크립트는 **X 로 Y 적합**을 사용하여 각 변수 쌍에 대한 단순 선형 회귀를 생성합니다 . **점 표시** 옵션이 해제되어 있으므로 산점도에서 어떤 데이터도 볼 수 없습니다 . 그림 5.2 에서는 각 회귀에 대한 모형 적합 및 기타 요약 정보를 보여 줍니다 .

그림 5.2 네 가지 모형

쌍 1 에 대한 모형

선형 적합

Y1 = 3.0000909 + 0.5000909*X1

적합 요약

R²	0.666542
Adj-R²	0.629492
제곱근 평균 제곱 오차	1.236603
반응 평균	7.500909
관측값 수(또는 가중합)	11

분산 분석

소스	DF	제곱합	평균 제곱	F 비
모형	1	27.510001	27.5100	17.9899
오차	9	13.762690	1.5292	Prob > F
수정 합계	10	41.272691		0.0022*

모수 추정값

항	추정값	표준 오차	t 비	Prob>|t|
절편	3.0000909	1.124747	2.67	0.0257*
X1	0.5000909	0.117906	4.24	0.0022*

쌍 3 에 대한 모형

선형 적합

Y3 = 3.0024545 + 0.4997273*X3

적합 요약

R²	0.666324
Adj-R²	0.629249
제곱근 평균 제곱 오차	1.236311
반응 평균	7.5
관측값 수(또는 가중합)	11

분산 분석

소스	DF	제곱합	평균 제곱	F 비
모형	1	27.470008	27.4700	17.9723
오차	9	13.756192	1.5285	Prob > F
수정 합계	10	41.226200		0.0022*

모수 추정값

항	추정값	표준 오차	t 비	Prob>|t|
절편	3.0024545	1.124481	2.67	0.0256*
X3	0.4997273	0.117878	4.24	0.0022*

쌍 2 에 대한 모형

선형 적합

Y2 = 3.0009091 + 0.5*X2

적합 요약

R²	0.666242
Adj-R²	0.629158
제곱근 평균 제곱 오차	1.237214
반응 평균	7.500909
관측값 수(또는 가중합)	11

분산 분석

소스	DF	제곱합	평균 제곱	F 비
모형	1	27.500000	27.5000	17.9656
오차	9	13.776291	1.5307	Prob > F
수정 합계	10	41.276291		0.0022*

모수 추정값

항	추정값	표준 오차	t 비	Prob>|t|
절편	3.0009091	1.125302	2.67	0.0258*
X2	0.5	0.117964	4.24	0.0022*

쌍 4 에 대한 모형

선형 적합

Y4 = 3.0017273 + 0.4999091*X4

적합 요약

R²	0.666707
Adj-R²	0.629675
제곱근 평균 제곱 오차	1.235695
반응 평균	7.500909
관측값 수(또는 가중합)	11

분산 분석

소스	DF	제곱합	평균 제곱	F 비
모형	1	27.490001	27.4900	18.0033
오차	9	13.742490	1.5269	Prob > F
수정 합계	10	41.232491		0.0022*

모수 추정값

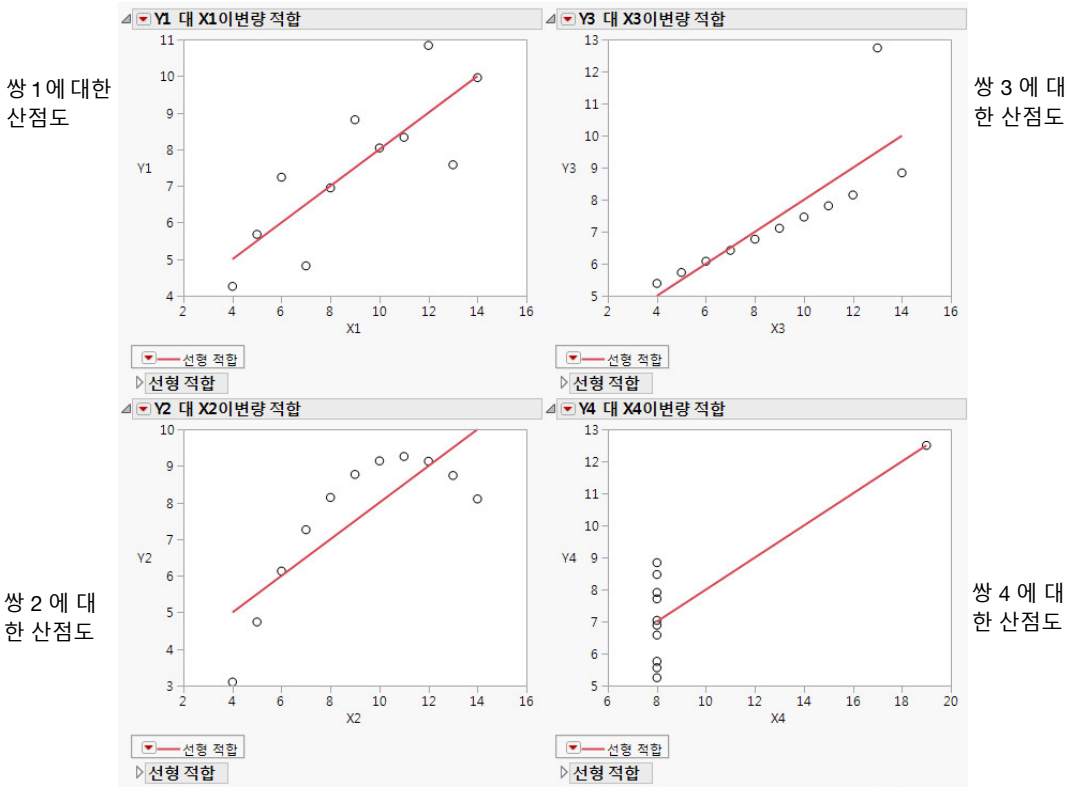
항	추정값	표준 오차	t 비	Prob>|t|
절편	3.0017273	1.123921	2.67	0.0256*
X4	0.4999091	0.117819	4.24	0.0022*

네 가지 모형과 R^2 값이 거의 동일하다는 것에 유의하십시오. 각 사례에서 적합 모형은 기본적으로 $Y = 3 + 0.5X$ 이며 R^2 값은 기본적으로 0.66 입니다. 데이터 분석에서 위의 요약 정보만 고려하면 각 사례의 X 와 Y 간 관계가 동일하다고 결론 내릴 수 있습니다. 그러나 이 시점에는 데이터를 시각화하지 않았습니다. 결론이 잘못되었을 수도 있습니다.

데이터 시각화를 위해 네 가지 산점도 모두에 점 추가

1. Ctrl 키를 누릅니다 .
2. " 이변량 적합 " 중 하나의 옆에 있는 빨간색 삼각형을 클릭하고 **점 표시**를 선택합니다 .

그림 5.3 점이 추가된 산점도



산점도는 관계를 나타내는 선이 같더라도 X 와 Y 간의 관계가 네 쌍에서 동일하지 않음을 보여 줍니다 .

- 산점도 1 은 선형 관계를 나타냅니다 .
- 산점도 2 는 비선형 관계를 나타냅니다 .
- 산점도 3 은 하나의 이상치를 제외하고 선형 관계를 나타냅니다 .
- 산점도 4 에서는 한 점을 제외하고 x = 8 에 모든 데이터가 있습니다 .

이 예는 통계만을 기반으로 한 결론이 부적절할 수 있음을 보여 줍니다. 모든 데이터 분석의 초기 단계에는 데이터를 시각적으로 탐색하는 과정이 있어야 합니다.

모델링 유형 이해

JMP 에서는 데이터가 여러 가지 유형일 수 있습니다 . JMP 에서는 이것을 데이터의 모델링 유형이라고 합니다 . 표 5.1 에서는 JMP 의 세 가지 모델링 유형을 보여 줍니다 .

표 5.1 모델링 유형

모델링 유형 및 설명	예	구체적인 예
연속형 숫자 데이터에만 해당됩니다. 합계 및 평균 같은 연산에서 사용됩니다.	높이 온도 시간	시험을 완료하는 데 소요되는 시간은 2 시간 또는 2.13 시간일 수 있습니다.
순서형 숫자 또는 문자 데이터에 해당됩니다. 값은 정렬된 범주에 속합니다.	월 (1,2,...,12) 문자 등급 (A, B,...F) 크기 (소, 중, 대)	월은 2(2 월) 또는 3(3 월) 일 수 있지만 2.13 일 수 없습니다. 2 월은 3 월 전에 옵니다.
명목형 숫자 또는 문자 데이터에 해당됩니다. 값은 범주에 속하지만 순서는 중요하지 않습니다.	성별 (남 또는 여) 색상 시험 결과 (통과 또는 실패)	성별은 남 또는 여가 될 수 있으며 순서는 없습니다. 성별 범주를 숫자(남 = 1 및 여 = 2) 로 나타낼 수도 있습니다.
다중 반응 문자 데이터에만 해당됩니다. 단일 셀에서 쉼표로 구분된 고유 항목입니다.	이를 닦는 시간 대학 학위 즐겨 하는 스포츠	하루에 여러 번 이를 닦을 수 있습니다. 아침에 맨 먼저, 아침 식사 후, 식사 후, 잠자기 전 또는 이들의 조합입니다.
비정형 텍스트 문자 데이터에만 해당됩니다. 일반적으로 텍스트 탐색기를 사용하여 분석해야 하는 모든 고유 값입니다.	제품 후기 노래 가사 설문 조사의 자유 의견 작성 필드	제품 후기는 대부분 고유하며 텍스트 탐색기를 사용하여 기본적인 유사성을 결정합니다.
벡터 표현식 데이터에만 해당됩니다. 셀의 값은 열 벡터 또는 행 벡터입니다.	예측 계산식	

표 5.1 모델링 유형 (계속)

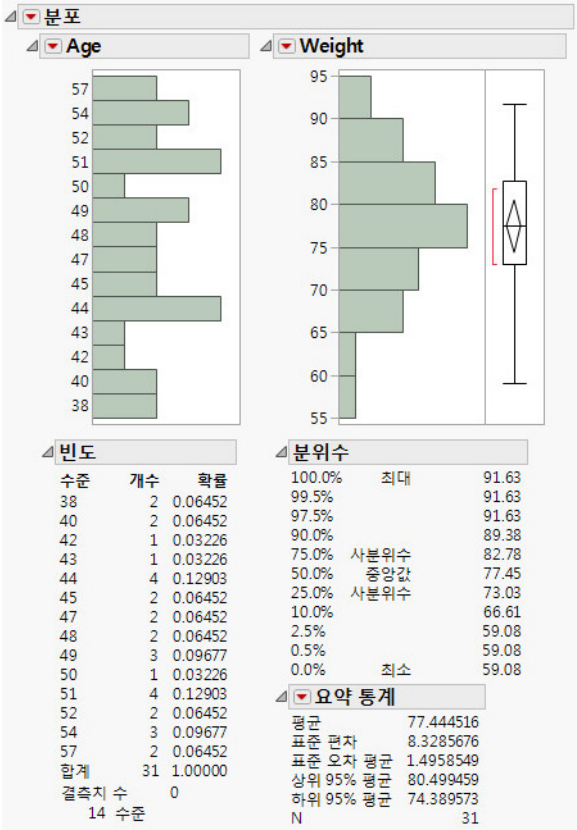
모델링 유형 및 설명	예	구체적인 예
없음 모든 데이터 유형에 해당됩니다. 열이 다른 모델링 유형으로 잘 표현되지 않는 시나리오에서 사용됩니다.	그림 ID 값	데이터 테이블의 그림 열은 모델링에 사용되지 않지만 그래프에서 표식 등으로 사용할 수 있습니다.

예 : 모델링 유형 결과 보기

모델링 유형이 다르면 JMP 에서 생성되는 결과도 달라집니다 . 차이점을 보려면 다음 단계를 수행하십시오 .

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Linnerud.jmp 를 엽니다 .
2. **분석** > **분포**를 선택합니다 .
3. Age 및 Weight 를 선택하고 **Y, 열**을 클릭합니다 .
4. **확인**을 클릭합니다 .

그림 5.4 Age 및 Weight 에 대한 분포 결과



Age 및 Weight 는 모두 숫자 변수이지만 동일하게 처리되지 않습니다. 표 5.2 에서는 Weight 및 Age 에 대한 결과 간의 차이점을 비교합니다.

표 5.2 Weight 및 Age 에 대한 결과

변수	모델링 유형	결과
Weight	연속형	히스토그램, 사분위수 및 요약 통계량
Age	순서형	막대 차트 및 빈도

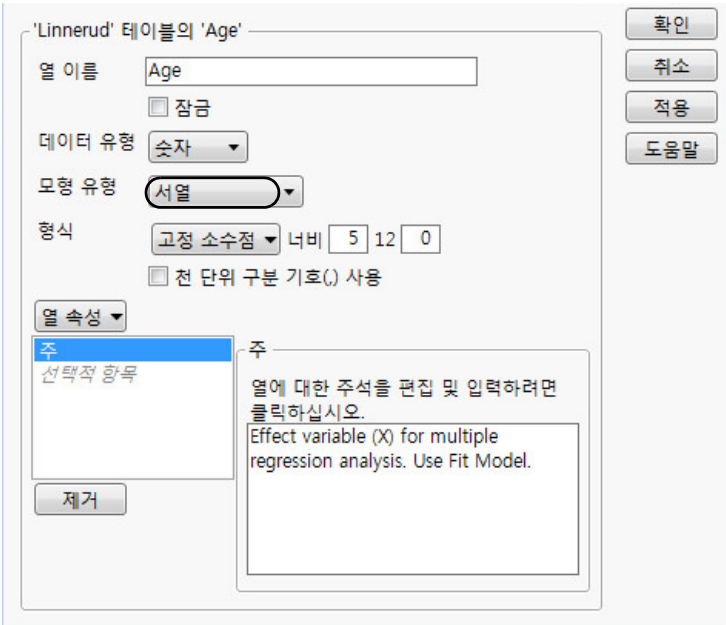
모델링 유형 변경

변수를 다르게 처리하려면 모델링 유형을 변경하십시오 . 예를 들어 그림 5.4 에서 Age 의 모델링 유형은 순서형입니다 . 순서형 변수에서 JMP 는 빈도 수를 계산합니다 . 빈도 수 대신 평균 연령

을 확인하려고 한다고 가정해 보겠습니다 . 모델링 유형을 평균 연령을 보여 주는 연속형으로 변경합니다 .

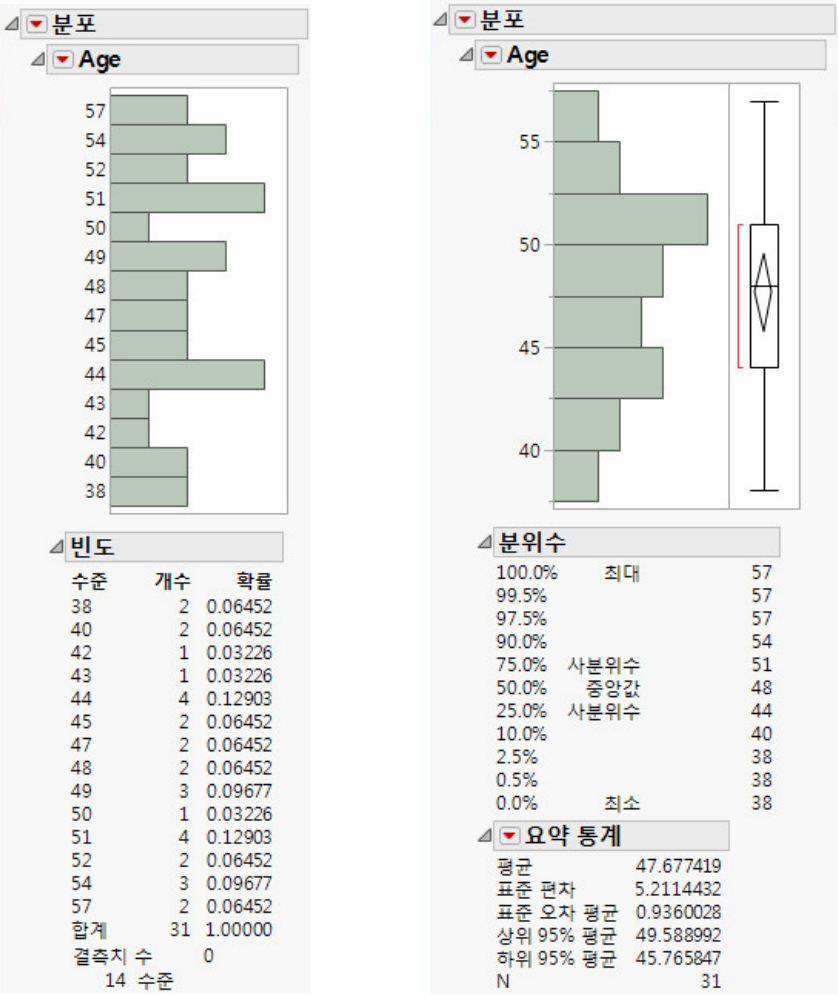
1. Age 열 머리글을 두 번 클릭합니다 . " 열 정보 " 창이 나타납니다 .
2. " 모델링 유형 " 을 **연속형**으로 변경합니다 .

그림 5.5 열 정보 창



3. **확인**을 클릭합니다 .
4. 이 예의 단계를 반복하여 (" 예 : 모델링 유형 결과 보기 "(133 페이지) 참조) 분포를 생성합니다 . 그림 5.6 에서는 Age 가 순서형일 때와 연속형일 때의 분포 결과를 보여 줍니다 .

그림 5.6 Age 의 서로 다른 모델링 유형



순서형 연속형

Age가 순서형이면 각 연령의 빈도 수를 볼 수 있습니다. 예를 들어 연령 48은 두 번 나타납니다. 연령이 연속형이면 평균 연령(이 예의 경우 약 48세, 즉 47.677세)을 확인할 수 있습니다.

분포 분석

단일 변수를 분석하려는 경우 분포 플랫폼을 사용하여 변수의 분포를 조사할 수 있습니다. 각 변수의 보고서 내용은 변수가 범주형(명목형 또는 순서형)인지 연속형인지에 따라 달라집니다.

참고 : 분포 플랫폼에 대한 자세한 내용은 **기본 분석**의 에서 확인하십시오 .

연속형 변수의 분포

연속형 변수를 분석할 때는 다음과 같은 질문을 던질 수 있습니다 .

- 데이터의 모양이 알려진 분포와 매칭됩니까 ?
- 데이터에 이상치가 있습니까 ?
- 데이터의 평균은 얼마입니까 ?
- 평균이 목표값 또는 과거 값과 통계적으로 다릅니까 ?
- 데이터가 얼마나 퍼져 있습니까 ? 즉 , 표준편차는 얼마입니까 ?
- 최소값과 최대값은 얼마입니까 ?

그래프, 요약 통계량 및 간단한 통계 검정을 통해 이러한 질문과 그 밖의 질문에 대한 답을 얻을 수 있습니다.

시나리오

이 예에서는 116가지 자동차 모델에 대한 정보가 들어 있는 Car Physical Data.jmp 데이터 테이블을 사용합니다.

계획 전문가가 철도 회사로부터 기차로 자동차를 운송할 때 발생할 수 있는 문제를 파악해 달라는 요청을 받았습니다 . 계획 전문가는 데이터를 사용하여 다음과 같은 질문에 대한 답을 구하려고 합니다 .

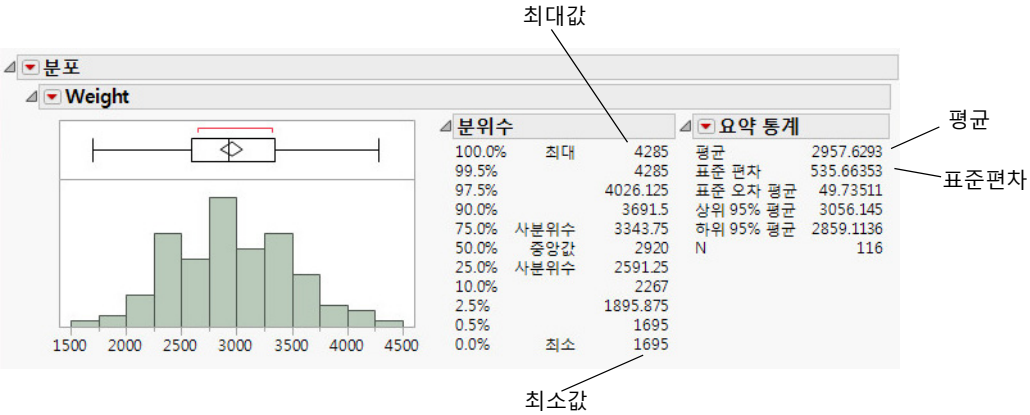
- 평균 자동차 중량은 얼마입니까 ?
- 자동차의 중량은 얼마나 퍼져 있습니까 (표준편차)?
- 자동차의 최소 중량과 최대 중량은 얼마입니까 ?
- 데이터에 이상치가 있습니까 ?

이러한 질문에 답하기 위해 중량에 대한 히스토그램을 사용합니다.

히스토그램 생성

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Car Physical Data.jmp 를 엽니다 .
2. **분석** > **분포**를 선택합니다 .
3. Weight 를 선택하고 **Y, 열**을 클릭합니다 .
4. **확인**을 클릭합니다 .
5. 보고서 창을 회전하려면 "Weight" 의 빨간색 삼각형을 클릭하고 **표시 옵션** > **가로 레이아웃**을 선택합니다 .

그림 5.7 Weight 분포



보고서 창에는 다음 세 가지 섹션이 있습니다 .

- 히스토그램과 상자 그림은 데이터를 시각화합니다 .
- 분위수 보고서는 분포의 백분위수를 보여 줍니다 .
- 요약 통계량 보고서는 평균 , 표준편차 및 기타 통계량을 보여 줍니다 .

분포 결과 해석

계획 전문가는 그림 5.7 에 제시된 결과를 기반으로 질문에 답할 수 있습니다 .

평균 자동차 중량은 얼마입니까 ? 히스토그램은 중량이 약 3,000 파운드임을 보여 줍니다 .

자동차의 중량은 얼마나 퍼져 있습니까 (표준편차)? 요약 통계량에서 중량은 약 2,958 파운드이고 표준편차는 약 536 파운드입니다 .

최소 중량과 최대 중량은 얼마입니까 ? 히스토그램에서 최소 중량은 약 1,500 파운드이고 최대 중량은 약 4,500 파운드입니다 . 분위수에서 최소 중량은 약 1,695 파운드이고 최대 중량은 약 4,285 파운드입니다 .

이상치가 있습니까 ? 없습니다 .

그림 5.7의 기본 보고서 창에서는 최소한의 그래프와 통계량이 제공됩니다. 추가적인 그래프 및 통계량은 빨간색 삼각형 메뉴를 클릭하여 확인할 수 있습니다.

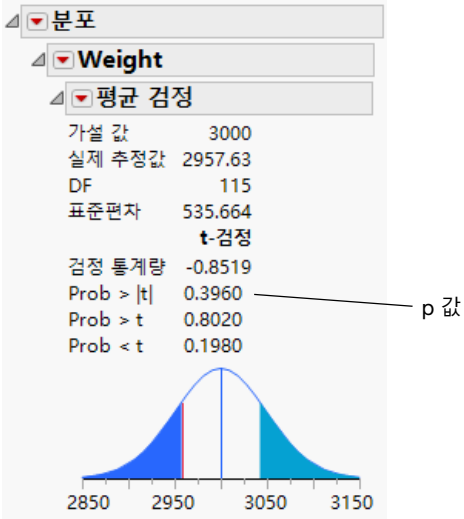
결론

다른 조사 결과에 따라, 철도 회사는 평균 3,000 파운드를 운송하는 것이 가장 효율적이라고 판단했습니다. 이제 계획 전문가는 운송할 수 있는 일반적인 자동차 모집단의 평균 자동차 중량이 3,000 파운드인지 알아내야 합니다. *t*-검정을 통해 이 모집단 표본을 바탕으로 보다 광범위한 모집단에 대한 추론을 이끌어 내야 합니다.

결론 검정

1. "Weight" 의 빨간색 삼각형을 클릭하고 **평균 검정**을 선택합니다 .
2. 창이 나타나면 " 가설 평균 지정 " 상자에 3000 을 입력합니다 .
3. **확인**을 클릭합니다 .

그림 5.8 평균 검정 결과



t- 검정 해석

t-검정의 기본 결과는 *p* 값입니다. 이 예에서 *p* 값은 0.396 이고 분석가는 0.05 의 유의 수준을 사용합니다. 0.396 은 0.05 보다 크기 때문에 더 광범위한 모집단에서 자동차 모델의 평균 중량이 3,000 파운드와 유의하게 다르다는 결론을 내릴 수는 없습니다. *p* 값이 유의 수준보다 낮았다면 계획 전문가는 더 광범위한 모집단의 평균 자동차 중량이 3,000 파운드와 유의하게 다르다는 결론을 내렸을 것입니다.

범주형 변수의 분포

범주형 (순서형 또는 명목형) 변수를 분석할 때는 다음과 같은 질문을 던질 수 있습니다 .

- 변수에는 몇 개의 수준이 있습니까 ?
- 각 수준에는 몇 개의 데이터 점이 있습니까 ?
- 데이터가 균등하게 분포되어 있습니까 ?
- 각 수준이 차지하는 비율은 어느 정도입니까 ?

시나리오

자세한 내용은 "연속형 변수의 분포"(137 페이지)의 시나리오에서 확인하십시오.

철도 회사에서는 자동차의 평균 중량이 목표 중량과 유의하게 다르지 않다고 판단했으므로 해결해야 할 질문이 더 있습니다.

계획 전문가는 철도 회사의 다음 질문에 대답하고자 합니다.

- 자동차의 종류는 무엇입니까?
- 원산지 국가는 어디입니까?

이러한 질문에 답하려면 Type 및 Country 분포를 보십시오.

분포 생성

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Car Physical Data.jmp 를 엽니다.
2. **분석** > **분포**를 선택합니다.
3. Country 및 Type 을 선택하고 **Y, 열**을 클릭합니다.
4. **확인**을 클릭합니다.

그림 5.9 Country 및 Type 분포

분포 결과 해석

보고서 창에는 Country 및 Type 에 대한 막대 차트와 " 빈도 " 보고서가 있습니다 . 막대 차트는 빈도 보고서에서 제공된 빈도 정보를 그래픽으로 나타낸 것입니다 . " 빈도 " 보고서에는 다음 항목이 포함됩니다 .

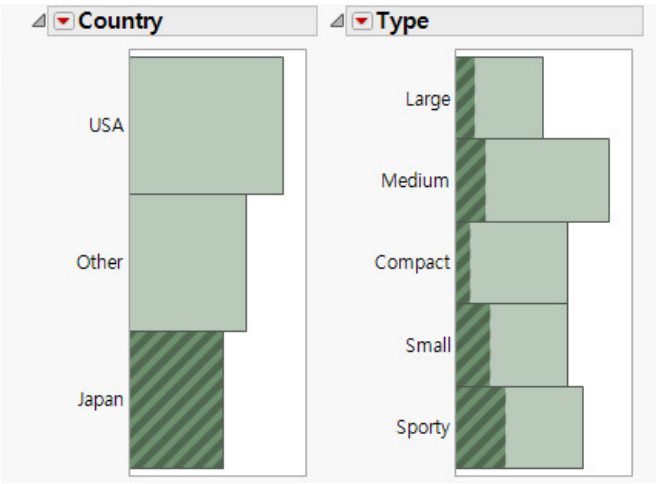
- 데이터의 범주. 예를 들어 "Japan" 은 "Country" 의 범주이고 "Sporty" 는 "Type" 의 범주입니다.
- 각 범주의 총 수
- 각 범주가 차지하는 전체 대비 비율

예를 들어 경차 ("Compact") 는 22 대가 있으며 116 개의 관측값 중 약 19% 를 차지합니다.

분포 결과에서의 상호 작용

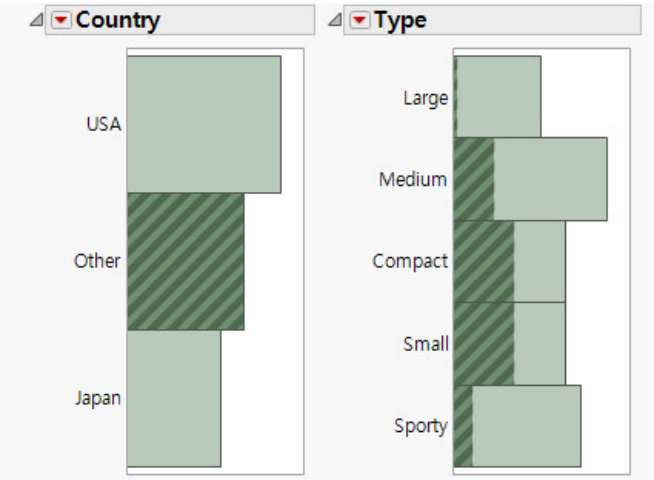
한 차트에서 막대를 선택하면 다른 차트에서 해당하는 데이터가 선택됩니다. 예를 들어 "Country" 막대 차트에서 "Japan" 막대를 선택하면 많은 수의 일본 자동차가 스포츠카인 것을 볼 수 있습니다.

그림 5.10 일본 자동차



"Other" 범주를 선택하면 이러한 자동차의 대다수가 소형 ("Small") 또는 경차 ("Compact") 이고 대형 ("Large") 은 거의 없음을 알 수 있습니다.

그림 5.11 기타 국가의 자동차



관계 분석

산점도 및 기타 그래프는 변수 간의 관계를 시각화하는 데 도움이 됩니다. 관계를 시각화한 후에는 다음 단계로 그 관계를 분석하여 수치로 설명할 수 있습니다. 변수 간의 관계를 수치로 정의한 것을 **모형**이라고 합니다. 더욱 중요한 것은 모형에서 한 변수(*Y*)의 평균 값을 다른 변수(*X*)의 값을 바탕으로 예측할 수도 있다는 것입니다. *X* 변수를 예측 변수라고도 합니다. 일반적으로 이러한 모형을 **회귀** 모형이라고 합니다.

JMP 의 **X 로 Y 적합** 플랫폼과 **모형 적합** 플랫폼에서는 회귀 모형을 생성합니다.

참고 : 여기서는 기본 플랫폼과 옵션만 다룹니다 . 모든 플랫폼 옵션에 대한 설명은 **기본 분석** , Essential Graphing, 그리고 " 내용 소개 "(129 페이지) 에 나열된 설명서에서 확인하십시오 .

표 5.3에서는 네 가지 기본 유형의 관계를 보여 줍니다.

표 5.3 관계 유형

X	Y	섹션
연속형	연속형	• " 하나의 예측 변수가 있는 회귀 사용 "(143 페이지) • "다중 예측 변수가 있는 회귀 사용"(157페이지)
범주형	연속형	• " 한 변수에 대한 평균 비교 "(148 페이지) • " 여러 변수의 평균 비교 "(153 페이지)

표 5.3 관계 유형 (계속)

X	Y	섹션
범주형	범주형	"비율 비교"(151 페이지)
연속형	범주형	로지스틱 회귀에 대해서는 자세한 설명이 필요합니다. 자세한 내용은 **기본 분석**의 에서 확인하십시오.

하나의 예측 변수가 있는 회귀 사용

시나리오

이 예에서는 제약 및 컴퓨터 업계의 32 개 회사에 대한 재무 데이터가 포함된 Companies.jmp 데이터 테이블을 사용합니다.

직원 수가 많은 회사가 직원 수가 적은 회사보다 많은 매출 수익을 창출한다는 것은 직관적으로 알 수 있습니다. 데이터 분석가는 직원 수에 따라 각 회사의 전체 매출 수익을 예측하려고 합니다.

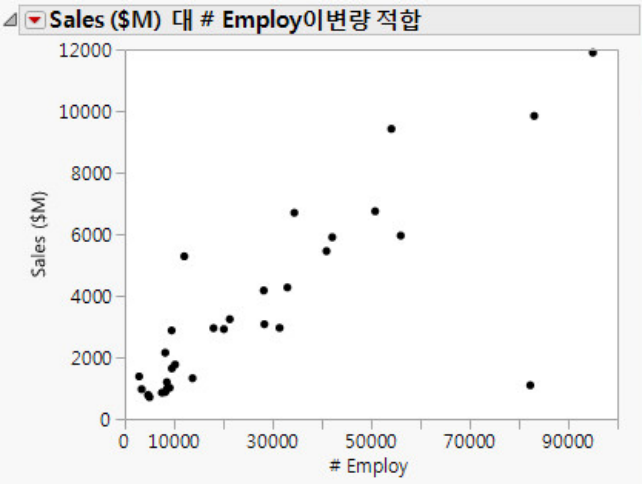
이 작업을 완수하려면 다음을 수행하십시오 .

- " 관계 발견 "(143 페이지)
- " 회귀 모형 적합 "(144 페이지)
- " 평균 매출 예측 "(146 페이지)

관계 발견

먼저, 직원 수와 매출 수익 간의 관계를 확인하기 위해 산점도를 생성합니다. 이 산점도는 " 데이터 시각화" 장의 "산점도 생성"(100 페이지)에서 생성되었습니다. 그림 5.12 의 산점도는 여기에서 이상치 하나(직원 수와 매출이 현저하게 높은 회사)를 숨기고 제외한 후의 결과를 보여 줍니다.

그림 5.12 Sales ($M) 대 # Employ 산점도

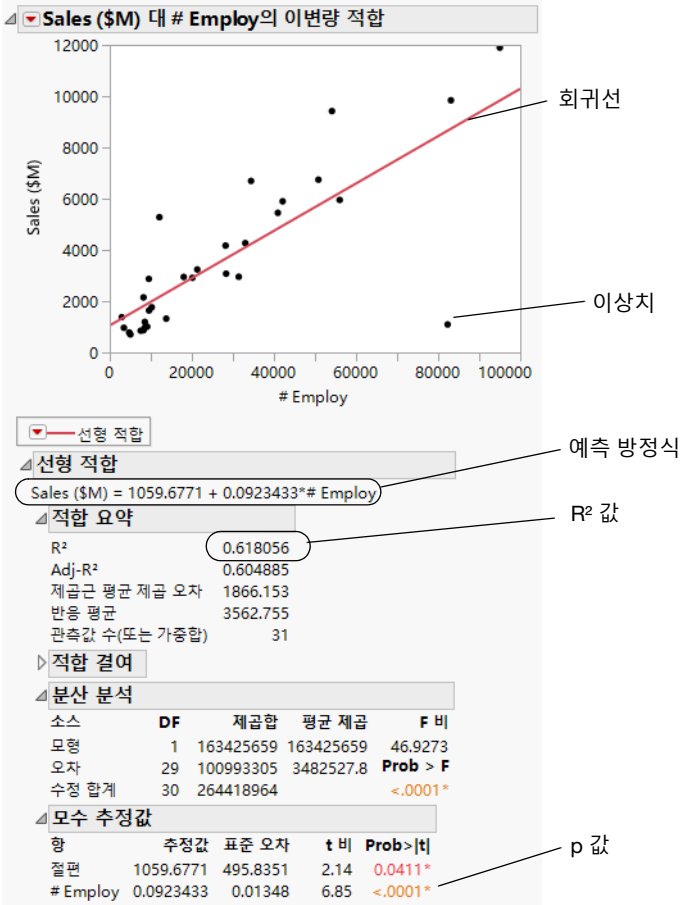


이 산점도는 매출과 직원 수 간의 관계를 명확하게 보여 줍니다. 예상대로, 회사의 직원이 많을수록 매출액이 높아질 수 있습니다. 이 산점도로 데이터 분석가의 추측을 시각적으로 확인할 수 있지만 특정 수의 직원에 대한 매출을 예측할 수는 없습니다.

회귀 모형 적합

직원 수로 매출 수익을 예측하려면 회귀 모형을 적합시켜야 합니다. "이변량 적합"의 빨간색 삼각형을 클릭하고 **선형 적합**을 선택합니다. 산점도에 회귀선이 추가되고 보고서 창에 보고서가 추가됩니다.

그림 5.13 회귀선



보고서에서 다음 결과를 확인하십시오 .

- *p* 값이 .0001 보다 작음
- R^2 값이 0.618 임

이러한 결과를 바탕으로 데이터 분석가는 다음과 같은 결론을 내릴 수 있습니다 .

- "# Employ" 모형 항의 *p* 값이 작습니다 . 이는 유의 수준 0.05 에서 "# Employ" 에 대한 계수가 0 이 아니라는 것을 뒷받침합니다 . 따라서 예측 모형에 직원 수를 포함하면 직원 수가 포함되지 않은 모형에 비해 평균 매출을 예측하는 능력이 크게 향상됩니다 .
- R^2 값 0.618 은 이 모형이 매출 변동의 약 62% 를 설명하고 있음을 나타냅니다 . R^2 값은 결정계수로서 , 모형으로 설명되는 종속 (반응) 변수의 분산 비율을 나타냅니다 . R^2 는 0 에서 1 사이입니다 . R^2 가 0 인 모형은 설명력이 없습니다 . R^2 가 1 인 모형은 반응을 완벽하게 예측합니다 .

평균 매출 예측

회귀 모형을 사용하면 특정 수의 직원이 있을 때 회사에서 기대할 수 있는 평균 매출을 예측할 수 있습니다 . 이 모형에 대한 예측 방정식이 보고서에 포함되어 있습니다 .

평균 매출 = 1059.68 + 0.092* 직원 수

예를 들어 직원 수가 70,000 명인 회사의 매출은 약 7,500 달러로 예측됩니다 .

7,499.68 달러 = 1059.68 + 0.092*70,000

현재 산점도의 오른쪽 아래에는 다른 회사의 일반적인 패턴을 따르지 않는 이상치가 있습니다. 데이터 분석가는 이 이상치가 제외될 때 예측 모형이 바뀌는지에 대해 알고 싶습니다.

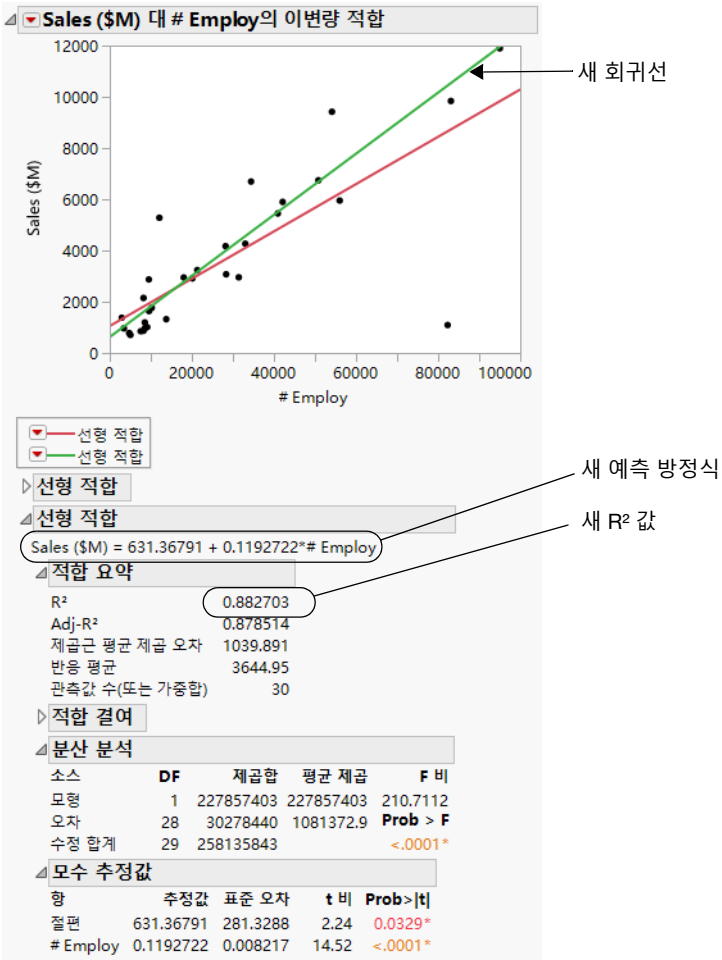
이상치 제외

1. 이상치를 클릭합니다 .
2. **행** > **제외 / 제외 해제**를 선택합니다 .
3. 이 모형을 적합시키려면 "Sales (SM) 대 # Employ 의 이변량 적합 " 옆의 빨간색 삼각형을 클릭하고 **선형 적합**을 선택합니다 .

다음 항목이 보고서 창에 추가됩니다(그림 5.14).

- 새 회귀선
- 다음을 포함하는 새 선형 적합 보고서
  - 새 예측 방정식
  - 새 R^2 값

그림 5.14 모형 비교



결과 해석

그림 5.14 의 결과를 바탕으로 데이터 분석가는 다음과 같은 결론을 내릴 수 있습니다 .

- 이상치는 큰 회사의 회귀선을 끌어내리고 작은 회사의 회귀선을 끌어올립니다 .
- 이상치가 없는 데이터에 대한 새 모형은 첫 번째 모형보다 강력한 모형입니다 . 새 R^2 값 0.88 은 초기 분석보다 높고 1 에 더 가깝습니다 .

결론

새 예측 방정식을 사용하면 7 만 명의 직원이 있는 회사의 예측 평균 매출을 다음과 같이 계산할 수 있습니다 .

8,961.37 달러 = 631.37 + 0.119*70,000

첫 번째 모형의 예측값은 약 7,500 달러입니다. 두 번째 모형에서는 총 매출을 첫 번째 모형에 비해 1,460 달러 늘어난 약 8,960 달러로 예측합니다.

이상치를 제거한 두 번째 모형은 직원 수를 기준으로 한 매출 총액을 첫 번째 모형보다 더욱 정확하게 설명하고 예측합니다. 데이터 분석가는 이제 사용하기 적절한 모형을 확보했습니다.

한 변수에 대한 평균 비교

연속형 Y 변수와 범주형 X 변수가 있으면 X 변수의 수준에서 평균을 비교할 수 있습니다.

시나리오

이 예에서는 제약 및 컴퓨터 업계의 32 개 회사에 대한 재무 데이터가 포함된 Companies.jmp 데이터 테이블을 사용합니다.

재무 분석가가 다음과 같은 질문에 대한 답을 구하려고 합니다 .

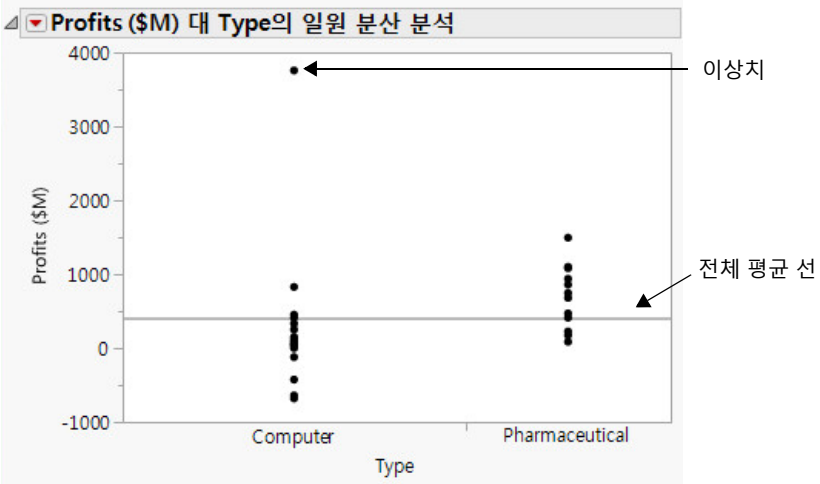
- 컴퓨터 회사의 수익은 제약 회사의 수익과 어떻게 비교됩니까 ?

이 질문에 답하려면 Type 을 기준으로 Profits ($M) 를 적합시킵니다.

관계 발견

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다 .
2. Companies.jmp 샘플 데이터 테이블을 열어 둔 상태라면 제외되었거나 숨겨진 행이 있을 수 있습니다 . 행을 기본 상태로 되돌려 숨겨진 행 없이 모든 행을 포함하려면 **행** > **행 상태 지우기**를 선택합니다 .
3. **분석** > **X 로 Y 적합**을 선택합니다 .
4. Profits ($M) 를 선택하고 **Y, 반응**을 클릭합니다 .
5. Type 을 선택하고 **X, 요인**을 클릭합니다 .
6. **확인**을 클릭합니다 .

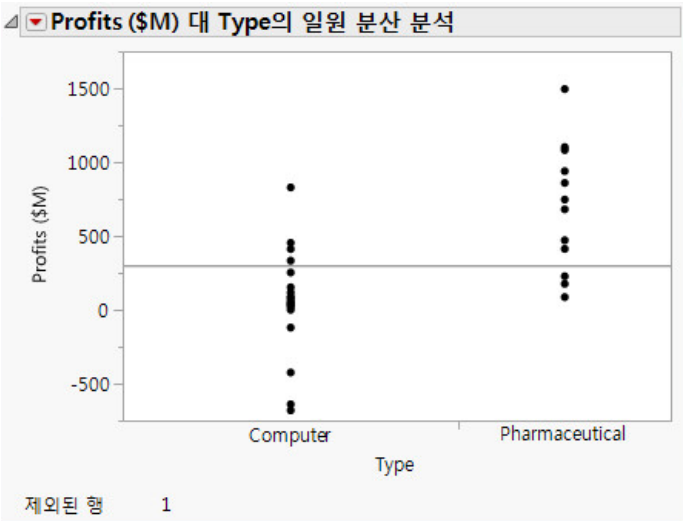
그림 5.15 회사 유형별 수익



컴퓨터 유형에 이상치가 있습니다 . 이 이상치는 산점도의 범위를 늘리고 수익을 비교하기 어렵게 만듭니다 . 이상치를 제외하고 숨깁니다 .

1. 이상치를 클릭합니다 .
2. **행** > **제외 / 제외 해제**를 선택합니다 . 해당 데이터 점이 더 이상 계산에 포함되지 않습니다 .
3. **행** > **숨기기 / 숨기기 해제**를 선택합니다 . 해당 데이터 점이 모든 그래프에서 숨겨집니다 .
4. 이상치를 제외하고 그림을 다시 생성하려면 "Profits ($M) 대 Type 의 일원 분석 " 을 클릭하고 **다시 실행** > **분석 다시 실행**을 선택합니다 . 원래의 산점도 창은 닫아도 됩니다 .

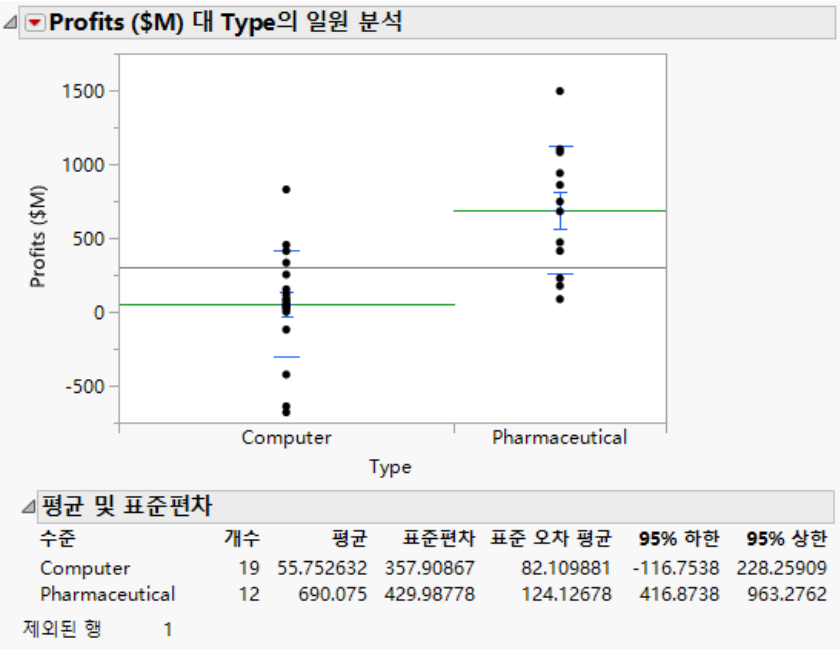
그림 5.16 업데이트된 그림



이상치를 제거하면 재무 분석가가 데이터를 보다 명확하게 파악할 수 있습니다 .

5. 관계 분석을 계속하려면 "Profits ($M) 대 Type 의 일원 분석 " 옆의 빨간색 삼각형에서 다음 옵션을 선택합니다 .
   - **표시 옵션** > **평균 선** . 이 옵션을 선택하면 산점도에 평균 선이 추가됩니다 .
   - **평균 및 표준편차** . 이 옵션을 선택하면 평균 및 표준편차를 제공하는 보고서가 표시됩니다 .

그림 5.17 평균 선 및 보고서



수준	개수	평균	표준편차	표준 오차 평균	95% 하한	95% 상한
Computer	19	55.752632	357.90867	82.109881	-116.7538	228.25909
Pharmaceutical	12	690.075	429.98778	124.12678	416.8738	963.2762

제외된 행 1

결과 해석

재무 분석가는 컴퓨터 회사의 수익이 제약 회사의 수익과 어떻게 비교되는지 알고자 했습니다. 업데이트된 산점도는 제약 회사의 평균 수익이 컴퓨터 회사보다 높다는 것을 보여 줍니다. 이 보고서의 경우 한 쪽 평균 값에서 다른 쪽의 평균 값을 뺀 수익 차이는 약 6 억 3,500 만 달러입니다. 이 산점도는 또한 컴퓨터 회사 중 일부의 수익이 마이너스이고 모든 제약 회사의 수익은 플러스임을 보여 줍니다.

t- 검정 수행

재무 분석가는 회사 표본 (데이터 테이블에 있는 회사) 만을 살펴보았습니다 . 재무 분석가는 이제 다음 질문을 검토하고자 합니다 .

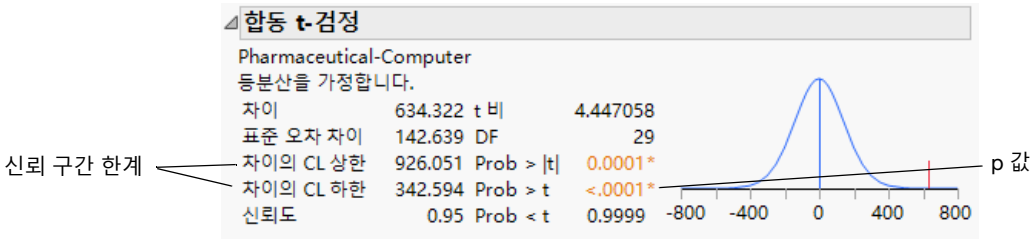
- 더 광범위한 모집단에서 차이가 있습니까 ? 아니면 6 억 3,500 만 달러의 차이는 우연입니까 ?

- 차이가 있다면 얼마입니까 ?

이 질문에 답하려면 2 표본 *t*-검정을 수행합니다. *t*-검정을 통해 표본의 데이터를 사용하여 더 많은 모집단에 대한 추론을 할 수 있습니다.

t-검정을 수행하려면 "일원 분석"의 빨간색 삼각형을 클릭하고 **평균/ANOVA/합동 t**를 선택합니다.

그림 5.18 t- 검정 결과



p 값 0.0001 은 유의 수준 0.05 보다 작으므로 통계적으로 유의함을 나타냅니다. 따라서 재무 분석가는 표본 데이터에 대해 관측된 평균 수익의 차이가 통계적으로 유의하다고 결론 내릴 수 있습니다. 즉, 더 큰 모집단에서 제약 회사의 평균 수익은 컴퓨터 회사의 평균 수익과 다릅니다.

결론

신뢰 구간 한계를 사용하여 두 회사 유형의 수익에 얼마나 많은 차이가 있는지 파악합니다. 그림 5.18 에서 **차이 CL 상한** 및 **차이 CL 하한** 값을 살펴봅니다. 재무 분석가는 제약 회사의 평균 수익이 컴퓨터 회사의 평균 수익보다 높은 3 억 4 천 3 백만 달러에서 9 억 2 천 6 백만 달러 사이라고 결론을 내립니다.

비율 비교

범주형 *X* 변수와 *Y* 변수를 사용할 때는 Y 변수 내 수준의 비율을 X 변수 내의 수준과 비교할 수 있습니다.

시나리오

이 예에서는 계속해서 Companies.jmp 데이터 테이블을 사용합니다. "한 변수에 대한 평균 비교"(148 페이지)에서 재무 분석가는 제약 회사가 컴퓨터 회사보다 평균적으로 더 높은 수익을 얻는 것으로 판단했습니다.

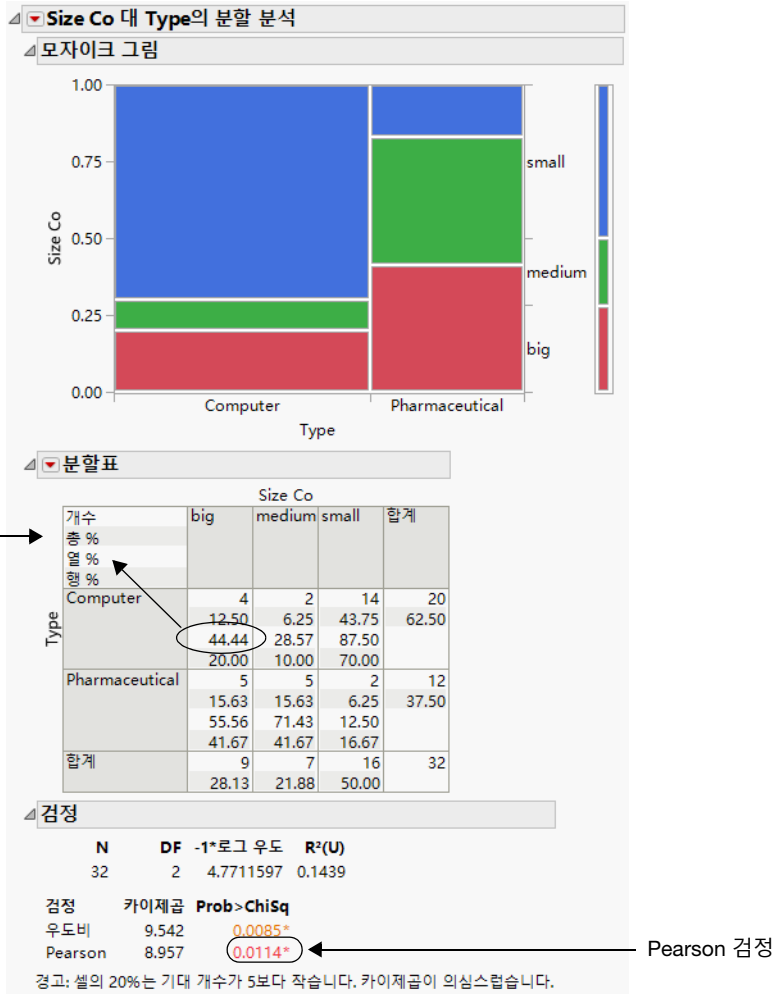
재무 분석가는 회사의 규모가 수익에 미치는 영향이 회사 유형별로 달라지는지 알아보려고 합니다. 그러나 이 질문을 검토하기 전에 재무 분석가는 컴퓨터 및 제약 회사의 모집단에서 소규모, 중간 규모 및 대규모 기업이 동일한 비율로 구성되었는지 알아야 합니다.

관계 발견

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다 .

2. 이전 예의 Companies.jmp 데이터 파일을 계속 열어 둔 상태라면 제외되거나 숨겨진 행이 있을 수 있습니다. 행을 기본 상태로 되돌려 숨겨진 행 없이 모든 행을 포함하려면 **행** > **행 상태 지우기**를 선택합니다.
3. **분석** > **X 로 Y 적합**을 선택합니다.
4. Size Co 를 선택하고 **Y, 반응**을 클릭합니다.
5. Type 을 선택하고 **X, 요인**을 클릭합니다.
6. **확인**을 클릭합니다.

그림 5.19 회사 유형별 회사 규모



분할표에는 이 예에 적용할 수 없는 정보가 포함되어 있습니다. "분할표"의 빨간색 삼각형을 클릭하고 **총 %** 및 **열 %**를 선택 취소하여 해당 정보를 제거합니다. 그림 5.20에서는 업데이트된 테이블을 보여 줍니다.

그림 5.20 업데이트된 분할표

분할표

Type	개수 행 %	big	medium	small	합계
		Size Co			
Computer		4 20.00	2 10.00	14 70.00	20
Pharmaceutical		5 41.67	5 41.67	2 16.67	12
합계		9	7	16	32

결과 해석

분할표의 통계는 모자이크 그림에 그래픽으로 표시됩니다. 모자이크 그림과 분할표는 두 업계의 소규모, 중간 규모 및 대규모 기업 비율을 비교합니다. 예를 들어 모자이크 그림은 컴퓨터 업계가 제약 업계에 비해 소규모 기업의 비율이 높음을 보여 줍니다. 분할표에서는 정확한 통계량을 보여 줍니다. 즉, 컴퓨터 회사의 70%가 소규모이고 제약 회사의 약 17%가 소규모입니다.

검정 해석

재무 분석가는 회사 표본(데이터 테이블에 있는 회사)만을 살펴보았습니다. 재무 분석가는 모든 컴퓨터 및 제약 회사의 더 광범위한 모집단에서 비율이 다른지 알아야 합니다.

이 질문에 답하려면 **검정** 보고서에서 Pearson 검정의 *p* 값을 사용해야 합니다(그림 5.19(152 페이지)). *p* 값 0.011 이 유의 수준 0.05 보다 작기 때문에 재무 분석가는 다음과 같이 결론을 내립니다.

- 표본 데이터의 차이는 통계적으로 유의합니다.
- 더 광범위한 모집단에서는 비율이 달라집니다.

이제 재무 분석가는 소규모, 중간 규모 및 대규모 기업의 비율이 다르다는 것을 알고 있으며 다음 질문에 답할 수 있습니다.

여러 변수의 평균 비교

"한 변수에 대한 평균 비교"(148 페이지) 섹션에서는 범주형 변수 하나의 여러 수준에서 평균을 비교했습니다. 한 번에 두 개 이상의 변수에 대해 여러 수준에서 평균을 비교하려면 **분산 분석** 기법(또는 ANOVA)을 사용해야 합니다.

시나리오

재무 분석가는 비율 비교 섹션에서 검토하기 시작한 질문에 대답할 수 있습니다. 유형(제약 또는 컴퓨터)에 따라 회사의 규모가 회사의 수익에 미치는 영향이 달라집니까?

이 질문에 대답하려면 다음 두 변수를 기준으로 회사 수익을 비교하십시오 .

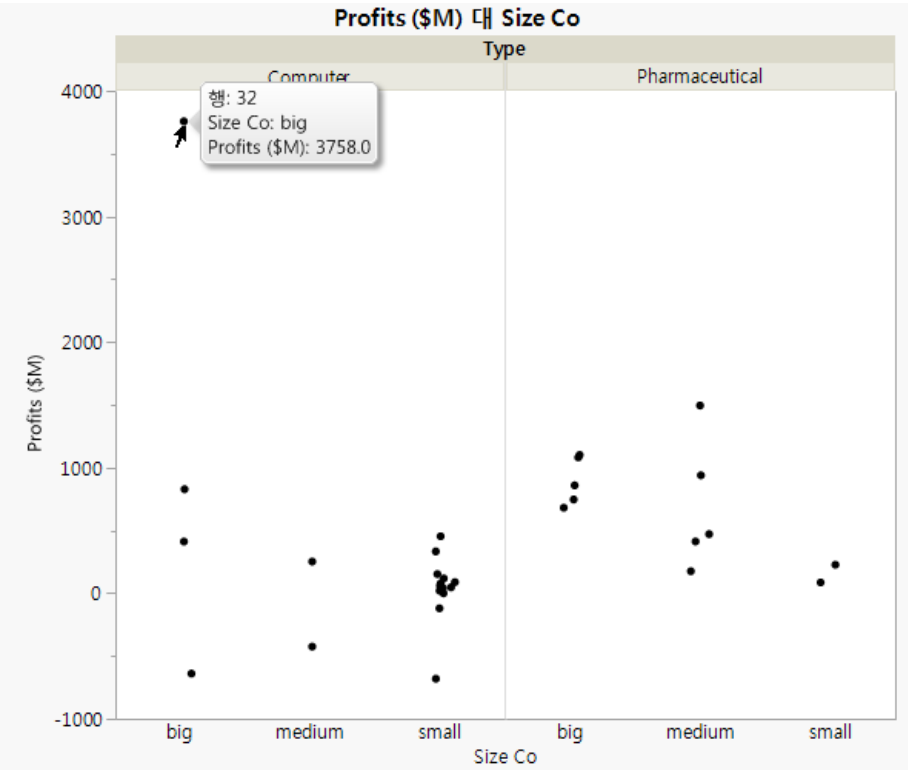
- Type(Pharmaceutical 또는 Computer)
- Size(small, medium, big)

관계 발견

유형과 크기의 모든 조합에 대해 수익 간의 차이를 시각화하려면 그래프를 사용하십시오 .

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다 .
2. **그래프** > **그래프 빌더**를 선택합니다 . " 그래프 빌더 " 창이 나타납니다 .
3. Profits ($M) 를 클릭하고 **Y** 영역으로 드래그하여 놓습니다 .
4. Size Co 를 클릭하고 **X** 영역으로 드래그하여 놓습니다 .
5. Type 을 클릭하고 **그룹 X** 영역으로 드래그하여 놓습니다 .

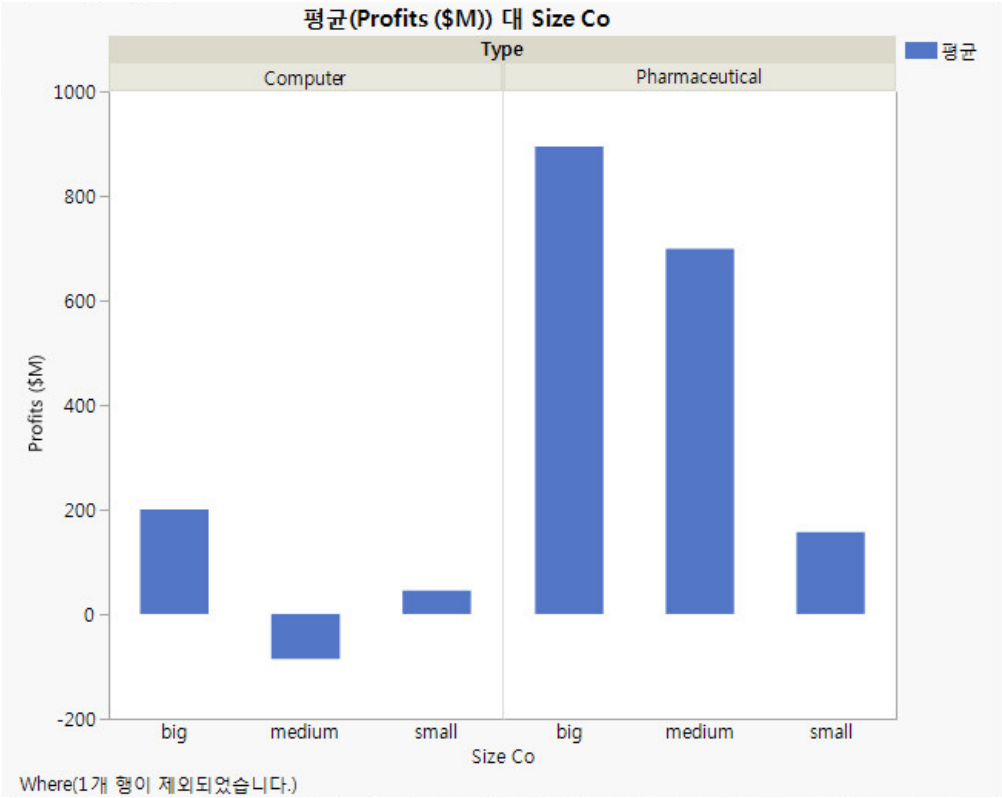
그림 5.21 회사 수익 그래프



그래프에서는 한 대규모 회사의 수익이 매우 크다는 것을 보여 줍니다 . 이 이상치는 그래프의 범위를 늘려 다른 데이터 점을 비교하기 어렵게 만듭니다 .

6. 이 이상치를 선택한 후 마우스 오른쪽 버튼을 클릭하고 **행 > 행 제외**를 선택합니다 . 해당 점이 제거되고 그래프 범위가 자동으로 업데이트됩니다 .
7. 막대 아이콘을 클릭합니다. 막대 차트를 사용하면 점을 사용할 때보다 쉽게 평균 수익을 비교할 수 있습니다 .

그림 5.22 이상치가 제거된 그래프



업데이트된 그래프는 제약 회사의 평균 수익이 더 높다는 것을 보여 줍니다. 이 그래프는 또한 제약 회사만 회사 규모에 따라 수익이 달라진다는 것을 보여 줍니다. 한 변수(회사 규모)의 효과가 다른 변수(회사 유형)의 여러 수준에 대해 변경될 때 이를 **교호작용**이라고 합니다.

관계 수량화

이 데이터는 표본일 뿐이므로 재무 분석가는 다음을 판단해야 합니다 .

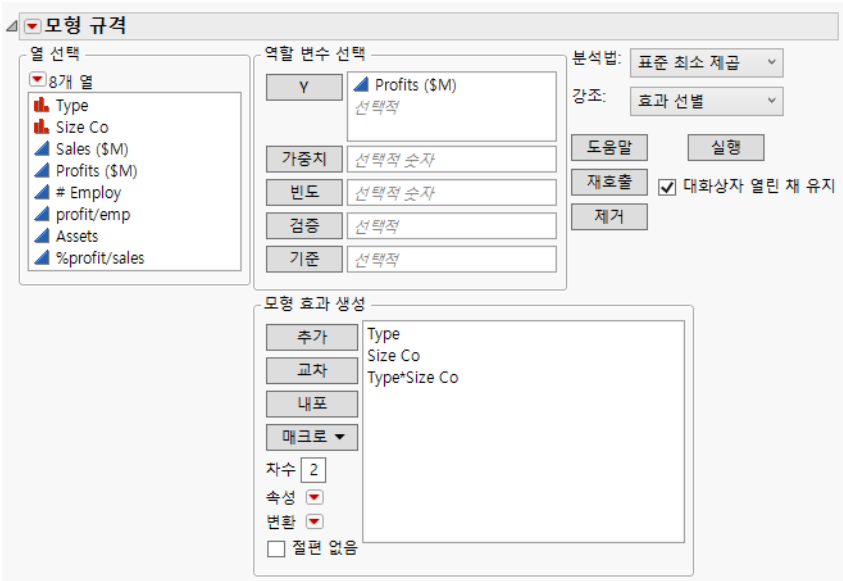
- 차이가 이 표본에 국한되고 우연에 기인하는지 여부

  또는

- 더 광범위한 모집단에 동일한 패턴이 존재하는지 여부

1. 이상치 데이터 점이 제거된 Companies.jmp 샘플 데이터로 돌아갑니다 . 자세한 내용은 " 관계 발견 "(154 페이지) 에서 확인하십시오 .
2. **분석** > **모형 적합**을 선택합니다 .
3. Profits ($M) 를 선택하고 **Y** 를 클릭합니다 .
4. Type 과 Size Co 를 모두 선택합니다 .
5. **매크로** 버튼을 클릭하고 **완전 요인**을 선택합니다 .
6. " 강조 " 메뉴에서 **효과 선별**을 선택합니다 .
7. **대화상자 열린 채 유지** 옵션을 선택합니다 .

그림 5.23 완료된 모형 적합 창



8. **실행**을 클릭합니다 . 보고서 창에 모형 결과가 나타납니다 .

수익의 차이가 실제로 발생했는지 아니면 우연히 발생했는지를 판단하려면 **효과 검정** 보고서를 검토해야 합니다.

참고 : 모든 모형 적합 결과에 대한 자세한 내용은 Fitting Linear Models 의 에서 확인하십시오 .

효과 검정 보기

"효과 검정" 보고서(그림 5.24)는 통계 검정 결과를 보여 줍니다. "모형 적합" 창에는 모형에 포함된 "Type", "Size Co" 및 "Type*Size Co" 효과에 대한 검정이 있습니다.

그림 5.24 효과 검정 보고서

효과 검정

소스	N모수	DF	제곱합	F 비	Prob > F
Type	1	1	1401847.4	10.1368	0.0039*
Size Co	2	2	724616.2	2.6198	0.0927
Type*Size Co	2	2	448061.5	1.6200	0.2180

먼저, 모형에서의 교호작용에 대한 검정인 "Type*Size Co" 효과를 살펴봅니다. 그림 5.22 에서는 제약 회사의 규모에 따라 수익이 다르게 나타난다는 것을 보여 주었습니다. 그러나 이 효과 검정은 수익과 관련하여 유형과 규모 간에 교호작용이 없음을 나타냅니다. p 값 0.218 은 매우 큰 값으로, 유의 수준 0.05 보다도 큽니다. 따라서 해당 효과를 모형에서 제거하고 모형을 다시 실행합니다.

1. " 모형 적합 " 창으로 돌아갑니다.
2. " 모형 효과 생성 " 상자에서 **Type*Size Co** 효과를 선택하고 **제거**를 클릭합니다.
3. **실행**을 클릭합니다.

그림 5.25 업데이트된 효과 검정 보고서

효과 검정

소스	N모수	DF	제곱합	F 비	Prob > F
Type	1	1	1356297.9	9.3768	0.0049*
Size Co	2	2	434161.3	1.5008	0.2410

"Size Co" 효과의 p 값은 크며, 이는 더 광범위한 모집단에서 크기에 따른 차이가 없음을 나타냅니다. "Type" 효과의 p 값은 작아서, 컴퓨터와 제약 회사 간의 데이터에서 확인된 차이가 우연히 발생한 것이 아님을 나타냅니다.

결론

재무 분석가는 유형 (제약 또는 컴퓨터) 에 따라 회사의 규모가 회사의 수익에 미치는 영향에 차이가 있는지 알아보려고 했습니다. 이제 재무 분석가는 질문에 다음과 같이 답할 수 있습니다.

- 더 광범위한 모집단에서 컴퓨터와 제약 회사 간의 수익에는 실제로 차이가 있습니다.
- 회사의 규모와 유형 및 수익 간에는 상관관계가 없습니다.

다중 예측 변수가 있는 회귀 사용

"하나의 예측 변수가 있는 회귀 사용"(143 페이지) 섹션에서는 하나의 예측 변수와 하나의 반응 변수로 구성된 단순 회귀 모형을 생성하는 방법을 소개했습니다. **다중 회귀**는 둘 이상의 예측 변수를 사용하여 평균 반응 변수를 예측합니다.

시나리오

이 예에서는 초코바의 영양 정보가 포함된 Candy Bars.jmp 데이터 테이블을 사용합니다.

영양사는 다음 정보를 사용하여 칼로리를 예측하려고 합니다 .

- 총 지방
- 탄수화물
- 단백질

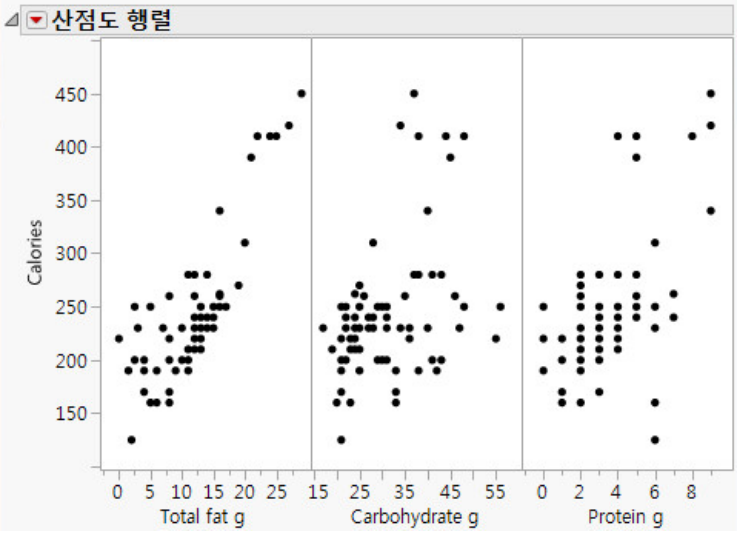
다중 회귀를 사용하여 이 세 가지 예측 변수를 통해 평균 반응 변수를 예측합니다.

관계 발견

칼로리와 총 지방 , 탄수화물 및 단백질 간의 관계를 시각화하려면 산점도 행렬을 생성합니다 .

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Candy Bars.jmp 를 엽니다 .
2. **그래프** > **산점도 행렬**을 선택합니다 .
3. Calories 를 선택하고 **Y, 열**을 클릭합니다 .
4. Total fat g, Carbohydrate g 및 Protein g 를 선택하고 **X** 를 클릭합니다 .
5. **확인**을 클릭합니다 .

그림 5.26 산점도 행렬 결과



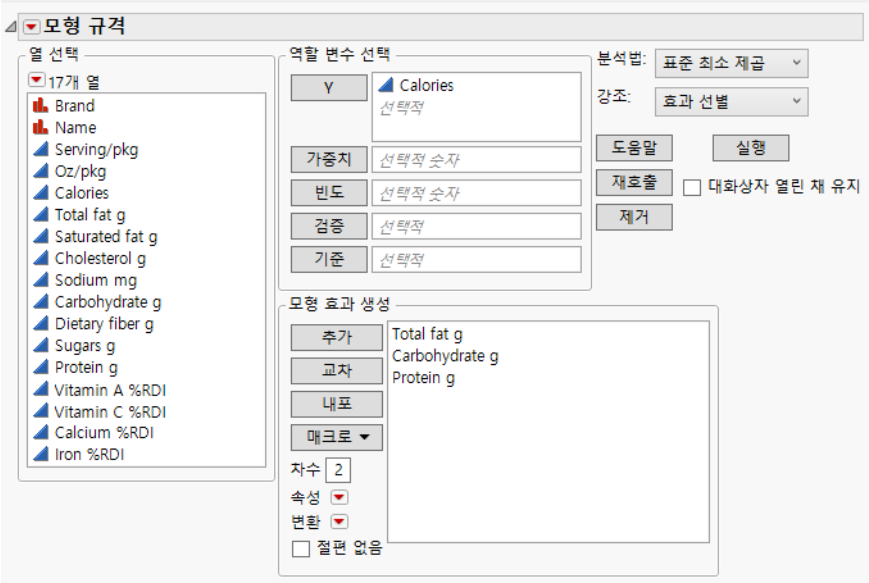
산점도 행렬은 칼로리와 세 변수 간에 양의 상관관계가 있음을 보여 줍니다. 칼로리와 총 지방 간의 상관관계가 가장 강합니다. 이제 영양사는 관계가 있다는 것을 알고 있으므로 평균 칼로리를 예측하기 위해 다중 회귀 모형을 생성할 수 있습니다.

다중 회귀 모형 생성

Candy Bars.jmp 샘플 데이터 테이블을 계속 사용합니다 .

1. **분석** > **모형 적합**을 선택합니다 .
2. Calories 를 선택하고 **Y** 를 클릭합니다 .
3. Total fat g, Carbohydrate g 및 Protein g 를 선택하고 **추가**를 클릭합니다 .
4. " 강조 " 옆에 있는 **효과 선별**을 선택합니다 .

그림 5.27 모형 적합 창



5. **실행**을 클릭합니다 .

보고서 창에 모형 결과가 나타납니다 . 모형 결과를 해석하려면 다음 영역에 중점을 둡니다 .

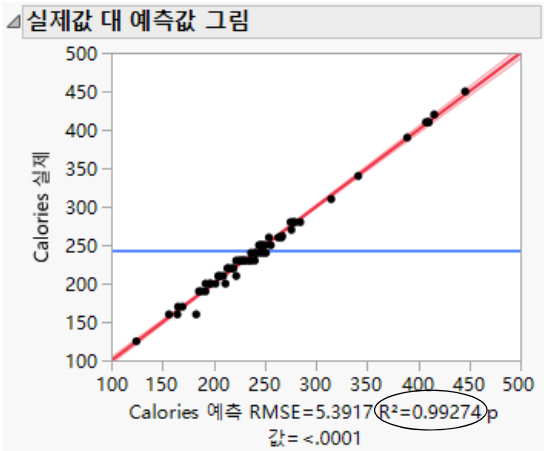
- " 실제값 대 예측값 그림 보기 "(159 페이지)
- " 모수 추정값 해석 "(160 페이지)
- " 예측 프로파일러 사용 "(161 페이지)

참고 : 모든 모형 결과에 대한 자세한 내용은 Fitting Linear Models 의 에서 확인하십시오 .

실제값 대 예측값 그림 보기

실제값 대 예측값 그림은 실제 칼로리와 예측된 칼로리를 보여 줍니다. 예측값이 실제값에 가까워질수록 산점도의 점이 빨간색 선 주위로 모입니다(그림 5.28). 점이 모두 선에 매우 가깝기 때문에 모형이 선택한 요인을 기반으로 칼로리를 정확하게 예측한다는 것을 알 수 있습니다.

그림 5.28 실제값 대 예측값 그림



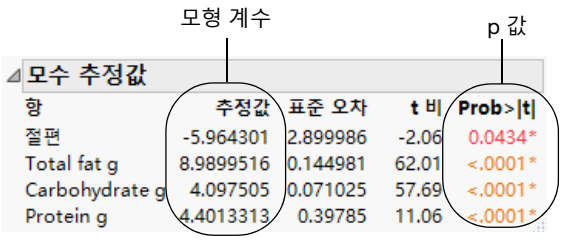
모형 정확도의 또 다른 측도는 R^2 값입니다. 이 값은 그림 5.28의 그림 아래에 나타납니다. R^2 값은 모형에 의해 설명된 대로 칼로리의 변동률을 측정합니다. 1에 가까운 값은 모형이 정확하게 예측한다는 것을 의미합니다. 이 예에서 R^2 값은 0.99입니다.

모수 추정값 해석

" 모수 추정값 " 보고서는 다음과 같은 정보를 보여 줍니다 .

- 모형 계수
- 각 모수의 *p* 값

그림 5.29 모수 추정값 보고서



항	추정값	표준 오차	t 비	Prob>|t|
절편	-5.964301	2.899986	-2.06	0.0434*
Total fat g	8.9899516	0.144981	62.01	<.0001*
Carbohydrate g	4.097505	0.071025	57.69	<.0001*
Protein g	4.4013313	0.39785	11.06	<.0001*

이 예에서 *p* 값은 모두 매우 작습니다(<.0001). 이것은 칼로리를 예측할 때 세 가지 효과(지방, 탄수화물 및 단백질)가 모두 유의하게 기여함을 나타냅니다.

모형 계수를 사용하여 지방 , 탄수화물 및 단백질의 특정 값에 대한 칼로리 값을 예측할 수 있습니다 . 예를 들어 다음 특성을 가진 초코바의 평균 칼로리를 예측한다고 가정해 보겠습니다 .

- 지방 = 11g
- 탄수화물 = 43g

- 단백질 = 2g

이 값을 사용하여 예측 평균 칼로리를 다음과 같이 계산할 수 있습니다.

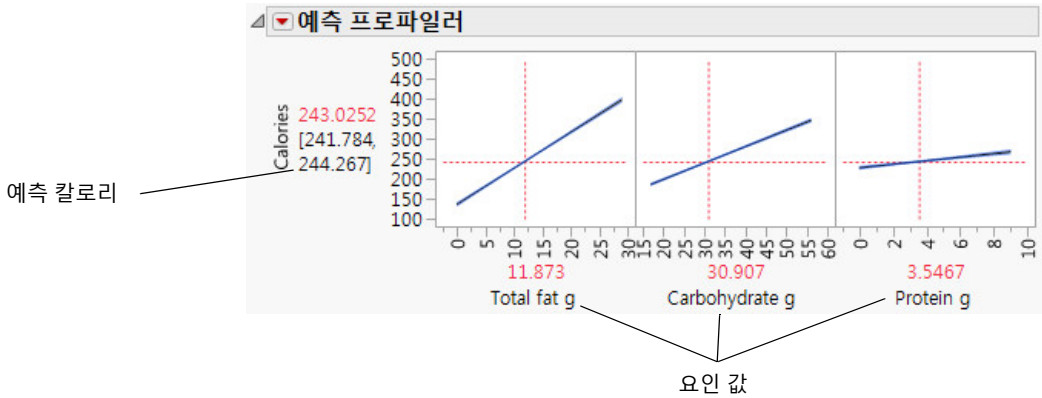
277.92 = -5.9643 + 8.99*11 + 4.0975*43 + 4.4013*2

이 예의 특성은 "Milky Way" 초코바(데이터 테이블의 59 행)와 동일합니다. "Milky Way"의 실제 칼로리는 280 이며 모형이 정확하게 예측한다는 것을 보여 줍니다.

예측 프로파일러 사용

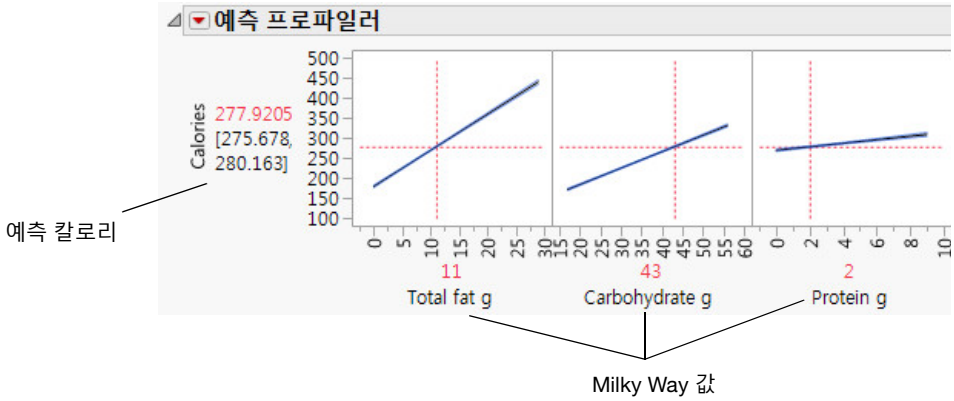
예측 프로파일러를 사용하여 요인의 변화가 예측값에 어떻게 영향을 미치는지 확인할 수 있습니다. 프로파일 선은 요인이 바뀜에 따라 변화되는 칼로리 크기를 보여 줍니다. Total fat g 의 선이 가장 가파르며, 이는 총 지방의 변화가 칼로리에 가장 큰 영향을 미친다는 것을 의미합니다.

그림 5.30 예측 프로파일러



각 요인의 수직선을 클릭하고 드래그하여 예측값이 어떻게 변하는지 확인하십시오. 현재 요인 값을 클릭하고 변경할 수도 있습니다. 예를 들어 요인 값을 클릭하고 "Milky Way" 초코바(59 행)의 값을 입력합니다.

그림 5.31 Milky Way 의 요인 값



참고 : 예측 프로파일러에 대한 자세한 내용은 Profilers 의 에서 확인하십시오 .

결론

영양사는 이제 총 지방, 탄수화물 및 단백질을 기준으로 초코바의 칼로리를 예측할 수 있는 적절한 모형을 확보했습니다.

6 장

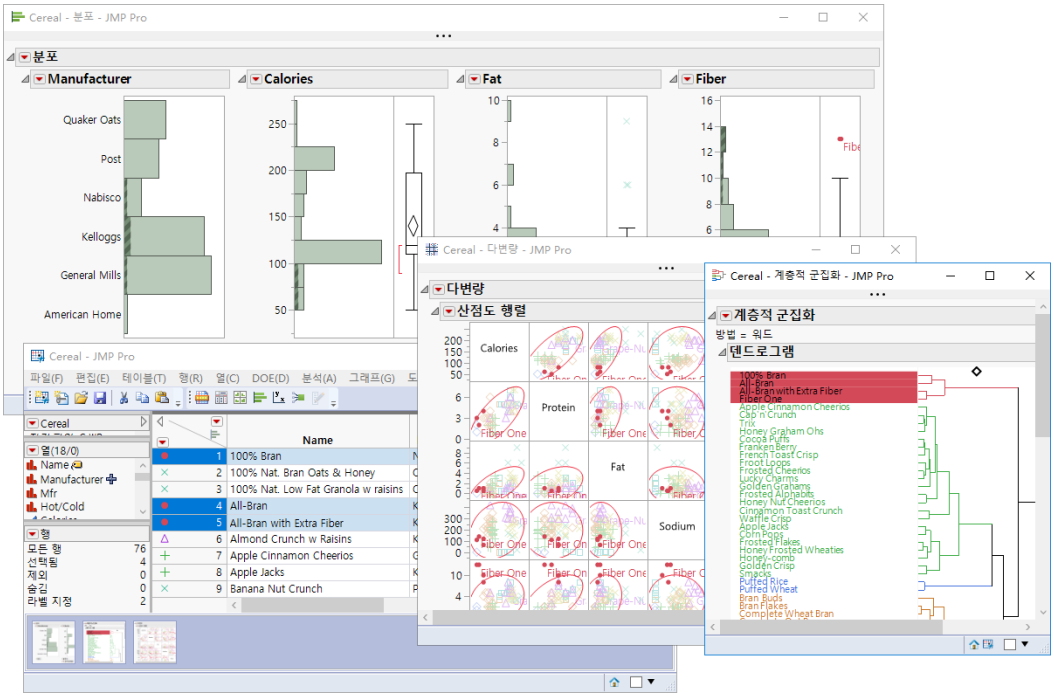
전체를 보는 시각

여러 플랫폼에서 데이터 탐색

JMP는 데이터의 다양한 측면을 탐색하는 데 도움이 되는 통계적 발견 플랫폼의 호스트를 제공합니다. 먼저 히스토그램에서 개별 변수를 간단하게 살펴본 다음 다변량 및 군집 분석으로 진행하여 더욱 자세히 파악할 수 있습니다. 각 단계를 거치면서 데이터에 대해 더 많이 알게 됩니다.

이 장에서는 JMP와 함께 설치되는 샘플 데이터 테이블 Cereal.jmp의 분석 과정을 단계별로 설명합니다. 이를 통해 분포, 다변량 및 계층적 군집화 플랫폼에서 데이터를 탐색하는 방법을 살펴봅니다.

그림 6.1 JMP 의 연결된 분석



목차

재미있는 사실 : 연결된 분석 165
여러 플랫폼에서 데이터 탐색 165
분포 플랫폼에서 분포 분석 165
다변량 플랫폼에서 패턴 및 관계 분석 169
군집화 플랫폼에서 유사 값 분석 172

재미있는 사실 : 연결된 분석

JMP 의 강력한 기능 중 하나는 연결된 분석입니다. 생성된 그래프와 보고서는 데이터 테이블을 통해 서로 연결됩니다. 그림 6.1 에 표시된 것처럼 데이터 테이블에서 선택된 데이터는 세 개의 보고서 창에서도 선택됩니다. 연결된 분석 기능 덕분에 한 창에서 데이터를 선택하면 다른 창에서 해당 데이터의 위치를 확인할 수 있습니다. 이 장의 예를 진행하면서 JMP 창을 열어 두면 이러한 상호 작용을 직접 볼 수 있습니다.

여러 플랫폼에서 데이터 탐색

어떤 시리얼이 건강한 식단에 도움이 될까요? Cereal.jmp 샘플 데이터 (인기 시리얼의 상자에서 수집한 실제 데이터)는 섬유질 함량, 칼로리 및 기타 영양 정보에 대한 통계량을 제공합니다. 가장 건강한 시리얼을 파악하려면 히스토그램 및 기술 통계량, 상관관계 및 이상치 탐지, 산점도 및 군집 분석을 단계별로 해석합니다.

분포 플랫폼에서 분포 분석

분포 플랫폼은 히스토그램 , 추가 그래프 및 보고서를 사용하여 단일 변수 (**단변량** 분석) 의 분포를 보여 줍니다 . **단변량**이란 관련 변수가 두 개 (이변량) 또는 여러 개 (다변량) 가 아니라 하나뿐임을 의미합니다 . 그러나 단일 보고서에서 여러 개별 변수의 분포를 검토할 수도 있습니다 . 각 변수의 보고서 내용은 변수가 범주형 (명목형 또는 순서형) 인지 연속형인지에 따라 달라집니다 .

- 범주형 변수의 경우 초기 그래프는 히스토그램입니다 . 히스토그램에는 순서형 또는 명목형 변수의 각 수준에 대한 막대가 표시됩니다 . 보고서에는 개수와 비율이 표시됩니다 .
- 연속형 변수의 경우 초기 그래프에는 히스토그램과 이상치 상자 그림이 표시됩니다 . 히스토그램에는 연속형 변수의 그룹화된 값에 대한 막대가 표시됩니다 . 보고서에는 선택한 사분위수와 요약 통계량이 표시됩니다 .

데이터가 어떻게 분포되어 있는지 알게 되면 적절한 유형의 분석을 계획할 수 있습니다.

참고 : 분포 플랫폼에 대한 자세한 내용은 **기본 분석**의 에서 확인하십시오 .

시나리오

더 건강한 음식을 먹을 수 있도록 시리얼의 영양가를 확인하려고 합니다 . 시리얼 데이터의 분포를 분석하면 다음 질문에 대한 답을 얻을 수 있습니다 .

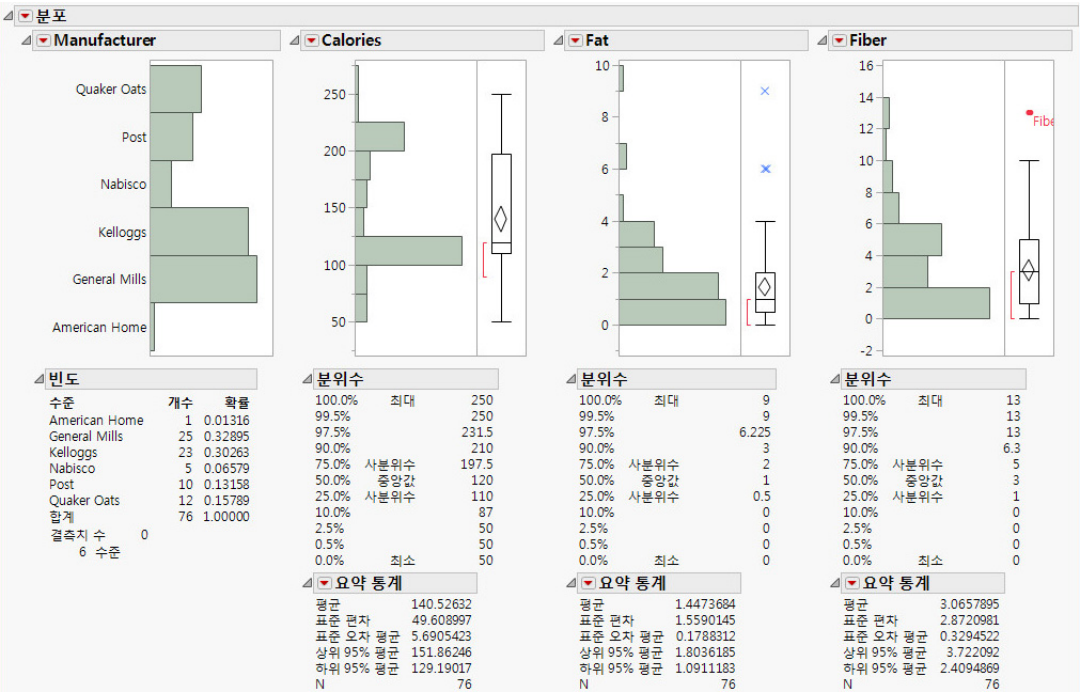
- 어떤 시리얼이 섬유질이 가장 많습니까 ?
- 평균 , 최소 및 최대 칼로리는 얼마입니까 ?
- 지방 함량의 중앙값은 얼마입니까 ?
- 어떤 시리얼이 가장 많은 지방을 함유하고 있습니까 ?

- 데이터에 이상치가 있습니까 ?

분포 생성

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Cereal.jmp 를 엽니다 .
2. **분석 > 분포**를 선택합니다 .
3. Ctrl 키를 누른 채로 Manufacturer, Calories, Fat 및 Fiber 를 클릭합니다 .
4. **Y, 열**을 클릭한 후 **확인**을 클릭합니다 .

그림 6.2 Manufacturer, Calories, Fat 및 Fiber 분포



"Fiber" 분포에서는 다음을 확인할 수 있습니다 .

– "Fiber" 상자 그림에 표시된 것처럼 "Fiber One" 과 "All-Bran with Extra Fiber" 가 가장 많은 섬유질을 함유하고 있습니다 . 이들 시리얼은 섬유질 함량에 있어서 이상치입니다 .

  Cereal.jmp 에서 "Fiber One" 이 포함된 행에는 라벨이 있습니다 . 이 라벨은 그래프의 데이터 점 옆에 시리얼 이름을 표시합니다 . 전체 라벨을 표시하려면 맨 오른쪽 세로 경계선을 오른쪽으로 드래그하십시오 . 라벨이 없는 데이터 점을 커서로 가리키면 "All Bran with Extra Fiber" 가 표시됩니다 .

"Fat" 분포에서는 다음을 확인할 수 있습니다 .

– "Fat" 상자 그림에서 맨 위 데이터 점 (x 표식) 을 커서로 가리키면 "100% Nat. Bran Oats & Honey" 가 지방 함량이 가장 높음을 알 수 있습니다 .

– "Fat" 의 " 분위수 " 보고서에서 지방 함량의 중앙값은 1g 입니다 .

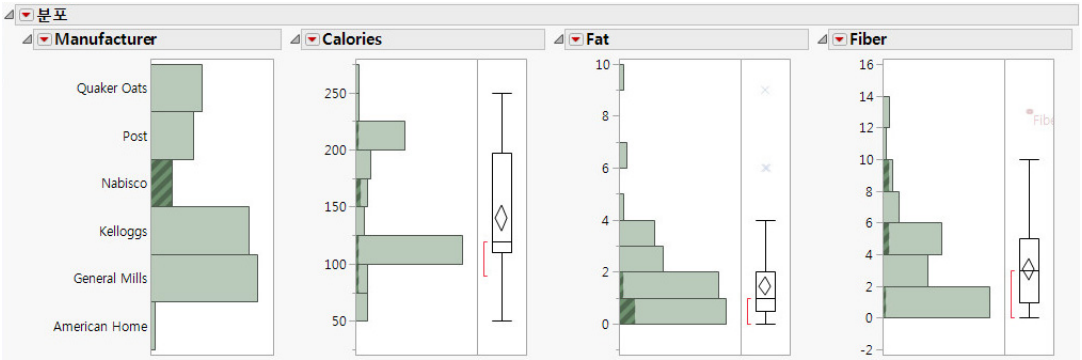
"Calories" 의 " 분위수 " 보고서에서는 다음을 확인할 수 있습니다 .

– 최대 칼로리는 250 입니다 .

– 최소 칼로리는 50 입니다 .

5. "Manufacturer" 히스토그램에서 "Nabisco" 막대를 클릭합니다 .

그림 6.3 Nabisco 시리얼의 분포

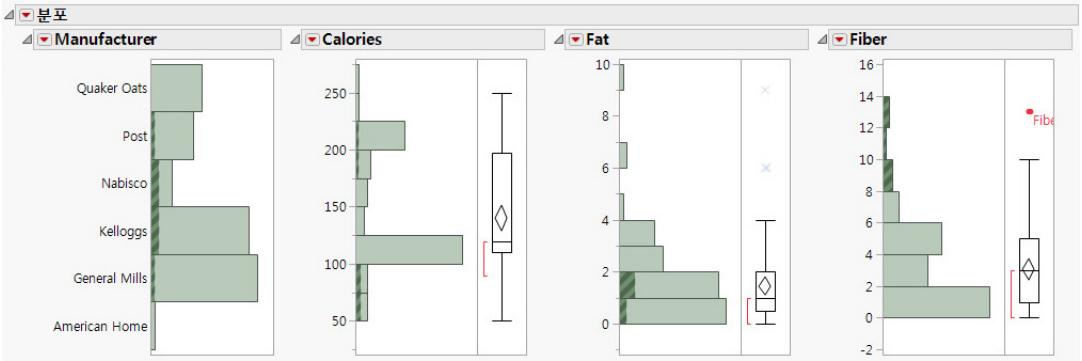


"Nabisco" 시리얼의 칼로리 ("Calories"), 지방 ("Fat") 및 섬유질 ("Fiber") 분포가 다른 히스토그램에서 강조 표시됩니다 . "Nabisco" 시리얼의 칼로리 , 지방 및 섬유질 분포를 전체 데이터의 칼로리 , 지방 및 섬유질 분포와 비교해서 볼 수 있습니다 . 예를 들어 "Nabisco" 시리얼의 지방 분포는 전체 데이터의 지방 분포보다 낮습니다 .

6. 마지막 "Fiber" 막대 아래를 클릭하여 모든 막대를 선택 취소합니다 .

7. Shift 키를 누른 상태로 "Fiber" 히스토그램에서 값이 8 보다 큰 모든 히스토그램 막대를 클릭합니다 .

그림 6.4 섬유질이 많은 시리얼

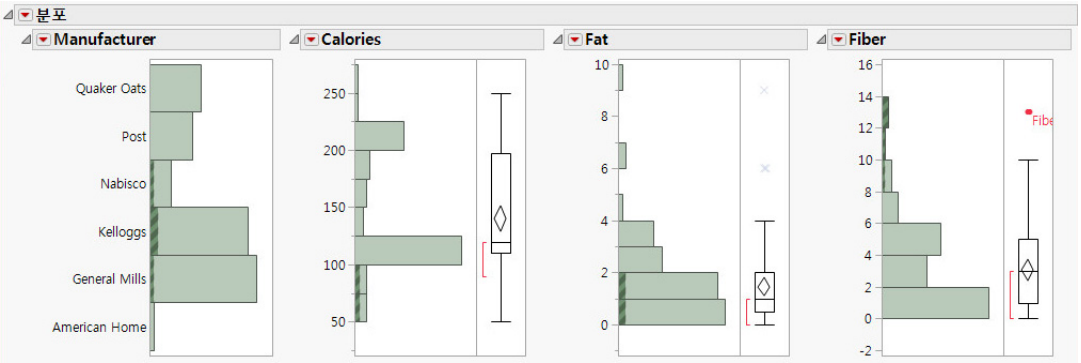


섬유질이 가장 많은 시리얼이 "Calories" 및 "Fat" 히스토그램에서 강조 표시됩니다 . 히스토그램이 연결되어 있기 때문에 섬유질이 많은 시리얼 중 일부는 지방도 적다는 것을 알 수 있습니다 .

8. Ctrl 키와 Shift 키를 누른 상태로 "Calories" 히스토그램에서 값이 200 또는 그 근처에 있는 두 개의 막대를 선택 취소합니다 .

   칼로리가 높은 시리얼이 히스토그램에서 제거됩니다 .

그림 6.5 고섬유질 저지방 시리얼



팁 : " 분포 " 보고서를 열어 두십시오 . 나중에 군집 분석에서 사용할 것입니다 . 자세한 내용은 " 군집화 플랫폼에서 유사 값 분석 "(172 페이지) 에서 확인하십시오 .

결과 해석

결과를 살펴보면 다음 질문에 대한 답을 얻을 수 있습니다 .

어떤 시리얼이 섬유질이 가장 많습니까 ? "Fiber" 상자 그림에서 "All-Bran with Extra Fiber" 와 "Fiber One" 이 가장 많은 섬유질을 함유하고 있음을 알 수 있습니다 . 이 두 시리얼은 이상치입니다 .

평균 , 최소 및 최대 칼로리는 얼마입니까 ? "Calories" 히스토그램에서 칼로리가 50 에서 275 사이임을 알 수 있습니다 . "Calories" 의 " 분위수 " 에서는 칼로리가 50 에서 250 사이이며 , 칼로리의 중앙값이 120 임을 알 수 있습니다 . 이 분포는 균등하지 않습니다 .

지방 함량의 중앙값은 얼마입니까 ? "Fat" 의 " 분위수 " 보고서에서 지방 함량의 중앙값이 1g 임을 알 수 있습니다 .

어떤 시리얼이 가장 많은 지방을 함유하고 있습니까 ? "Fat" 상자 그림에서 "100% Nat. Bran Oats & Honey" 가 지방 함량이 가장 높음을 알 수 있습니다 . 이 시리얼은 이상치입니다 .

결론

섬유질 섭취량을 늘리기 위해 "All-Bran with Extra Fiber" 와 "Fiber One" 을 먹어 보기로 결정했습니다. 이들 시리얼은 칼로리와 지방 함량이 낮습니다. 대부분의 시리얼이 지방 섭취량을 크게

늘리지 않지만, 지방 함량이 높은 "100% Nat. Bran Oats & Honey"는 피하기로 했습니다. 또한 대부분의 시리얼이 지방 함량이 비교적 적지만 칼로리가 반드시 낮은 것은 아닙니다.

다변량 플랫폼에서 패턴 및 관계 분석

이제는 먹어야 할 시리얼과 피해야 할 시리얼을 확인했으므로 시리얼 변수 간에 어떤 관련이 있는지 파악해 보려고 합니다. 다변량 플랫폼을 사용하면 변수 간의 패턴과 관계를 관찰할 수 있습니다. " 다변량 " 보고서를 통해 다음이 가능합니다.

- " 상관 " 테이블을 사용하여 각 반응 변수 쌍 간의 선형 관계 강도를 요약할 수 있습니다.
- " 산점도 행렬 " 을 사용하여 종속성, 이상치 및 군집을 식별할 수 있습니다.
- 다른 기법을 사용하여 부분 상관, 역상관, 쌍별 상관, 공분산 행렬 및 주성분 같은 여러 변수를 검토할 수 있습니다.

참고 : 다변량 플랫폼에 대한 자세한 내용은 Multivariate Methods 의 에서 확인하십시오.

시나리오

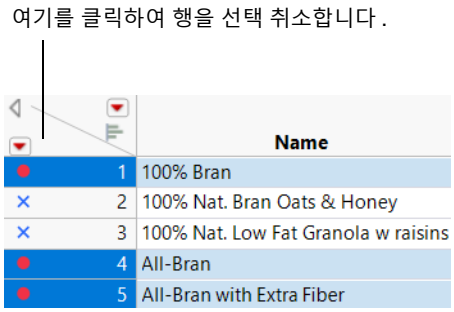
지방과 칼로리 같은 변수 사이의 관계를 확인하려고 합니다. 다변량 플랫폼에서 시리얼 데이터를 분석하면 다음 질문에 대한 답을 얻을 수 있습니다.

- 어떤 변수 쌍이 상관관계가 높습니까?
- 상관관계가 없는 변수 쌍은 무엇입니까?

다변량 보고서 생성

1. Cereal.jmp 데이터 테이블에서 " 열 " 패널 상단에 있는 아래쪽 삼각형을 클릭하여 행을 선택 취소합니다.

그림 6.6 행 선택 취소



2. **분석 > 다변량 방법 > 다변량**을 선택합니다.
3. Calories 부터 Potassium 까지 선택하고 **Y, 열**을 클릭한 후 **확인**을 클릭합니다.

" 다변량 " 보고서가 나타납니다 . 이 보고서에는 기본적으로 " 상관 " 보고서와 " 산점도 행렬 " 이 포함됩니다 . " 상관 " 보고서는 각 반응 변수 (Y) 쌍 간 선형 관계의 강도가 요약된 상관계수의 행렬입니다 . 숫자 색상이 어두울수록 상관관계가 낮음을 나타냅니다 .

그림 6.7 상관 보고서

상관관계

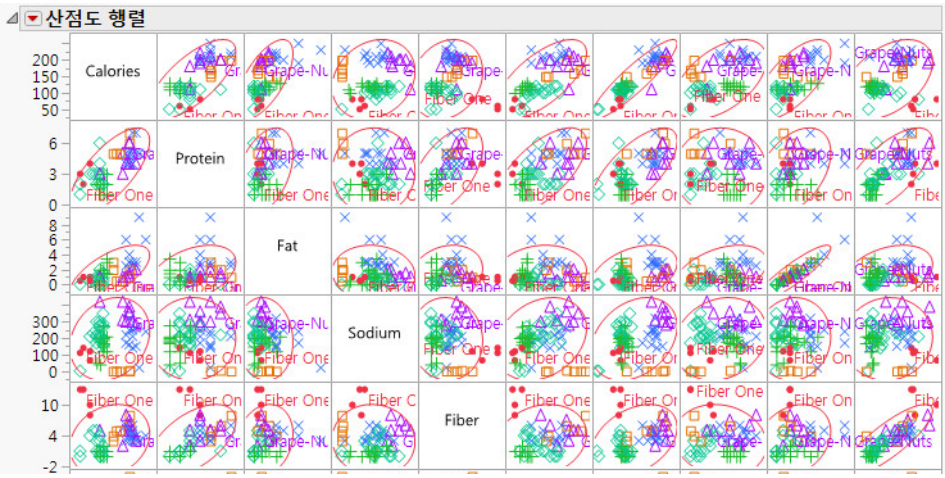
	Calories	Protein	Fat	Sodium	Fiber	Complex Carbos	Tot Carbo	Sugars	Calories fr Fat	Potassium
Calories	1.0000	0.7041	0.6460	0.1996	0.1953	0.6688	0.9076	0.5060	0.6709	0.4451
Protein	0.7041	1.0000	0.4080	0.0050	0.5470	0.6486	0.6937	-0.0010	0.4288	0.6782
Fat	0.6460	0.4080	1.0000	-0.0768	0.1824	0.1037	0.3860	0.4148	0.9013	0.3420
Sodium	0.1996	0.0050	-0.0768	1.0000	-0.0448	0.2619	0.3066	0.1767	0.0572	0.0459
Fiber	0.1953	0.5470	0.1824	-0.0448	1.0000	0.1769	0.3668	-0.1264	0.2553	0.8326
Complex Carbos	0.6688	0.6486	0.1037	0.2619	0.1769	1.0000	0.7773	-0.1601	0.1558	0.2693
Tot Carbo	0.9076	0.6937	0.3860	0.3066	0.3668	0.7773	1.0000	0.4263	0.4636	0.5375
Sugars	0.5060	-0.0010	0.4148	0.1767	-0.1264	-0.1601	0.4263	1.0000	0.4369	0.1166
Calories fr Fat	0.6709	0.4288	0.9013	0.0572	0.2553	0.1558	0.4636	0.4369	1.0000	0.3694
Potassium	0.4451	0.6782	0.3420	0.0459	0.8326	0.2693	0.5375	0.1166	0.3694	1.0000

다음 사항에 유의하십시오 .

- "Calories" 열에서 칼로리는 나트륨 및 섬유질을 제외한 모든 변수와 높은 상관관계가 있습니다 .
- "Fiber" 열에서 섬유질과 칼륨은 높은 상관관계가 있는 것으로 나타납니다 .
- "Sodium" 열에서 나트륨은 다른 변수와 높은 상관관계가 없습니다 .

" 산점도 행렬 " 의 밀도 타원은 변수 간의 관계를 더 자세히 나타냅니다 .

그림 6.8 산점도 행렬의 일부



기본적으로 각 산점도에는 95% 의 이변량 정규 밀도 타원이 있습니다 . 각 변수 쌍이 이변량 정규분포를 따른다고 가정하면 이 타원은 전체 점의 약 95% 를 둘러쌉니다 . 타원이 상당히 둥근 형태이고 대각선 방향이 아니면 해당 변수 간에 상관관계가 없는 것입니다 . 타원이 좁고 대각선 방향이면 해당 변수 간에 상관관계가 높은 것입니다 .

다음 사항을 알 수 있습니다 .

– "Sodium" 행에서는 타원이 상당히 둥근 형태입니다 . 이 모양은 나트륨이 다른 변수와 상관관계가 없음을 나타냅니다 .

– "Fat" 행에서는 "Nat. Bran Oats & Honey", "Cracklin' Oat Bran" 및 "Banana Nut Crunch" 를 나타내는 파란색 x 표식이 타원 바깥쪽에 나타납니다 . 이러한 배치는 해당 데이터가 시리얼의 지방 함량 때문에 이상치임을 나타냅니다 .

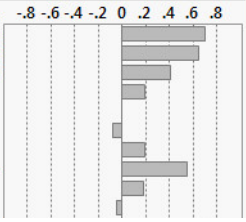
나중에 산점도 행렬을 자세히 살펴볼 것입니다 .

4. " 다변량 " 의 빨간색 삼각형을 클릭하고 **쌍별 상관**을 선택하여 " 쌍별 상관 " 보고서를 표시합니다 .

그림 6.9 쌍별 상관 보고서의 일부

쌍별 상관관계

변수	기준 변수	상관관계	개수	하위 95%	상위 95%	유의성 확률
Protein	Calories	0.7041	76	0.5690	0.8022	<.0001*
Fat	Calories	0.6460	76	0.4922	0.7607	<.0001*
Fat	Protein	0.4080	76	0.2011	0.5801	0.0003*
Sodium	Calories	0.1996	76	-0.0270	0.4068	0.0838
Sodium	Protein	0.0050	76	-0.2207	0.2302	0.9655
Sodium	Fat	-0.0768	76	-0.2971	0.1513	0.5099
Fiber	Calories	0.1953	76	-0.0315	0.4030	0.0908
Fiber	Protein	0.5470	76	0.3668	0.6877	<.0001*
Fiber	Fat	0.1824	76	-0.0449	0.3918	0.1147
Fiber	Sodium	-0.0448	76	-0.2675	0.1825	0.7010



" 쌍별 상관 " 보고서에는 각 Y 변수 쌍에 대한 Pearson 곱적률 상관 계수가 나열됩니다 . 또한 이 보고서는 유의 확률을 보여 주고 막대 차트로 상관관계를 비교합니다 .

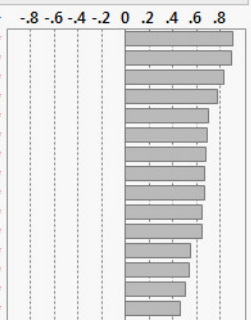
5. 상관관계가 높은 쌍을 빠르게 보려면 보고서를 마우스 오른쪽 버튼으로 클릭하고 **열별 정렬** , **유의 확률** , **오름차순** 체크박스를 선택한 후 **확인**을 클릭합니다 .

가장 관련성이 높은 쌍이 보고서의 맨 위에 나타납니다 . 쌍의 *p* 값이 작은 것은 상관관계의 증거를 나타냅니다 . 가장 유의한 상관관계는 "Tot Carbo"(총 탄수화물) 와 "Calories" 사이의 상관관계입니다 .

그림 6.10 *p* 값이 작은 변수 쌍

쌍별 상관관계

변수	기준 변수	상관관계	개수	하위 95%	상위 95%	유의성 확률
Tot Carbo	Calories	0.9076	76	0.8576	0.9406	<.0001*
Calories fr Fat	Fat	0.9013	76	0.8483	0.9365	<.0001*
Potassium	Fiber	0.8326	76	0.7474	0.8908	<.0001*
Tot Carbo	Complex Carbos	0.7773	76	0.6691	0.8532	<.0001*
Protein	Calories	0.7041	76	0.5690	0.8022	<.0001*
Tot Carbo	Protein	0.6937	76	0.5551	0.7949	<.0001*
Potassium	Protein	0.6782	76	0.5344	0.7838	<.0001*
Calories fr Fat	Calories	0.6709	76	0.5248	0.7786	<.0001*
Complex Carbos	Calories	0.6688	76	0.5221	0.7771	<.0001*
Complex Carbos	Protein	0.6486	76	0.4956	0.7625	<.0001*
Fat	Calories	0.6460	76	0.4922	0.7607	<.0001*
Fiber	Protein	0.5470	76	0.3668	0.6877	<.0001*
Potassium	Tot Carbo	0.5375	76	0.3551	0.6805	<.0001*
Sugars	Calories	0.5060	76	0.3167	0.6566	<.0001*
Calories fr Fat	Tot Carbo	0.4636	76	0.2659	0.6238	<.0001*



결과 해석

결과를 살펴보면 다음 질문에 대한 답을 얻을 수 있습니다 .

어떤 변수 쌍이 상관관계가 높습니까 ? " 상관 " 보고서와 " 산점도 행렬 " 은 "Calories" 가 "Sodium" 및 "Fiber" 를 제외한 모든 변수와 높은 상관관계에 있음을 보여 줍니다 . " 쌍별 상관 " 보고서는 "Tot Carbo"(총 탄수화물) 와 "Calories" 가 가장 상관관계가 높은 변수 쌍이라는 것을 보여 줍니다 .

상관관계가 없는 변수 쌍은 무엇입니까 ? " 상관 " 보고서와 " 산점도 행렬 " 은 "Sodium" 이 다른 변수와 상관관계가 없음을 보여 줍니다 .

결론

지방 함량이 높은 "100% Nat. Bran Oats & Honey" 를 피해야 한다는 이전 결정에 더 확신을 갖게 되었습니다. "All-Bran with Extra Fiber" 와 "Fiber One" 을 먹어 본다는 것도 현명한 결정이었습니다. 이 두 가지 고섬유질 시리얼에는 칼로리, 지방 및 당분 섭취량을 줄이고 칼륨 섭취량을 높여 주는 추가적인 이점이 있습니다. 또한 칼로리가 높은 고탄수화물 시리얼도 피하기로 결정했습니다.

군집화 플랫폼에서 유사 값 분석

군집화는 여러 변수 중에서 유사한 값을 공유하는 관측값들을 함께 그룹화하는 다변량 기법입니다. 계층적 군집화는 행을 계층적 순서로 결합하여 트리로 표현합니다. 고섬유질과 같은 특정 특성을 지닌 시리얼을 군집으로 그룹화하여 시리얼 간의 유사점을 확인할 수 있습니다.

참고 : 계층적 군집화에 대한 자세한 내용은 Multivariate Methods 의 에서 확인하십시오 .

시나리오

어떤 시리얼들이 서로 유사하고 어떤 시리얼들이 서로 유사하지 않은지 알아보려고 합니다 . 시리얼 데이터의 군집을 분석하면 다음 질문에 대한 답을 얻을 수 있습니다 .

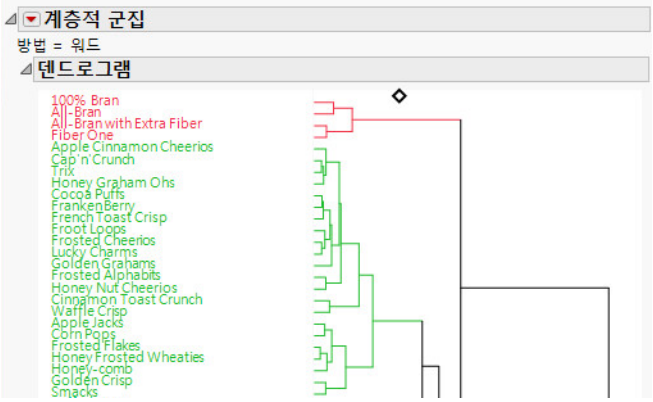
- 어떤 시리얼 군집의 영양가가 적습니까 ?
- 어떤 시리얼 군집이 비타민과 무기질이 풍부하고 당분과 지방은 적습니까 ?
- 어떤 시리얼 군집이 섬유질이 풍부하고 칼로리가 낮습니까 ?

계층적 군집화 그래프 생성

1. Cereal.jmp 가 표시된 상태에서 **분석 > 군집화 > 계층적 군집화**를 선택합니다 .
2. Calories 부터 Enriched 까지 선택하고 **Y, 열**을 클릭한 후 **확인**을 클릭합니다 .

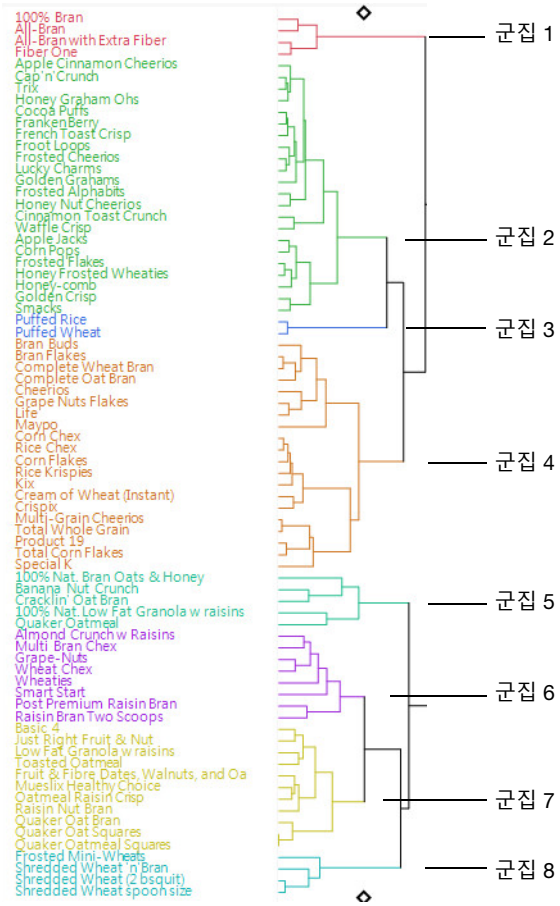
   " 계층적 군집화 " 보고서가 나타납니다 . 데이터 테이블 행 상태에 따라 군집에 색상이 적용됩니다 .

그림 6.11 계층적 군집화 보고서의 일부



3. " 계층적 군집화 " 의 빨간색 삼각형을 클릭하고 **군집 색 표시**를 선택합니다 .
   덴드로그램의 관계에 따라 군집에 색상이 적용됩니다 .

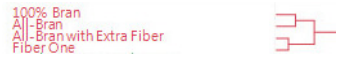
그림 6.12 색상이 적용된 군집



각 군집 내의 시리얼은 유사한 특성을 갖습니다 . 예를 들어 군집 1 에 있는 시리얼은 시리얼 이름으로 판단할 때 섬유질이 많다고 추측할 수 있습니다 .

"All-Bran with Extra Fiber" 와 "Fiber One" 이 군집 1 로 그룹화된 방식에 주목하십시오 . 이들 시리얼은 해당 군집 내 다른 두 시리얼보다 더 유사합니다 .

그림 6.13 군집 1 의 유사한 시리얼



4. 군집 1 을 선택하려면 오른쪽의 빨간색 수평선을 클릭합니다 .

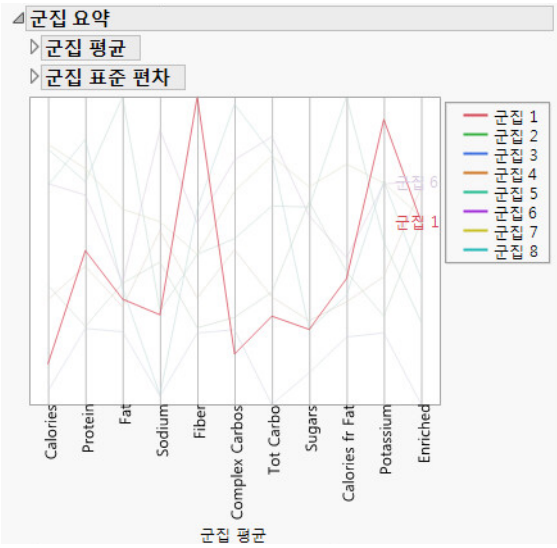
   4 개의 시리얼이 빨간색으로 강조 표시됩니다 .

그림 6.14 군집 선택

5. 군집 내 유사 특성을 확인하려면 " 계층적 군집화 " 의 빨간색 삼각형을 클릭하고 **군집 요약**을 선택합니다 .

   보고서 하단의 " 군집 요약 " 그래프에 각 군집의 변수별 평균 값이 표시됩니다 . 예를 들어 이 군집의 시리얼은 다른 군집의 시리얼보다 섬유질 및 칼륨 함량이 많습니다 .

그림 6.15 군집 요약

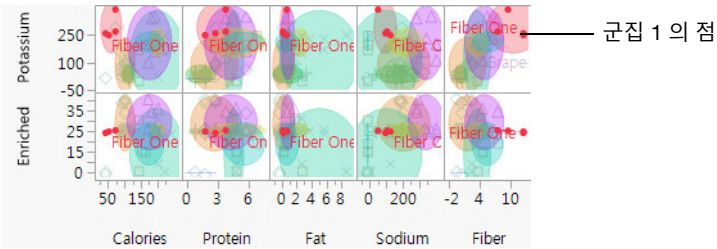


6. " 계층적 군집화 " 의 빨간색 삼각형을 클릭하고 **산점도 행렬**을 선택합니다 .

   다변량 플랫폼에서 산점도 행렬을 만드는 대신 이 옵션을 사용할 수 있습니다 .

   "Potassium" 행의 "Fiber" 그림에 주목하십시오 . 선택한 시리얼은 그림 오른쪽의 8g 과 13g 사이에 있습니다 . 이 위치는 군집 1 의 시리얼에 섬유질과 칼륨이 많다는 것을 나타냅니다 .

그림 6.16 군집 1 특성



참고 : 해당 점은 이전에 생성한 산점도 행렬이 아직 열려 있으면 여기서도 선택됩니다 .

결과 해석

군집을 각각 클릭하고 " 군집 요약 " 보고서를 보면 다음과 같은 특성을 알 수 있습니다 .

- "Fiber One" 과 "All-Bran" 같은 군집 1 의 시리얼은 섬유질 및 칼륨 함량은 높고 칼로리는 낮습니다 .
- 어린이들이 좋아하는 시리얼이 많이 포함된 군집 2 의 시리얼은 당분 함량이 높고 섬유질 , 복합 탄수화물 및 단백질 함량은 낮습니다 .
- 군집 3 의 시리얼 ("Puffed Rice" 와 "Puffed Wheat") 은 칼로리는 낮지만 영양가가 거의 없습니다 .
- "Total Corn Flakes" 와 "Multi-Grain Cheerios" 같은 군집 4 의 시리얼은 하루에 필요한 비타민과 무기질을 100% 공급합니다 . 이들 시리얼은 지방 , 섬유질 및 당분 함량이 낮습니다 .
- 군집 5 의 시리얼은 단백질 및 지방 함량이 높고 나트륨 함량은 적습니다 . 이 군집은 "Banana Nut Crunch" 와 "Quaker Oatmeal" 같은 시리얼로 구성되어 있습니다 .
- 군집 6 의 시리얼은 지방 함량은 낮고 나트륨 및 탄수화물 함량은 높습니다 . "Wheaties" 와 "Grape-Nuts" 같은 전통적인 시리얼이 이 군집에 속합니다 .
- 군집 7 의 시리얼은 칼로리는 높고 섬유질 함량은 낮습니다 . 말린 과일이 들어 있는 많은 시리얼이 이 군집에 속합니다 ("Mueslix Healthy Choice", "Low Fat Granola w Raisins", "Oatmeal Raisin Crisp", "Raisin Nut Bran" 및 "Just Right Fruit & Nut").
- 군집 8 의 시리얼은 나트륨 및 당분 함량은 낮고 복합 탄수화물 , 단백질 및 칼륨 함량은 풍부합니다 . "Shredded Wheat" 및 "Mini-Wheat" 시리얼이 이 군집에 속합니다 .

덴드로그램에서 결합 상태를 보면 각 군집에서 어떤 시리얼들이 가장 유사한지 알 수 있습니다.

- 군집 1 에서는 "Fiber One" 이 영양가 면에서 "All-Bran with Extra Fiber" 와 유사합니다 . "100% Bran" 과 "All-Bran" 도 유사합니다 . 유사한 각 쌍의 시리얼들은 서로 다른 회사에서 제조되므로 서로 경쟁 관계에 있습니다 .
- 군집 2 에서 "Frosted Flakes" 와 "Honey Frosted Wheaties" 는 하나는 옥수수 플레이크이고 다른 하나는 밀 플레이크임에도 불구하고 서로 유사합니다 . "Lucky Charms" 와 "Frosted Cheerios" 도 유사하고 , "Cap'n'Crunch" 와 "Trix" 도 유사합니다 .

결론

섬유질은 더 많이, 칼로리는 더 적게 섭취하려는 바람에 따라 군집 1의 시리얼을 먹어 보기로 결정합니다. 튀긴 밀과 튀긴 쌀로 구성되고 영양가가 거의 없는 군집 3의 시리얼은 피하려고 합니다. 또한 영양가가 높은 군집 4의 시리얼도 먹어 볼 것입니다.

7 장

작업 저장 및 공유

결과 저장 및 다시 생성

JMP에서는 데이터에서 생성된 결과를 여러 가지 방법으로 다른 사람들과 공유할 수 있습니다. 다음은 작업을 공유할 수 있는 몇 가지 방법입니다.

- 플랫폼 결과를 저널 , 프로젝트 또는 웹 보고서로 저장
- 결과 , 데이터 테이블 및 기타 파일을 프로젝트에 저장
- 스크립트를 저장하여 데이터 테이블에서 결과 재현
- 결과를 대화식 HTML(.htm, html) 로 저장
- 결과를 PowerPoint 프레젠테이션 (.pptx) 으로 저장
- 결과를 대시보드에서 공유

그림 7.1 웹 보고서의 예

목차

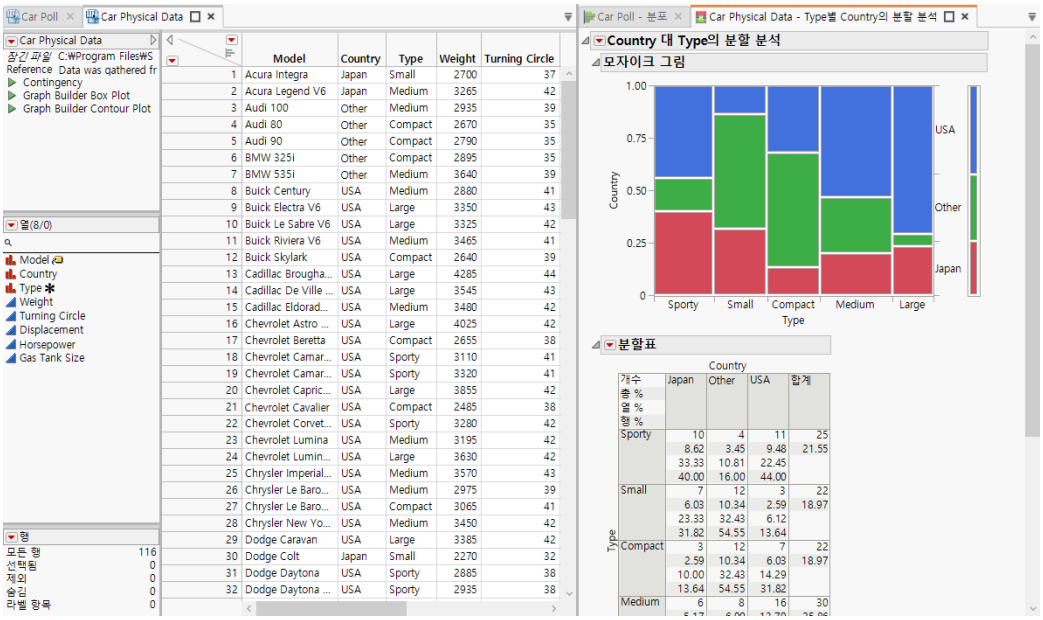
프로젝트 작업 179
새 프로젝트 생성 180
프로젝트에서 파일 열기 180
프로젝트에서 파일 재배열 180
프로젝트 저장 182
프로젝트 작업 영역 182
프로젝트 책갈피 183
프로젝트 내용 184
프로젝트의 최근 사용한 파일 184
프로젝트 로그 185
파일을 프로젝트로 이동 185
프로젝트 내의 열기 스크립트 작성 186
새 프로젝트 생성의 예 186
플랫폼 결과를 저널로 저장 189
예: 저널 생성 189
저널에 분석 추가 190
스크립트 저장 및 실행 191
예: 스크립트 저장 및 실행 191
스크립트 및 JSL 192
보고서를 대화식 HTML 로 저장 192
대화식 HTML의 데이터 포함 193
예: 대화식 HTML 생성 194
보고서를 대화식 HTML로 공유 195
보고서를 PowerPoint 프레젠테이션으로 저장 199
대시보드 생성 199
예: 창 결합 200
예: 두 개의 보고서가 포함된 대시보드 생성 201

프로젝트 작업

JMP 프로젝트를 사용하여 다음을 수행할 수 있습니다.

- 탭 형식의 단일 JMP 창에서 데이터를 더 효율적으로 탐색합니다.
- 관련 파일 및 보고서 집합을 빠르게 저장하고 다시 엽니다.
- 자체 포함 프로젝트 파일에 테이블과 스크립트를 포함하여 작업을 쉽게 공유합니다.

그림 7.2 데이터 테이블 및 보고서가 포함된 프로젝트 파일



이 섹션에는 다음과 같은 정보가 포함되어 있습니다.

- " 새 프로젝트 생성 "
- " 프로젝트에서 파일 열기 "
- " 프로젝트에서 파일 재배열 "
- " 프로젝트 저장 "
- " 프로젝트 작업 영역 "
- " 프로젝트 책갈피 "
- " 프로젝트 내용 "
- " 프로젝트의 최근 사용한 파일 "
- " 프로젝트 로그 "
- " 파일을 프로젝트로 이동 "

- " 프로젝트 내의 열기 스크립트 작성 "
- " 새 프로젝트 생성의 예 "

새 프로젝트 생성

비어 있는 새 JMP 프로젝트를 생성하려면 **파일** > **새로 만들기** > **프로젝트**(Windows) 또는 **파일** > **새로 만들기** > **새 프로젝트**(macOS)를 선택합니다.

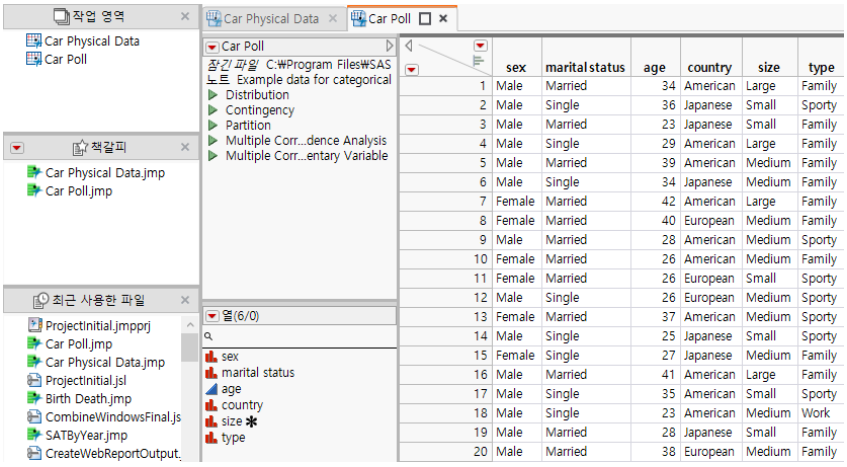
프로젝트에서 파일 열기

JMP 프로젝트에서 데이터 테이블을 열고 데이터에 대한 분석을 실행할 수 있습니다. 각 데이터 테이블과 분석 보고서는 새 탭에서 열립니다.

1. 프로젝트 창에서 **파일** > **열기**를 선택하고 열려는 데이터 테이블로 이동합니다.
2. 데이터 테이블에서 분석을 실행합니다.

컴퓨터에서 데이터 테이블이 업데이트된 경우 프로젝트를 다시 열면 프로젝트의 모든 관련 보고서가 업데이트됩니다.

그림 7.3 초기 프로젝트



팁 : 프로젝트에서 Ctrl+S 또는 Ctrl+W 같은 JMP 바로 가기 키를 사용하여 파일을 저장하거나 활성 창 페인을 닫을 수 있습니다. 전체 목록을 보려면 **도움말** > **빠른 참조 카드**를 선택하십시오.

프로젝트에서 파일 재배열

JMP 프로젝트에서 데이터 테이블과 보고서는 개별 탭에 나타납니다. 탭을 도킹 영역으로 드래그하여 재배열할 수 있습니다.

팁 : 도킹된 항목을 실행 취소하거나 다시 실행하려면 **프로젝트** > **레이아웃 실행 취소** 또는 **프로젝트** > **레이아웃 다시 실행**을 선택합니다 . 모든 테이블과 보고서를 개별 탭으로 되돌리려면 **프로젝트** > **레이아웃 재설정**을 선택합니다 .

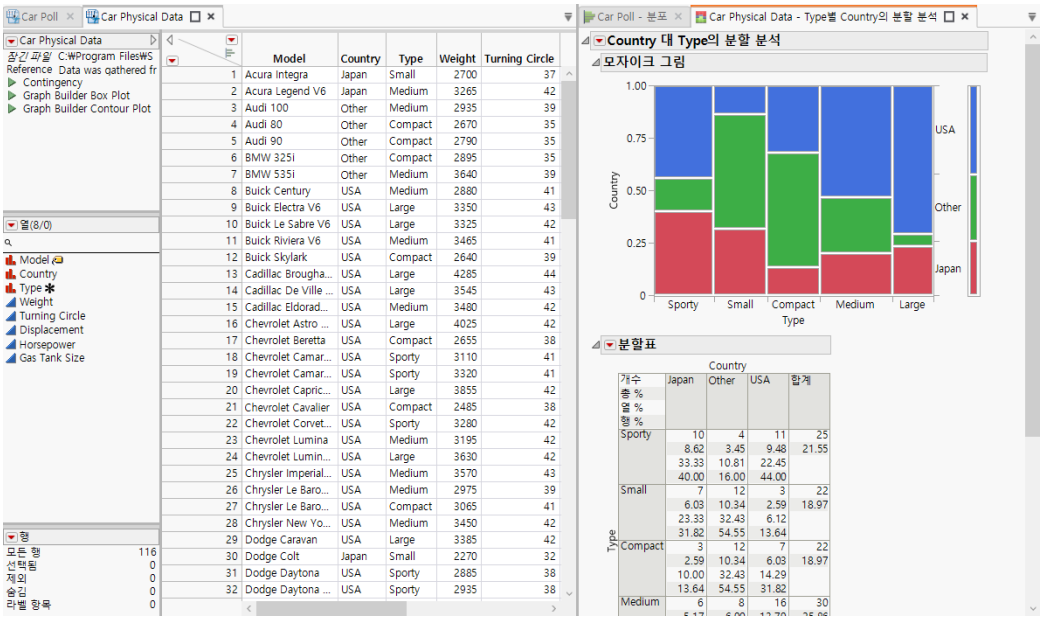
프로젝트 파일 재배열의 예

1. **파일** > **새로 만들기** > **프로젝트** (Windows) 또는 **파일** > **새로 만들기** > **새 프로젝트** (macOS) 를 선택합니다 .
2. **도움말** > **샘플 데이터 라이브러리**를 선택하고 Car Physical Data.jmp 및 Car Poll.jmp를 엽니다.
3. Car Poll.jmp 데이터 테이블에서 Distribution 스크립트를 실행합니다 .
4. Car Physical Data.jmp 데이터 테이블에서 Contingency 스크립트를 실행합니다 .
5. **Car Poll - 분포** 보고서 탭을 오른쪽으로 드래그하여 오른쪽 도킹 영역에 놓습니다 .

   프로젝트 창 오른쪽에 " 분포 " 보고서 페인이 나타납니다 . 보고서가 도킹되어 탭 간에 전환할 때 계속 표시됩니다 .
6. **Car Physical Data - 분할 분석** 탭을 **Car Poll - Distribution** 보고서의 가운데로 드래그하여 탭 도킹 영역에 놓습니다 .

   팁 : 데이터 테이블 창의 크기를 조정하여 두 보고서를 모두 표시할 수 있습니다 .

그림 7.4 탭으로 그룹화된 보고서 및 데이터 테이블



두 데이터 테이블과 마찬가지로 "분포" 보고서와 "분할 분석" 보고서가 탭 형식으로 함께 표시됩니다.

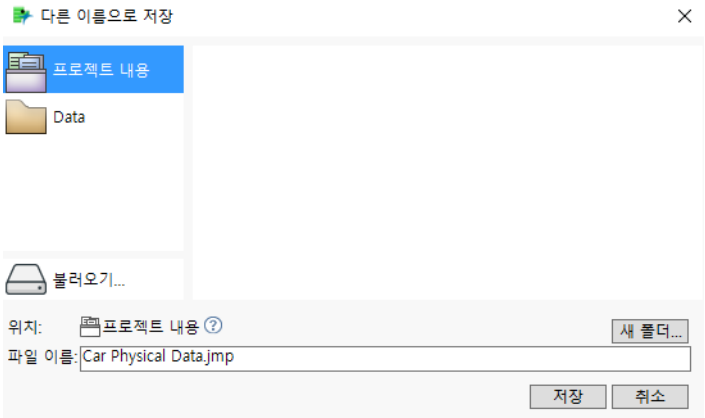
프로젝트 저장

프로젝트를 공유, 배포 또는 보관하려면 모든 데이터 테이블과 스크립트를 프로젝트 내용에 저장하여 자체 포함 프로젝트를 생성합니다. 이렇게 하면 모든 프로젝트 파일과 폴더를 단일 프로젝트 파일에 포함하여 다른 JMP 사용자와 공유할 수 있습니다.

1. (선택 사항) 프로젝트를 자체 포함 프로젝트로 저장하려면 **파일** > **다른 이름으로 저장**을 선택한 후 **프로젝트 내용**을 선택하여 각 데이터 테이블과 스크립트를 프로젝트에 저장합니다. **새 폴더**를 클릭하여 프로젝트 내용에 폴더를 생성할 수 있습니다.

   참고 : 보고서는 프로젝트에 자동 저장되므로 저장할 필요가 없습니다.

그림 7.5 프로젝트 내용에 데이터 테이블 저장



2. **파일** > **프로젝트 저장**을 선택하여 프로젝트 파일을 저장합니다.

그림 7.6 프로젝트 파일 저장

프로젝트 작업 영역

JMP 프로젝트에서 현재 열려 있는 모든 파일을 "작업 영역" 도구 페인에서 볼 수 있습니다. 보고서는 해당 데이터 테이블 아래에 나타납니다. 활성 데이터 테이블은 굵게 표시됩니다.

팁 : " 작업 영역 " 도구 페인을 표시하거나 숨기려면 **프로젝트** > **작업 영역 표시**를 선택합니다 . 프로젝트를 생성할 때 기본적으로 표시되는 도구 페인을 지정하려면 **파일** > **환경 설정** > **프로젝트**를 선택하고 초기 도구 페인을 선택합니다 .

"작업 영역" 도구 페인에서 다음을 수행할 수 있습니다.

- 파일을 두 번 클릭하여 활성화합니다 .
- 파일을 마우스 오른쪽 버튼으로 클릭하여 파일 닫기 , 숨기기 (탭을 제거하고 파일 이름을 흐리게 표시), 책갈피 지정 등을 수행합니다 .

프로젝트 책갈피

JMP 프로젝트에서 파일 또는 폴더의 바로 가기를 " 책갈피" 도구 페인에 생성할 수 있습니다. 책갈피 그룹을 생성하여 책갈피 지정된 파일과 폴더를 구성할 수도 있습니다.

팁 : " 책갈피 " 도구 페인을 표시하거나 숨기려면 **프로젝트** > **책갈피 표시**를 선택합니다 . 프로젝트를 생성할 때 기본적으로 표시되는 도구 페인을 지정하려면 **파일** > **환경 설정** > **프로젝트**를 선택하고 초기 도구 페인을 선택합니다 .

프로젝트에서 열린 파일 책갈피 지정

- 탭을 마우스 오른쪽 버튼으로 클릭하고 **책갈피**를 선택합니다 .
- 열려 있는 모든 파일을 책갈피 지정하려면 탭을 마우스 오른쪽 버튼으로 클릭하고 **모두 책갈피**를 선택합니다 .

컴퓨터의 파일 또는 폴더 책갈피 지정

컴퓨터의 파일 또는 폴더를 " 책갈피" 도구 페인으로 드래그하거나, " 책갈피" 의 빨간색 삼각형 메뉴를 클릭하고 **파일 추가** 또는 **폴더 추가**를 선택합니다.

참고 : 컴퓨터의 책갈피 지정된 폴더에서 파일을 추가하거나 제거하면 폴더 내용이 프로젝트에서 자동으로 업데이트됩니다 .

책갈피 지정된 파일 열기

" 책갈피" 도구 페인에서 파일을 두 번 클릭합니다.

책갈피 제거

" 책갈피" 도구 페인에서 파일을 마우스 오른쪽 버튼으로 클릭하고 **책갈피 제거**를 선택합니다.

책갈피 그룹 생성

1. " 책갈피 " 의 빨간색 삼각형 메뉴를 클릭하고 **새 그룹**을 선택합니다 .

2. 그룹 이름을 지정하고 **확인**을 클릭합니다 .
3. 새 책갈피 또는 기존 책갈피를 그룹으로 드래그하거나 , " 책갈피 " 도구 페인에서 그룹을 마우스 오른쪽 버튼으로 클릭하고 **파일 추가** 또는 **폴더 추가**를 선택합니다 .

프로젝트 내용

자체 포함 JMP 프로젝트에서 프로젝트 내용에 저장된 모든 파일이 "내용" 도구 페인에 나타납니다.

팁 : " 내용 " 도구 페인을 표시하거나 숨기려면 **프로젝트** > **내용 표시**를 선택합니다 . 프로젝트를 생성할 때 기본적으로 표시되는 도구 페인을 지정하려면 **파일** > **환경 설정** > **프로젝트**를 선택하고 초기 도구 페인을 선택합니다 .

프로젝트 내용에 있는 파일 열기

"내용" 도구 페인에서 파일을 두 번 클릭합니다.

프로젝트 내용에 폴더 생성

1. " 내용 " 의 빨간색 삼각형 메뉴를 클릭하고 **새 폴더**를 선택합니다 .
2. 폴더 이름을 지정하고 **확인**을 클릭합니다 .

파일을 폴더로 이동

"내용" 도구 페인에서 파일을 폴더로 드래그합니다.

컴퓨터의 파일 또는 폴더를 프로젝트 내용에 복사

1. " 내용 " 의 빨간색 삼각형 메뉴를 클릭하고 **프로젝트에 파일 복사** 또는 **프로젝트에 폴더 복사**를 선택합니다 .
2. 파일을 탐색한 후 **열기**를 클릭하거나 폴더를 탐색한 후 **선택**을 클릭합니다 .

프로젝트 내용에 있는 파일 이름 바꾸기

"내용" 도구 페인에서 파일을 마우스 오른쪽 버튼으로 클릭하고 **이름 바꾸기**를 선택합니다.

프로젝트 내용에 있는 파일 삭제

"내용" 도구 페인에서 파일을 마우스 오른쪽 버튼으로 클릭하고 **삭제**를 선택합니다.

프로젝트의 최근 사용한 파일

JMP 프로젝트의 "최근 사용한 파일" 도구 페인에서 파일을 열 수 있습니다.

팁 : " 최근 사용한 파일 " 도구 페인을 표시하거나 숨기려면 **프로젝트** > **최근 사용한 파일 표시**를 선택합니다 . 프로젝트를 생성할 때 기본적으로 표시되는 도구 페인을 지정하려면 **파일** > **환경 설정** > **프로젝트**를 선택하고 초기 도구 페인을 선택합니다 .

파일 열기

"최근 사용한 파일" 도구 페인에서 파일을 두 번 클릭하거나 프로젝트로 드래그합니다.

참고: 프로젝트 내용에 저장된 열린 파일은 "최근 사용한 파일" 도구 페인에 표시되지 않습니다.

프로젝트 로그

JMP 프로젝트의 "로그" 도구 페인에서 로그 메시지를 볼 수 있습니다.

팁 :

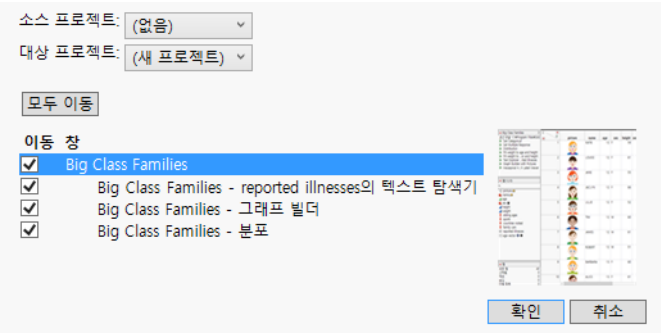
- " 로그 " 도구 페인을 표시하거나 숨기려면 **프로젝트** > **로그 표시**를 선택합니다 . 프로젝트를 생성할 때 기본적으로 표시되는 도구 페인을 지정하려면 **파일** > **환경 설정** > **프로젝트**를 선택하고 초기 도구 페인을 선택합니다 .
- 기본적으로 프로젝트에서 발생하는 로그 메시지는 기본 JMP 로그가 아니라 프로젝트 로그에만 나타납니다 . 이 설정을 변경하려면 **파일** > **환경 설정** > **프로젝트**를 선택하고 **프로젝트 로그 페인 사용**을 **열려 있는 경우** 또는 **안 함**으로 업데이트합니다 .

파일을 프로젝트로 이동

JMP에서는 파일을 프로젝트로 이동하거나 프로젝트 외부 또는 프로젝트 간에 이동할 수 있습니다.

1. **도움말** > **샘플 데이터 라이브러리**를 선택하고 Big Class Families.jmp 를 엽니다 .
2. "Distribution", "Text Explorer" 및 "Graph Builder" 스크립트를 실행합니다 .
3. 어느 창에서든 **창** > **프로젝트 간 이동**을 선택합니다 .
4. 프로젝트에 열려 있지 않은 파일을 보려면 " 소스 프로젝트 " 를 **(없음)** 으로 그대로 둡니다 .
5. 선택한 파일을 새로 생성한 프로젝트로 이동하려면 " 대상 프로젝트 " 를 **(새 프로젝트)** 로 그대로 둡니다 .
6. Big Class Families.jmp 옆의 체크박스를 선택합니다 . 데이터 테이블과 연결된 그래프도 선택됩니다 .

그림 7.7 프로젝트 간 창 이동



7. 확인을 클릭합니다. 선택한 테이블과 보고서로 새 프로젝트가 생성됩니다.

프로젝트 내의 열기 스크립트 작성

"프로젝트 내의 열기 스크립트" 상자에는 이 특정 프로젝트를 열 때마다 실행할 JSL 스크립트를 추가할 수 있습니다. 예를 들어 보고서를 생성 또는 수정하고, 사용자에게 정보를 요청하거나, 사용 정보를 로그 창에 쓰는 스크립트를 생성할 수 있습니다.

1. **프로젝트** > **프로젝트 설정**을 선택합니다.
2. JSL 스크립트를 붙여 넣고 **확인**을 클릭합니다.

팁 :

- 자세한 내용은 Scripting Guide 의 확인하십시오. 여기에서는 특정 프로젝트 대신 모든 프로젝트를 열 때 시작 스크립트를 생성하는 방법을 설명합니다.
- 기존 프로젝트를 새 프로젝트의 템플릿으로 사용하려면 **파일** > **환경 설정** > **프로젝트**에서 기존 프로젝트를 새 프로젝트 템플릿으로 지정합니다.

새 프로젝트 생성의 예

이 예에서는 프로젝트 생성, 데이터 가져오기, 프로젝트 창에서 분석 생성 및 보고서 도킹, 부분 집합 테이블 생성을 차례로 수행한 후 프로젝트를 저장하고, 닫았다가 다시 엽니다.

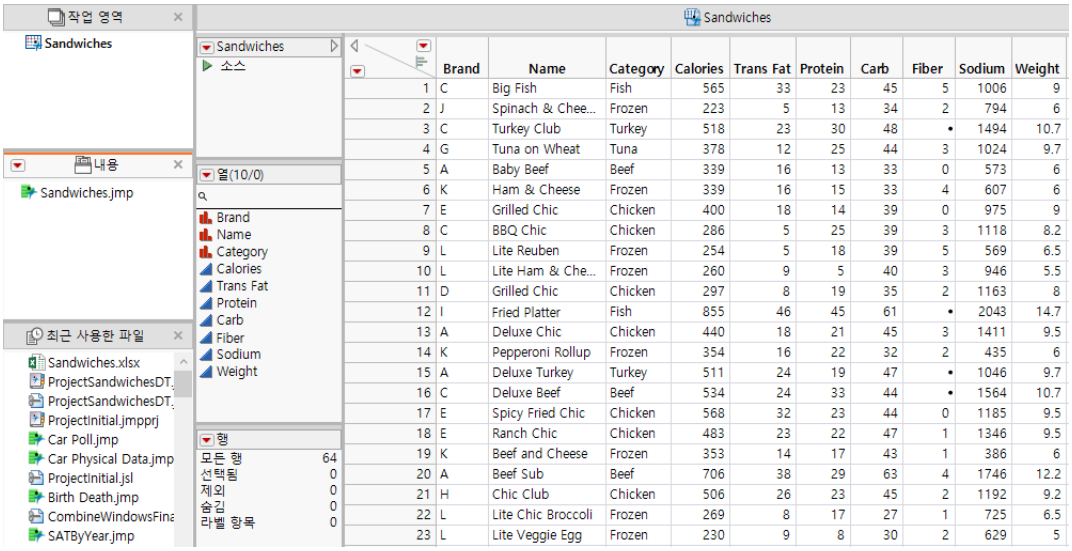
1. **파일** > **새로 만들기** > **프로젝트** (Windows) 또는 **파일** > **새로 만들기** > **새 프로젝트** (macOS) 를 선택합니다.
2. 프로젝트 창에서 **파일** > **열기**를 선택합니다.
3. sandwiches.xlsx 파일을 엽니다. 이 파일의 기본 위치는 다음과 같습니다.
   C:/Program Files/SAS/JMP/16/Samples/Import Data

   팁 : 하단에서 **모든 JMP 파일**을 **Excel 파일**로 변경해야 할 수도 있습니다.

4. 가져오기를 클릭합니다.

5. **파일** > **저장**을 선택합니다 .
6. **프로젝트 내용**이 선택되어 있는지 확인합니다 .
7. 파일 이름을 Sandwiches.jmp 로 변경하고 **저장**을 클릭합니다 .

그림 7.8 Sandwiches 데이터를 가져온 프로젝트

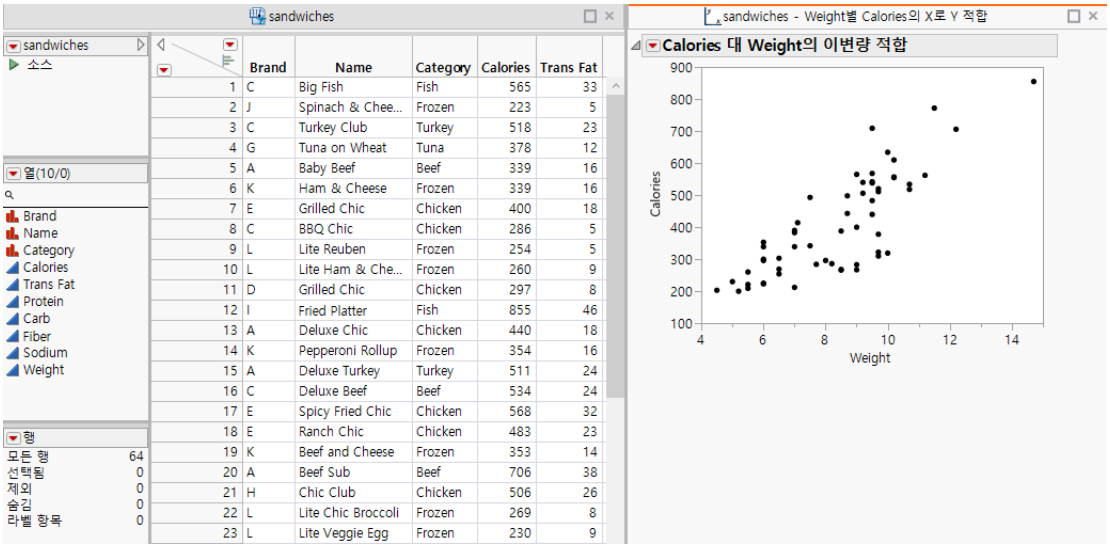


	Brand	Name	Category	Calories	Trans Fat	Protein	Carb	Fiber	Sodium	Weight
1	C	Big Fish	Fish	565	33	23	45	5	1006	9
2	J	Spinach & Chee...	Frozen	223	5	13	34	2	794	6
3	C	Turkey Club	Turkey	518	23	30	48	•	1494	10.7
4	G	Tuna on Wheat	Tuna	378	12	25	44	3	1024	9.7
5	A	Baby Beef	Beef	339	16	13	33	0	573	6
6	K	Ham & Cheese	Frozen	339	16	15	33	4	607	6
7	E	Grilled Chic	Chicken	400	18	14	39	0	975	9
8	C	BBQ Chic	Chicken	286	5	25	39	3	1118	8.2
9	L	Lite Reuben	Frozen	254	5	18	39	5	569	6.5
10	L	Lite Ham & Che...	Frozen	260	9	5	40	3	946	5.5
11	D	Grilled Chic	Chicken	297	8	19	35	2	1163	8
12	I	Fried Platter	Fish	855	46	45	61	•	2043	14.7
13	A	Deluxe Chic	Chicken	440	18	21	45	3	1411	9.5
14	K	Pepperoni Rollup	Frozen	354	16	22	32	2	435	6
15	A	Deluxe Turkey	Turkey	511	24	19	47	•	1046	9.7
16	C	Deluxe Beef	Beef	534	24	33	44	•	1564	10.7
17	E	Spicy Fried Chic	Chicken	568	32	23	44	0	1185	9.5
18	E	Ranch Chic	Chicken	483	23	22	47	1	1346	9.5
19	K	Beef and Cheese	Frozen	353	14	17	43	1	386	6
20	A	Beef Sub	Beef	706	38	29	63	4	1746	12.2
21	H	Chic Club	Chicken	506	26	23	45	2	1192	9.2
22	L	Lite Chic Broccoli	Frozen	269	8	17	27	1	725	6.5
23	L	Lite Veggie Egg	Frozen	230	9	8	30	2	629	5

8. 분석 > X 로 Y 적합을 선택합니다 .
9. Calories 를 선택하고 **Y, 반응**을 클릭합니다 .
10. Weight 를 선택하고 **X, 요인**을 클릭합니다 .
11. 확인을 클릭합니다 .
12. **Sandwiches - X 로 Y 적합** 탭을 오른쪽으로 드래그하여 오른쪽 도킹 영역에 놓습니다 .

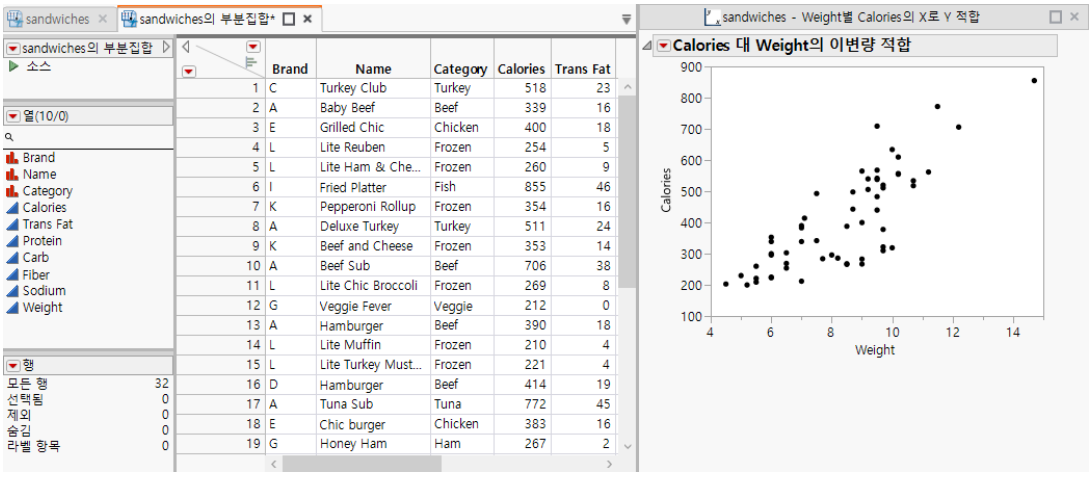
팁 : 전체 보고서를 표시하려면 데이터 테이블과 보고서 사이의 선을 왼쪽으로 드래그합니다 .

그림 7.9 오른쪽에 도킹된 X 로 Y 적합 (이변량) 보고서



13. **테이블** > **부분집합**을 선택합니다 .
14. " 행 " 아래에서 **랜덤 - 표집 비율 : 0.5** 를 선택합니다 .
15. 확인을 클릭합니다 .

그림 7.10 저장되지 않은 부분집합 테이블이 있는 프로젝트



16. **파일** > **프로젝트 저장**을 선택합니다 .
17. 프로젝트를 저장할 폴더로 이동하여 프로젝트 파일 이름을 지정하고 **저장**을 클릭합니다 .
18. 프로젝트를 닫습니다 .

19. **파일** > **열기**를 선택하고 프로젝트 파일을 엽니다 .

저장하지 않은 부분집합 테이블이 프로젝트 파일에 자동으로 저장되어 있습니다.

플랫폼 결과를 저널로 저장

보고서 창의 저널을 생성하여 플랫폼 보고서를 나중에 볼 수 있도록 저장합니다. 저널은 보고서 창의 사본입니다. 기존 저널을 편집하거나 기존 저널에 보고서를 추가할 수 있습니다. 저널은 데이터 테이블에 연결되지 않습니다. 저널은 여러 보고서 창의 결과를 단일 보고서 창에 저장하여 다른 사람들과 간편하게 공유할 수 있는 기능입니다.

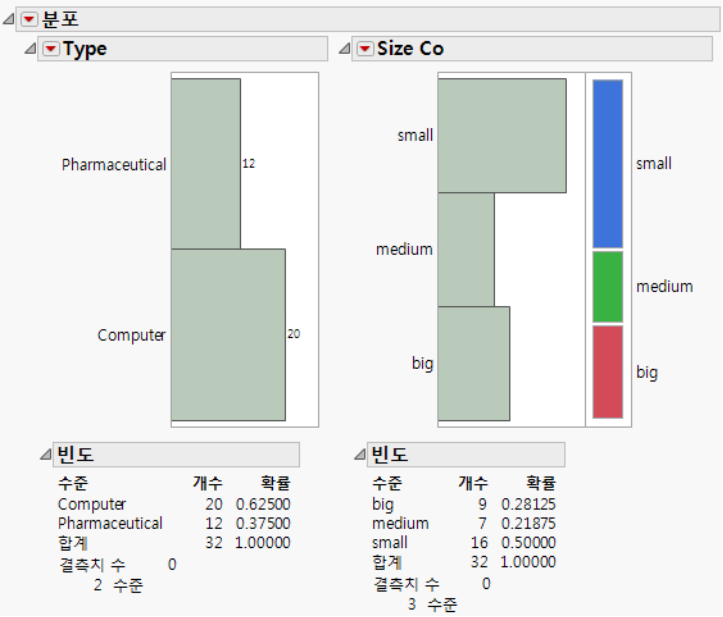
이 섹션에는 다음과 같은 정보가 포함되어 있습니다 .

- " 예 : 저널 생성 "
- " 저널에 분석 추가 "

예 : 저널 생성

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다 .
2. **분석** > **분포**를 선택합니다 .
3. Type 및 Size Co 를 모두 선택하고 **Y, 열**을 클릭합니다 .
4. **확인**을 클릭합니다 .
5. "Type" 의 빨간색 삼각형을 클릭하고 **히스토그램 옵션** > **개수 표시**를 선택합니다 .
6. "Size Co" 의 빨간색 삼각형을 클릭하고 **모자이크 그림**을 선택합니다 .
7. **편집** > **저널**을 선택하여 이 결과에 대한 저널을 생성합니다 . 결과가 저널 창에 복제됩니다 .

그림 7.11 분포 결과의 저널



저널의 결과는 데이터 테이블에 연결되지 않습니다. "Type" 막대 차트에서 "Computer" 막대를 클릭해도 데이터 테이블에서 행이 선택되지 않습니다.

저널은 결과의 사본이기 때문에 빨간색 삼각형 메뉴가 대부분 존재하지 않습니다. 그러나 저널에 새로 추가한 보고서에는 빨간색 삼각형 메뉴가 있습니다. 이 메뉴에는 두 가지 옵션이 있습니다.

새 창에서 다시 실행 원래 보고서를 생성하는 데 사용된 원래 데이터 테이블이 있을 때 이 옵션은 분석을 다시 실행합니다. 그 결과로 새 보고서 창이 표시됩니다.

스크립트 편집 이 옵션은 분석을 다시 생성하는 JSL 스크립트가 포함된 스크립트 창을 엽니다. JSL 에 대해서는 좀 더 자세한 설명이 필요하므로 Scripting Guide 및 JSL Syntax Reference 에서 별도로 다룹니다.

저널에 분석 추가

다른 분석을 수행하는 경우 분석 결과를 기존 저널에 추가할 수 있습니다.

1. 저널이 열린 상태에서 **분석** > **분포**를 선택합니다.
2. profit/emp 를 선택하고 **Y, 열**을 클릭합니다.
3. **확인**을 클릭합니다.
4. **편집** > **저널**을 선택합니다. 결과가 저널 하단에 추가됩니다.

스크립트 저장 및 실행

JMP 의 플랫폼 옵션은 대부분 스크립트 가능합니다 . 즉 , 수행하는 대부분의 작업을 JSL(JMP 스크립트 언어) 스크립트로 저장할 수 있습니다 . 스크립트를 사용하여 언제든지 작업 또는 결과를 재현할 수 있습니다 .

이 섹션에는 다음과 같은 정보가 포함되어 있습니다.

- " 예 : 스크립트 저장 및 실행 "
- " 스크립트 및 JSL"

예 : 스크립트 저장 및 실행

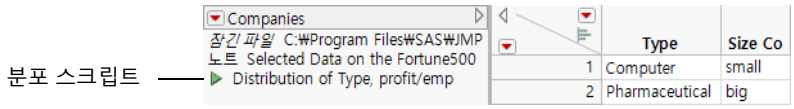
보고서 생성

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다 .
2. **분석** > **분포**를 선택합니다 .
3. Type 및 profit/emp 를 선택하고 **Y, 열**을 클릭합니다 .
4. **확인**을 클릭합니다 .
5. "Type" 의 빨간색 삼각형을 클릭하고 다음 옵션을 선택합니다 .
   - **히스토그램 옵션** > **개수 표시**
   - **신뢰 구간** > **0.95**
6. "profit/emp" 의 빨간색 삼각형을 클릭하고 다음 옵션을 선택합니다 .
   - **이상치 상자 그림** (이상치 상자 그림 제거)
   - **CDF 그림**
7. " 분포 " 의 빨간색 삼각형을 클릭하고 **쌓기**를 선택합니다 .

스크립트를 데이터 테이블에 저장하고 실행

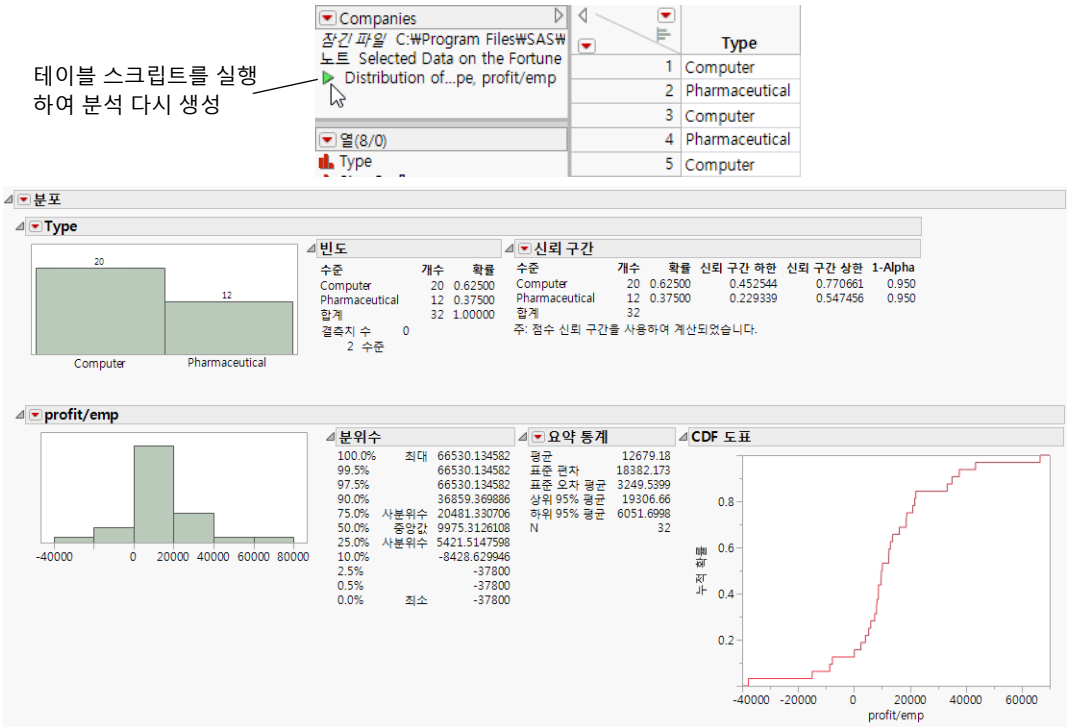
1. 이 분석을 저장하려면 " 분포 " 의 빨간색 삼각형을 클릭하고 **스크립트 저장 > 데이터 테이블에**를 선택합니다 . 테이블 패널에 새 스크립트가 나타납니다 .

그림 7.12 분포 스크립트



2. " 분포 " 보고서 창을 닫습니다 .
3. 분석을 다시 생성하려면 " 분포 " 스크립트 옆의 녹색 삼각형을 클릭합니다 .

그림 7.13 분포 스크립트 실행



팁 : 테이블 스크립트를 마우스 오른쪽 버튼으로 클릭하면 추가 옵션이 표시됩니다 .

스크립트 및 JSL

이 섹션에서 저장한 스크립트에는 JSL(JMP 스크립트 언어) 명령이 포함되어 있습니다. JSL에 대해서는 좀 더 자세한 설명이 필요하므로 Scripting Guide 및 JSL Syntax Reference 에서 별도로 다룹니다.

보고서를 대화식 HTML 로 저장

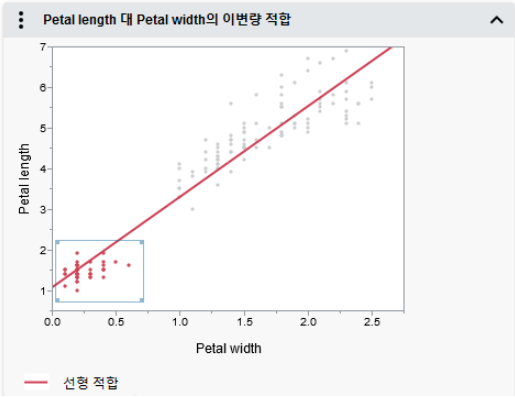
JMP 사용자가 대화식 HTML 을 통해 동적 그래프가 포함된 보고서를 공유하면 JMP 를 사용하지 않는 사람도 데이터를 탐색할 수 있습니다. JMP 보고서는 사용자에게 전자 우편으로 보내거나 웹 사이트에 게시할 수 있는 HTML 5 형식의 웹 페이지로 저장됩니다. 그러면 사용자들이 JMP 에서처럼 데이터를 탐색할 수 있습니다.

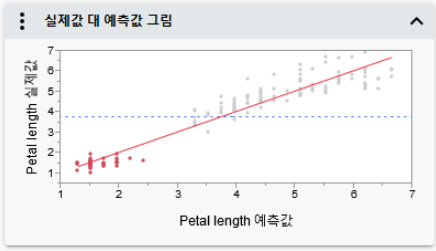
대화식 HTML 은 다음과 같은 JMP 의 일부 기능을 제공합니다 .

- 히스토그램 막대를 선택하고 데이터 값을 보는 등 대화식 그래프 기능을 탐색합니다 .

- 브러시 방식으로 데이터를 확인합니다.
- 보고서 섹션을 표시하거나 숨깁니다.
- 보고서를 커서로 가리켜 툴팁을 봅니다.
- 표식 크기를 늘립니다.

그림 7.14 대화식 HTML 에서 브러시 방식으로 데이터 확인





정렬된 변수, 가로 히스토그램, 배경 색상 및 색상이 적용된 데이터 점 같은 그래프의 많은 변경 사항이 웹 페이지에 저장됩니다. 내용을 저장할 때 닫힌 그래프와 테이블은 사용자가 열 때까지 웹 페이지에서 닫힌 채로 유지됩니다.

대화식 HTML 의 데이터 포함

JMP 에서 보고서를 대화식 HTML 로 저장하면 데이터가 HTML 에 포함됩니다. 웹 브라우저는 암호화된 데이터를 읽을 수 없기 때문에 내용이 암호화되지 않습니다. 중요한 데이터를 공유하지 않으려면 결과를 대화식이 아닌 웹 페이지로 저장하십시오. **파일 > 내보내기 > 데이터가 포함된 대화식 HTML** 을 선택하십시오.

예 : 대화식 HTML 생성

보고서 생성

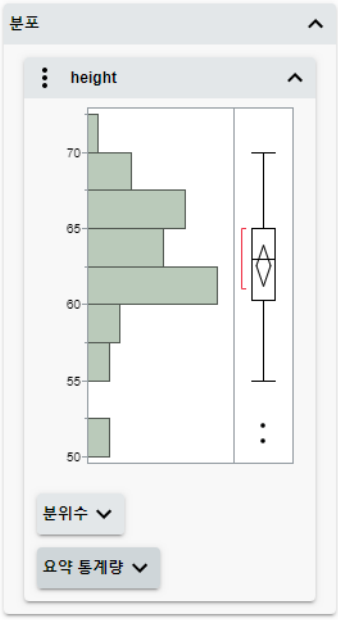
1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Big Class.jmp 를 엽니다 .
2. **분석 > 분포**를 선택합니다 .
3. height 를 선택하고 **Y, 열**을 클릭합니다 .
4. **확인**을 클릭합니다 .

대화식 HTML 로 저장

1. (Windows) **파일 > 내보내기**를 선택하고 **데이터가 포함된 대화식 HTML** 을 선택한 후 **확인**을 클릭합니다 .
2. (macOS) **파일 > 내보내기**를 선택하고 **데이터가 포함된 대화식 HTML** 을 선택한 후 **다음**을 클릭합니다 .
3. " 내보내기 " 창에서 **저장 후 파일 열기**가 선택되어 있지 않으면 이 항목을 선택합니다 .
4. 파일 이름을 지정하고 파일을 저장합니다 .

   출력이 기본 브라우저에 나타납니다 .

그림 7.15 대화식 HTML 출력



대화식 HTML 출력을 탐색하는 방법에 대한 자세한 내용은 https://www.jmp.com/interactive 에서 확인하십시오.

보고서를 대화식 HTML 로 공유

대화식 HTML 을 통해 동적 그래프가 포함된 JMP 보고서를 공유하면 JMP 를 사용하지 않는 사람도 데이터를 탐색할 수 있습니다. JMP 보고서는 공유 네트워크 드라이브나 웹 사이트에서 또는 이메일을 통해 다른 사람과 공유할 수 있는 대화식 웹 페이지로 저장됩니다. 그러면 사용자들이 JMP 에서처럼 데이터를 탐색할 수 있습니다.

대화식 HTML 의 데이터 포함

보고서를 대화식 HTML 로 내보내거나 게시하면 데이터가 HTML 에 포함됩니다. 웹 브라우저는 암호화된 데이터를 읽을 수 없기 때문에 내용이 암호화되지 않습니다. 중요한 데이터를 공유하지 않으려면 **파일 > 내보내기 > HTML** 을 선택하여 결과를 대화식이 아닌 웹 페이지로 내보내십시오.

대화식으로 지원되는 기능

대화식 HTML 은 다음과 같은 JMP 의 일부 기능을 제공합니다.

- 보고서의 기능이 완전히 지원되면 경고 없이 웹 페이지가 생성됩니다 .
- 보고서에 지원되지 않는 기능이 포함되어 있으면 내보내기 또는 게시 창에 대화식 HTML 이 부분적으로 구현되었다는 메시지가 표시됩니다 . 자세한 내용은 JMP 로그 (**보기** > **로그**) 에서 확인하십시오 .
- 부분적으로 지원되거나 지원되지 않는 기능은 웹 페이지에 정적 상태로 나타납니다. 웹 페이지에서 지원되지 않는 기능을 커서로 가리키면 해당 기능이 아직 대화식이 아니라는 툴팁이 표시됩니다 .

대화식 HTML 출력을 탐색하는 방법에 대한 자세한 내용은 https://www.jmp.com/interactive 에서 확인하십시오.

단일 보고서를 대화식 HTML 로 내보내기

단일 JMP 보고서를 대화식 HTML 로 내보내려면 " 데이터가 포함된 대화식 HTML 로 내보내기 " 옵션을 사용하여 단일 웹 페이지를 생성합니다.

1. JMP 에서 보고서를 생성하고 해당 창을 활성화합니다 .

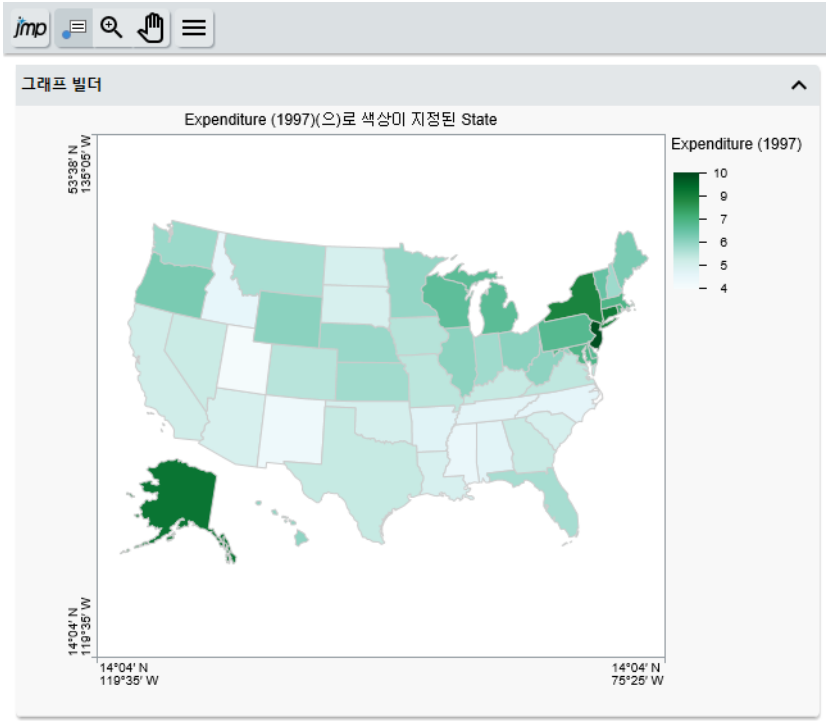
   참고 : 보고서에 로컬 데이터 필터가 포함되어 있으면 웹 보고서에서 정적 상태이며 사용자가 선택을 변경할 수 없습니다 . 로컬 데이터 필터를 대화식으로 만들려면 **포함** 모드를 선택 취소하십시오 . 그래프 빌더에서 **포함** 모드를 표시하려면 " 로컬 데이터 필터 " 의 빨간색 삼각형 메뉴에서 **모드 표시**를 선택합니다 .

2. 파일 > 내보내기를 선택하고 데이터가 포함된 대화식 HTML 을 선택한 후 다음을 클릭합니다 .
3. 파일 이름을 지정합니다 .
4. (선택 사항) HTML 파일을 내보낸 후 기본 브라우저에서 열려면 **저장 후 파일 열기**를 선택합니다 .

5. 저장을 클릭합니다.

   출력이 선택한 폴더에 저장됩니다.

그림 7.16 단일 대화식 HTML 보고서에 대한 웹 페이지



여러 보고서를 대화식 HTML 로 게시

여러 JMP 보고서를 대화식 HTML 로 게시하려면 "파일에 게시" 옵션을 사용하여 보고서가 포함된 인덱스 페이지를 생성합니다.

팁 : 여러 JMP 보고서를 게시하는 경우 인덱스 페이지와 보고서 파일을 저장할 위치를 지정해야 합니다. 공유 네트워크 폴더를 선택하고 다른 사람에게 위치를 알려주거나, 공유하기 전에 사용자 컴퓨터의 폴더를 선택하고 파일을 압축할 수 있습니다. 이 방법은 JMP 를 사용하지 않는 사람과 공유할 때 특히 유용합니다.

1. JMP 에서 보고서를 생성합니다.

   참고 : 보고서에 로컬 데이터 필터가 포함되어 있으면 웹 보고서에서 정적 상태이며 사용자가 선택을 변경할 수 없습니다. 로컬 데이터 필터를 대화식으로 만들려면 **포함** 모드를 선택 취소하십시오. 그래프 빌더에서 **포함** 모드를 표시하려면 " 로컬 데이터 필터 " 의 빨간색 삼각형 메뉴에서 **모드 표시**를 선택합니다.

2. 보고서 창에서 **파일** > **게시** > **파일에 게시**를 선택합니다 .
3. 게시할 보고서를 선택합니다 .
4. (선택 사항) 상위 폴더 위치를 변경하거나 보고서를 포함할 하위 폴더의 이름을 변경합니다 .
5. 다음을 클릭합니다 .
6. 인덱스 페이지의 제목을 입력합니다 . 보고서 제목을 업데이트할 수도 있습니다 .
7. (선택 사항) 추가 옵션을 변경합니다 . 자세한 내용은 " 대화식 HTML 보고서 옵션 " 에서 확인하십시오 .
8. 게시를 클릭합니다 .

그림 7.17 여러 대화식 HTML 보고서에 대한 인덱스 페이지

9. 썸네일을 클릭하여 보고서를 엽니다 .

대화식 보고서 작업에 대한 자세한 내용을 보려면 HTML 보고서에서 ☰ > **도움말**을 클릭합니다. 그러면 도움말 페이지 (https://www.jmp.com/interactive) 가 열립니다.

대화식 HTML 보고서 옵션

제목 인덱스 페이지 (보고서가 여러 개인 경우만 해당) 또는 보고서의 제목을 추가합니다 .

설명 (선택 사항) 인덱스 페이지 또는 보고서에 설명을 추가합니다 . 이 설명은 제목 아래에 표시됩니다 .

사용자 정의 (보고서가 여러 개인 경우에만 표시) 웹 페이지의 모양을 변경합니다 . 스타일 , 테마 , 글꼴 , 로고 및 날짜 / 시간 표시 여부를 변경할 수 있습니다 . 자세한 내용은 " 인덱스 페이지 사용자 정의 옵션 " 에서 확인하십시오 .

데이터 게시 대화식 보고서의 경우 이 옵션을 선택합니다. 이 옵션을 선택 취소하면 정적 보고서가 됩니다 .

참고 : 중요한 데이터를 공유하지 않으려면 결과를 대화식이 아닌 웹 페이지로 저장하십시오 . **파일 > 내보내기 > HTML** 을 선택합니다 .

이미지 추가 웹 페이지 아래쪽에 이미지를 추가합니다 .

게시된 웹 보고서 열기 " 게시 " 를 클릭하면 브라우저에서 웹 페이지를 엽니다 .

실행 후 보고서 닫기 웹 보고서를 게시하면 JMP 에서 보고서를 닫습니다 .

삭제 아이콘 (보고서가 여러 개인 경우에만 표시) 보고서를 삭제합니다 .

화살표 아이콘 (보고서가 여러 개인 경우에만 표시) 보고서 순서를 변경합니다 .

인덱스 페이지 사용자 정의 옵션

스타일 형식 보고서의 레이아웃을 결정합니다 .

큰 목록 보고서를 큰 썸네일 그래픽과 함께 열 하나에 표시합니다 .

작은 목록 보고서를 작은 썸네일 그래픽과 함께 열 하나에 표시합니다 .

격자 보고서를 여러 행에 표시합니다 .

사용자 CSS 웹 페이지의 형식을 지정하기 위한 CSS 파일을 지정할 수 있습니다 . CSS 파일은 _css 라는 하위 폴더에 복사됩니다 . 보고서와 지원 파일을 다른 사용자에게 전송하면서 CSS 형식 지정을 유지할 수 있도록 CSS 파일에 대한 링크는 상대적입니다 .

색상 테마 웹 페이지 , 머리글 및 테두리의 색상을 지정합니다 . 기본 웹 페이지의 배경은 흰색 , 머리글은 주황색 , 테두리는 파란색입니다 .

글꼴 변경 보고서에 적용된 글꼴을 변경합니다 .

로고 변경 보고서와 함께 표시할 이미지를 지정합니다. 이미지를 보고서의 위 또는 아래로 이동하려면 이미지 옆의 위쪽 또는 아래쪽 화살표를 클릭합니다 .

날짜 / 시간 표시 웹 페이지가 생성된 날짜 및 시간을 표시하거나 숨깁니다 .

보고서를 PowerPoint 프레젠테이션으로 저장

JMP 결과를 PowerPoint 프레젠테이션 (.pptx) 으로 저장하여 프레젠테이션을 생성할 수 있습니다 . .pptx 파일로 저장한 후에는 PowerPoint 에서 JMP 내용을 재배열하고 텍스트를 편집합니다 . JMP 보고서를 PowerPoint 로 내보낼 경우 다음과 같이 섹션이 달라집니다 .

- 보고서 머리글은 편집 가능한 텍스트 상자로 내보내집니다 .
- 그래프는 이미지로 내보내집니다 . 범례와 같은 특정 그래픽 요소는 별도의 이미지로 내보내집니다 . PowerPoint 의 슬라이드에 맞게 이미지의 크기가 조정됩니다 .

프레젠테이션에 저장할 섹션을 선택하려면 선택 도구를 사용합니다 . PowerPoint 에서 파일을 연 후에는 원치 않는 콘텐츠를 삭제하십시오 .

참고 : Windows 에서 JMP 를 통해 생성한 .pptx 파일을 열려면 PowerPoint 2007 이상이 필요합니다 . macOS 에서는 PowerPoint 2011 이상이 필요합니다 .

1. JMP 에서 보고서를 생성합니다 .
2. **파일 > 내보내기**를 선택하고 **Microsoft PowerPoint** 를 선택한 후 **다음**을 클릭합니다 .
3. 목록에서 그래픽 파일 형식을 선택합니다 .

   Windows 에서는 EMF 가 기본 형식입니다 . macOS 에서는 PDF 가 기본 형식입니다 .
4. 파일 이름을 지정하고 파일을 저장합니다 . macOS 에서는 파일 이름을 지정하고 **내보내기**를 클릭합니다 .

   저장 후 파일 열기가 기본적으로 선택되어 있으므로 파일이 Microsoft PowerPoint 에서 열립니다 .

참고 : Windows 에서 생성된 기본 EMF 그래픽은 macOS 에서 지원되지 않습니다 . macOS 에서 생성된 기본 PDF 그래픽은 Windows 에서 지원되지 않습니다 . 플랫폼 간 호환성을 위해서는 **파일 > 환경 설정 > 일반**을 선택하여 기본 그래픽 파일 형식을 변경하십시오 . 그런 다음 **PowerPoint 용 이미지 형식**을 PNG 또는 JPEG 로 변경하십시오 .

대시보드 생성

대시보드는 정기적으로 보고서를 실행하고 표시할 수 있는 시각적 도구입니다. 보고서, 데이터 필터, 선택 필터, 데이터 테이블 및 그래픽을 대시보드에 표시할 수 있습니다. 대시보드를 열면 대시보드에 표시된 내용이 업데이트됩니다.

이 섹션에는 다음과 같은 정보가 포함되어 있습니다.

- " 예 : 창 결합 "
- " 예 : 두 개의 보고서가 포함된 대시보드 생성 "

예 : 창 결합

JMP 에서 열려 있는 여러 창을 병합하여 신속하게 대시보드를 생성할 수 있습니다 . 창을 결합하면 통계량 요약을 볼 수 있는 옵션이 제공되고 선택 필터가 포함됩니다 .

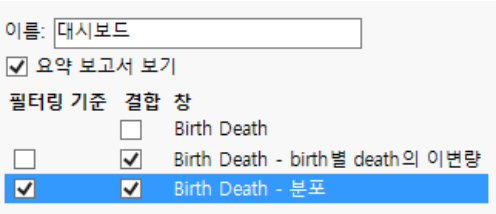
1. 도움말 > 샘플 데이터 라이브러리를 선택하고 Birth Death.jmp 를 엽니다 .
2. "Distribution" 및 "Bivariate" 테이블 스크립트를 실행합니다 .
3. 보고서 창 중 하나에서 **창 > 창 결합**을 선택합니다 .

   " 창 결합 " 창이 나타납니다 .

   팁 : Windows 에서는 JMP 창의 오른쪽 하단에 있는 " 배열 메뉴 " 옵션에서 " 창 결합 " 을 선택할 수도 있습니다 .

4. **요약 보고서 보기**를 선택하여 그래프를 표시하고 통계 보고서를 생략합니다 .
5. " 결합 " 열에서 **Birth Death - birth 별 death 의 이변량** 및 **Birth Death - 분포**를 선택합니다 .
6. " 필터링 기준 " 열에서 **Birth Death - 분포**를 선택합니다 .

그림 7.18 창 결합 옵션



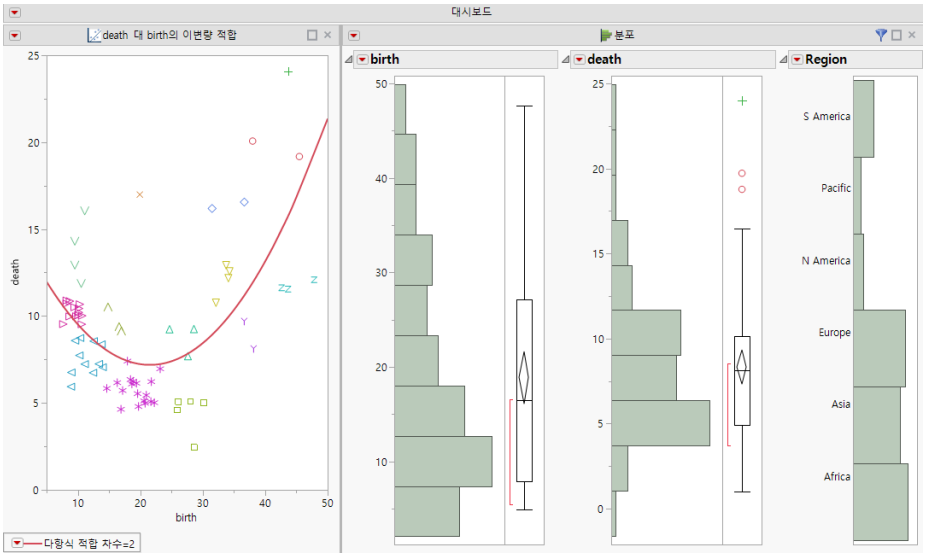
7. 확인을 클릭합니다 .

   두 개의 보고서가 하나의 창으로 결합됩니다 . " 분포 " 보고서 상단의 필터 아이콘 을 확인합니다 . 히스토그램 중 하나에서 막대를 선택하면 이변량 그래프의 해당 데이터가 선택됩니다 .

참고 :

- Windows 에서 보고서를 결합하려면 JMP 창의 오른쪽 하단에 있는 " 배열 메뉴 " 옵션에서 창 결합을 선택해도 됩니다 .
- 보고서에 그래프만 표시하고 통계량은 생략하려면 " 창 결합 " 창에서 **요약 보고서 보기**를 선택하십시오 .

그림 7.19 결합된 창



예 : 두 개의 보고서가 포함된 대시보드 생성

두 개의 보고서를 생성했고 다음날 업데이트된 데이터 집합에 대해 보고서를 다시 실행하려고 한다고 가정해 보겠습니다 . 이 예에서는 대시보드 빌더의 보고서에서 대시보드를 생성하는 방법을 보여 줍니다 .

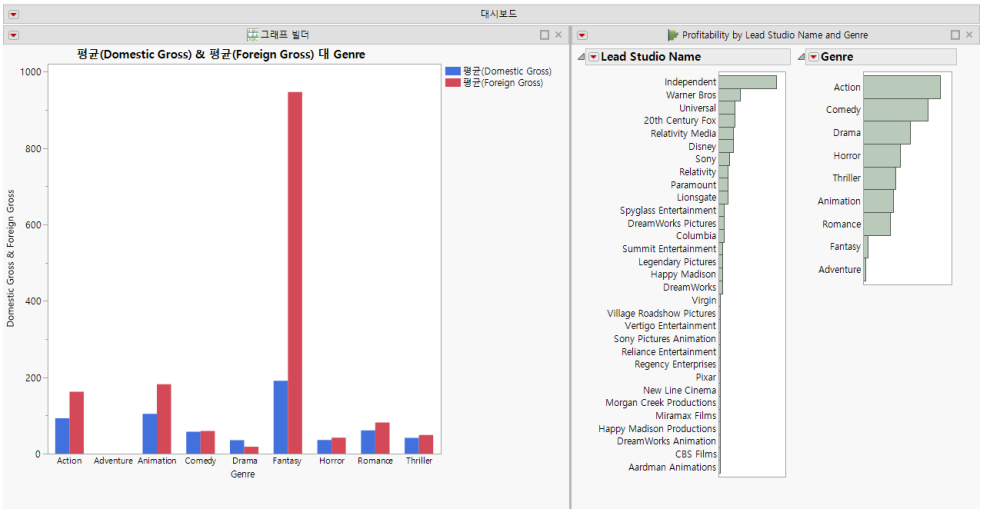
1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Hollywood Movies.jmp 를 엽니다 .
2. "Distribution: Profitability by Lead Studio and Genre" 및 "Graph Builder: World and Domestic Gross by Genre" 라는 테이블 스크립트를 실행합니다 .
3. 어느 창에서든 **파일 > 새로 만들기 > 대시보드**를 선택합니다 .

   일반적인 레이아웃을 위한 템플릿이 나타납니다 .
4. **2x1 대시보드** 템플릿을 선택합니다 .

   두 개의 보고서를 위한 공간이 있는 상자가 작업 공간에 나타납니다 .
5. " 보고서 " 목록에서 보고서 썸네일을 두 번 클릭하여 대시보드에 배치합니다 .
6. " 대시보드 빌더 " 의 빨간색 삼각형을 클릭하고 **미리보기 모드**를 선택합니다 .

   대시보드의 미리보기가 나타납니다 . 그래프가 서로 연결되고 데이터 테이블과 연결되어 있습니다 . 또한 " 분포 " 및 " 그래프 빌더 " 플랫폼과 동일한 빨간색 삼각형 옵션이 있습니다 .
7. **미리보기 닫기**를 클릭합니다 .

그림 7.20 두 개의 보고서가 있는 대시보드



대시보드 생성에 대한 자세한 내용은 JMP 사용의 에서 확인하십시오.

8 장

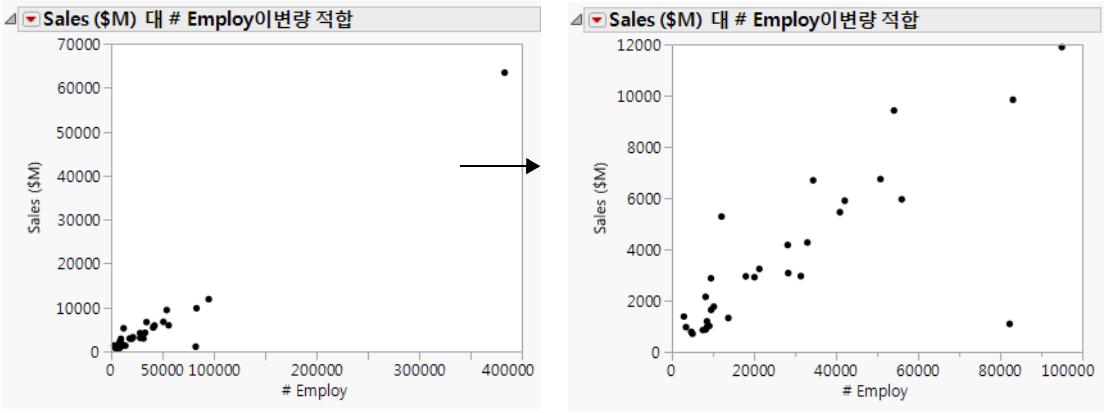
특수 기능

자동 분석 업데이트 및 SAS Integration

JMP 의 일부 특수 기능을 사용하여 다음을 수행할 수 있습니다.

- 자동으로 분석 또는 그래프 업데이트
- 플랫폼 결과 사용자 정의
- SAS 와 통합하여 고급 분석 기능 사용

그림 8.1 특수 기능의 예




```
DATA Candy_Bars; INPUT  Calories Total_fat_g Carbohydrate_g Protein_g; Lines;
310 20 28 6
230 12 27 4
220 12 24 3
170 8 21 3
200 2.5 43 1
260 16 25 5
190 1.5 42 2
190 11 21 2
230 12 28 3
;
RUN;

PROC GLM  DATA=Candy_Bars ALPHA=0.05;
MODEL  Calories =  Total_fat_g Carbohydrate_g Protein_g;
RUN;
```

목차

분석 및 그래프 자동 업데이트 205
예: 자동 재계산 사용 205
환경 설정 변경 209
예: 환경 설정 변경 209
JMP 와 SAS 통합 212
예: SAS 코드 생성 212
예: SAS 코드 전송 213

분석 및 그래프 자동 업데이트

데이터 테이블을 변경할 때 자동 재계산 기능을 사용하여 데이터 테이블과 연결된 분석 및 그래프를 자동으로 업데이트할 수 있습니다. 예를 들어 데이터 테이블에서 값을 제외, 포함 또는 삭제하면 해당 변경 사항이 연결된 분석 또는 그래프에 즉시 반영됩니다. 다음 사항에 유의하십시오.

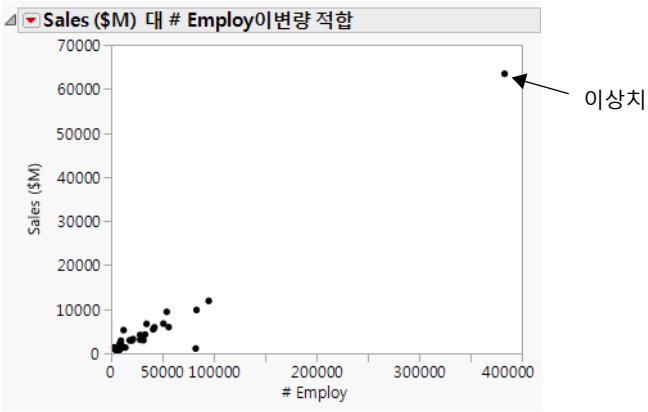
- 일부 플랫폼은 자동 재계산 기능을 지원하지 않습니다. 자세한 내용은 JMP 사용의 에서 확인하십시오.
- **분석** 메뉴의 지원되는 플랫폼에서는 자동 재계산 기능이 기본적으로 해제되어 있습니다. 그러나 **품질 및 공정** 메뉴에서 지원되는 플랫폼의 경우에는 계량형 / 계수형 게이지 차트, 공정 능력 및 관리도를 제외하고는 자동 재계산 기능이 기본적으로 설정되어 있습니다.
- **그래프** 메뉴에서 지원되는 플랫폼의 경우에는 자동 재계산 기능이 기본적으로 설정되어 있습니다.

예 : 자동 재계산 사용

이 예에서는 제약 및 컴퓨터 업계의 32 개 회사에 대한 재무 데이터가 포함된 Companies.jmp 샘플 데이터 테이블을 사용합니다.

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다.
2. **분석** > **X 로 Y 적합**을 선택합니다.
3. Sales ($M) 를 선택하고 **Y, 반응**을 클릭합니다.
4. # Employ 를 선택하고 **X, 요인**을 클릭합니다.
5. **확인**을 클릭합니다.

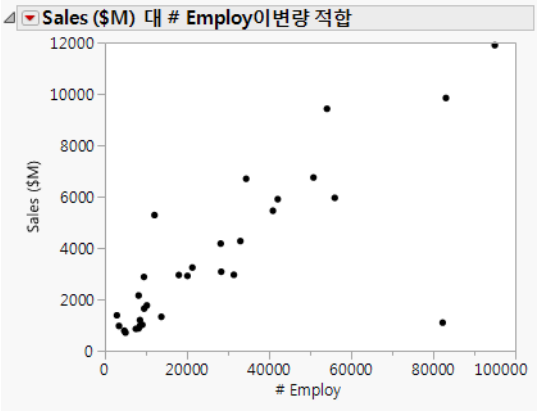
그림 8.2 초기 산점도



초기 산점도에서는 한 회사의 직원 수와 매출이 다른 모든 회사보다 훨씬 많음을 보여 줍니다. 이 회사가 이상치라고 판단하고 해당 데이터 점을 제외하고 싶습니다. 변경 시 산점도가 자동으로 업데이트되도록 하려면 해당 점을 제외하기 전에 자동 재계산 기능을 설정해야 합니다.

6. 자동 재계산 기능을 설정하려면 "Sales ($M) 대 # Employ 의 이변량 적합 " 옆의 빨간색 삼각형을 클릭하고 **다시 실행** > **자동 재계산**을 선택합니다 .
7. 이상치를 클릭하여 선택합니다 .
8. **행** > **제외 / 제외 해제**를 선택합니다 . 해당 점이 분석에서 제외되고 산점도가 자동으로 업데이트됩니다 .

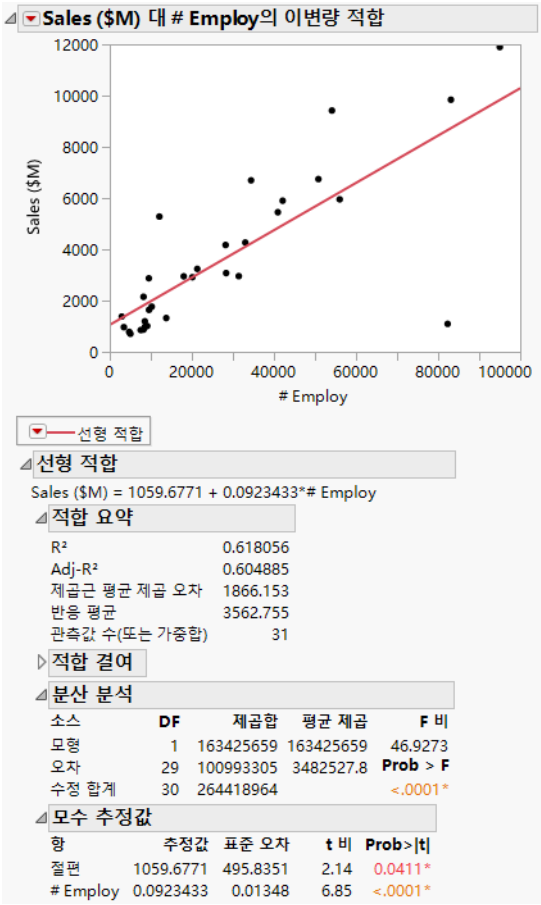
그림 8.3 업데이트된 산점도



데이터에 회귀선을 적합시킬 때 오른쪽 하단의 점은 이상치이며 선의 기울기에 영향을 줍니다 . 자동 재계산 기능이 설정된 상태에서 이상치를 제외하면 선의 기울기가 변경되는 것을 볼 수 있습니다 .

9. 회귀선을 적합시키려면 "Sales ($M) 대 # Employ 의 이변량 적합 " 옆의 빨간색 삼각형을 클릭하고 **선형 적합**을 선택합니다 . 그림 8.4 에서는 보고서 창에 추가된 회귀선과 분석 결과를 보여 줍니다 .

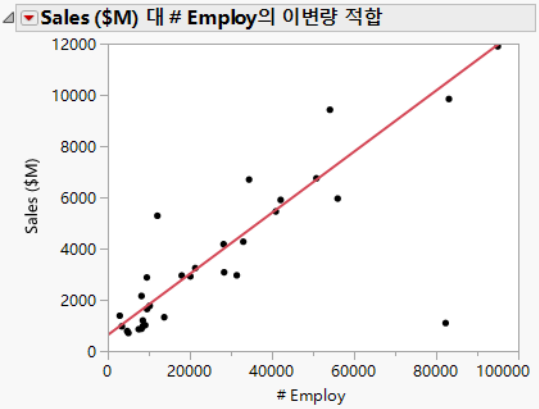
그림 8.4 회귀선 및 분석 결과



10. 이상치를 클릭하여 선택합니다 .
11. **행** > **제외 / 제외 해제**를 선택합니다 . 해당 점이 제외되었음을 반영하여 회귀선과 분석 결과가 자동으로 업데이트됩니다 .

팁 : 점을 제외하면 분석은 해당 데이터 점 없이 다시 계산되지만 해당 데이터 점이 산점도에서 숨겨지지는 않습니다 . 산점도에서도 해당 점을 숨기려면 점을 선택한 후 **행** > **숨기고 제외하기**를 선택하십시오 .

그림 8.5 업데이트된 회귀선 및 분석 결과



선형 적합

선형 적합

Sales ($M) = 631.36791 + 0.1192722*# Employ

적합 요약

R^2	0.882703
Adj-R^2	0.878514
제곱근 평균 제곱 오차	1039.891
반응 평균	3644.95
관측값 수(또는 가중합)	30

적합 결여

분산 분석

소스	DF	제곱합	평균 제곱	F 비
모형	1	227857403	227857403	210.7112
오차	28	30278440	1081372.9	**Prob > F**
수정 합계	29	258135843		<.0001*

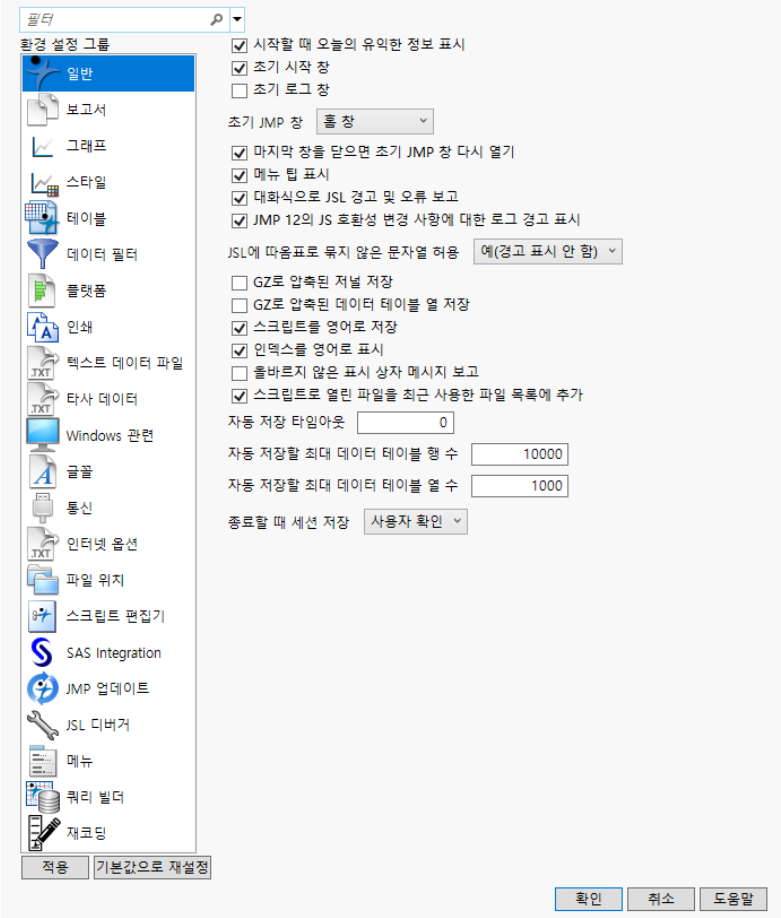
모수 추정값

항	추정값	표준 오차	t 비	Prob>|t|
절편	631.36791	281.3288	2.24	0.0329*
# Employ	0.1192722	0.008217	14.52	<.0001*

환경 설정 변경

" 환경 설정 " 창을 사용하여 JMP 의 환경 설정을 변경할 수 있습니다 . " 환경 설정 " 창을 열려면 **파일** > **환경 설정** (Windows) 또는 **JMP** > **환경 설정** (macOS) 을 선택합니다 .

그림 8.6 환경 설정 창



"환경 설정" 창의 왼쪽에는 "환경 설정 그룹" 목록이 있습니다. 창의 오른쪽에는 선택한 범주에 대해 변경할 수 있는 모든 환경 설정이 있습니다.

예 : 환경 설정 변경

모든 플랫폼 보고서 창에는 설정하거나 해제할 수 있는 옵션이 있습니다. 그러나 이러한 옵션에 대한 변경 사항은 저장되지 않으므로 다음에 플랫폼을 사용할 때는 적용되지 않습니다. JMP 에

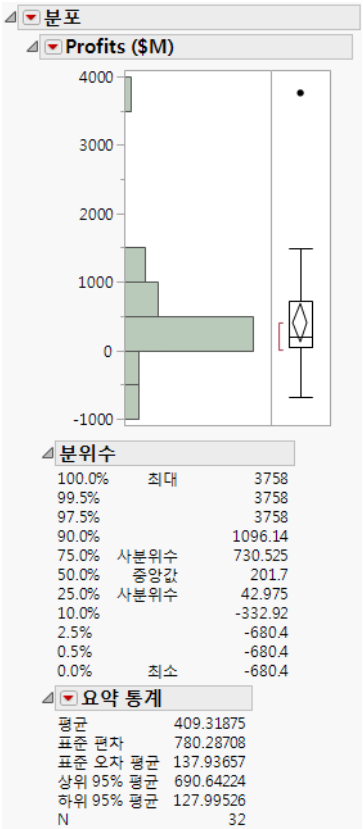
서 변경 사항을 저장하여 플랫폼을 사용할 때마다 적용하도록 하려면 "환경 설정" 창에서 해당 옵션을 변경하십시오.

이 예에서는 분포 플랫폼에서 이상치 상자 그림이 초기 보고서에 추가되지 않도록 설정하는 방법을 보여 줍니다.

기본 환경 설정을 사용하여 분포 생성

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Companies.jmp 를 엽니다.
2. **분석** > **분포**를 선택합니다.
3. Profits ($M) 를 선택하고 **Y, 열**을 클릭합니다.
4. **확인**을 클릭합니다.

그림 8.7 분포 보고서 창

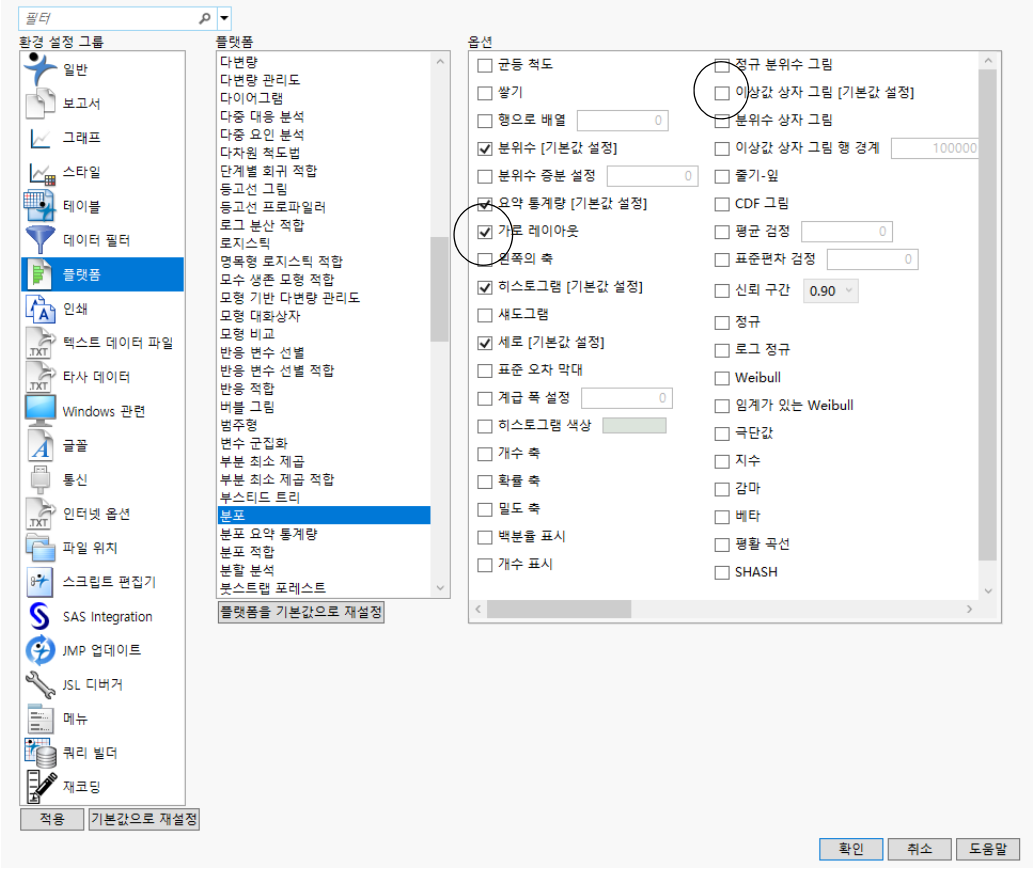


세로 히스토그램이 표시되며 그래프에는 이상치 상자 그림이 포함됩니다. 히스토그램을 가로로 변경하고 이상치 상자 그림을 제거하려면 "Profits ($M)" 의 빨간색 삼각형 메뉴에서 적절한 옵션을 선택합니다. 그러나 플랫폼을 사용할 때마다 해당 환경 설정이 적용되도록 하려면 "환경 설정" 창에서 해당 설정을 변경하십시오.

이상치 상자 그림에 대한 환경 설정 변경 및 분포 재실행

1. **파일** > **환경 설정** (Windows) 또는 **JMP** > **환경 설정** (macOS) 을 선택합니다 .
2. " 환경 설정 그룹 " 에서 **플랫폼**을 선택합니다 .
3. " 플랫폼 " 목록에서 **분포**를 선택합니다 .
4. **가로 레이아웃** 옵션을 선택하여 해당 기능을 설정합니다 .
5. **이상치 상자 그림** 옵션을 선택 취소하여 해당 기능을 해제합니다 .

그림 8.8 분포 환경 설정



6. **확인**을 클릭합니다 .
7. 분포 분석을 반복합니다 . 자세한 내용은 " 기본 환경 설정을 사용하여 분포 생성 "(210 페이지) 에서 확인하십시오 .

이제 가로 히스토그램이 표시되며 이상치 상자 그림은 표시되지 않습니다. 이러한 환경 설정은 변경하기 전까지 동일하게 유지됩니다.

모든 환경 설정에 대한 자세한 내용은 JMP 사용의 에서 확인하십시오.

JMP 와 SAS 통합

참고 : JMP 를 통해 SAS 를 사용하려면 로컬 컴퓨터나 서버에서 SAS 에 액세스해야 합니다 .

JMP 를 사용하면 다음과 같은 방식으로 SAS 와 상호 작용할 수 있습니다 .

- JMP 에서 SAS 코드 작성 또는 생성
- JMP 에서 SAS 코드 전송 및 결과 보기
- 원격 컴퓨터에서 SAS Metadata 서버 또는 SAS 서버에 연결
- 로컬 컴퓨터에서 SAS 에 연결
- SAS 데이터 집합 열기 및 불러오기
- SAS 에서 생성된 데이터 집합 가져오기 및 보기

JMP 와 SAS 의 통합에 대한 자세한 내용은 JMP 사용의 에서 확인하십시오.

예 : SAS 코드 생성

이 예에서는 초코바의 영양 데이터가 포함된 Candy Bars.jmp 샘플 데이터 테이블을 사용합니다.

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Candy Bars.jmp 를 엽니다 .
2. **분석** > **모형 적합**을 선택합니다 .
3. Calories 를 선택하고 **Y** 를 클릭합니다 .
4. Total fat g, Carbohydrate g 및 Protein g 를 선택하고 **추가**를 클릭합니다 .
5. " 모형 규격 " 의 빨간색 삼각형을 클릭하고 **SAS 작업 생성**을 선택합니다 .

그림 8.9 에서는 SAS 코드를 보여 줍니다. 모든 데이터가 표시되지는 않았습니다.

그림 8.9 SAS 코드

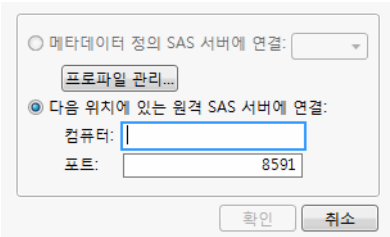
```
DATA Candy_Bars; INPUT  Calories Total_fat_g Carbohydrate_g Protein_g; Lines;
310  20  28  6
230  12  27  4
220  12  24  3
170  8  21  3
200  2.5  43  1
260  16  26  5
190  1.5  42  2
190  11  21  2
230  12  28  3
;
RUN;

PROC GLM  DATA=Candy_Bars ALPHA=0.05;
MODEL  Calories =  Total_fat_g Carbohydrate_g Protein_g;
RUN;
```

예 : SAS 코드 전송

1. **도움말 > 샘플 데이터 라이브러리**를 선택하고 Candy Bars.jmp 를 엽니다 .
2. **분석** > **모형 적합**을 선택합니다 .
3. Calories 를 선택하고 **Y** 를 클릭합니다 .
4. Total fat g, Carbohydrate g 및 Protein g 를 선택하고 **추가**를 클릭합니다 .
5. " 모형 규격 " 의 빨간색 삼각형을 클릭하고 **SAS 로 전송**을 선택합니다 .
6. **SAS 서버에 연결** 창 (그림 8.10) 에서 SAS 에 연결할 방법을 선택합니다 (아직 연결되지 않은 경우). 이 예에서는 **이 컴퓨터의 SAS 에 연결**을 선택합니다 .

그림 8.10 SAS 서버에 연결



7. **확인**을 클릭합니다 .

JMP 가 SAS 에 연결됩니다. SAS 가 모형을 실행하고 결과를 다시 JMP 로 보냅니다. 결과는 SAS 출력 , HTML, RTF, PDF 또는 JMP 보고서 형식 (JMP 환경 설정에서 형식을 선택할 수 있음) 으로 나타날 수 있습니다. 그림 8.11 에서는 JMP 보고서로 형식이 지정된 결과를 보여 줍니다. 자세한 내용은 JMP 사용의 에서 확인하십시오.

그림 8.11 JMP 보고서로 형식이 지정된 SAS 결과

SAS 시스템
The GLM Procedure

GLM 프로시저

Data

Number of Observations

Number of Observations Read	75
Number of Observations Used	75

Dependent Variable: Calories

Analysis of Variance

Calories

Overall ANOVA

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	282358	94119.3	3237.58	<.0001*
Error	71	2064.03	29.0709	.	.
Corrected Total	74	284422	.	.	.

Fit Statistics

R-Square	Coeff Var	Root MSE	Calories Mean
0.99274	2.21858	5.39174	243.027

Type I Model ANOVA

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Total_fat_g	1	185260	185260	6372.68	<.0001*
Carbohydrate_g	1	93540.4	93540.4	3217.67	<.0001*
Protein_g	1	3557.86	3557.86	122.386	<.0001*

Type III Model ANOVA

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Total_fat_g	1	111777	111777	3844.97	<.0001*
Carbohydrate_g	1	96756.1	96756.1	3328.28	<.0001*
Protein_g	1	3557.86	3557.86	122.386	<.0001*

Solution

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	-5.9643	2.89999	-2.0567	0.0434*
Total_fat_g	8.98995	0.14498	62.0078	<.0001*
Carbohydrate_g	4.0975	0.07102	57.6913	<.0001*
Protein_g	4.40133	0.39785	11.0628	<.0001*

부록 A

기술 공지사항

- Scintilla - Copyright © 1998-2017 by Neil Hodgson <neilh@scintilla.org>.

  All Rights Reserved.

  Permission to use, copy, modify, and distribute this software and its documentation for any purpose and without fee is hereby granted, provided that the above copyright notice appear in all copies and that both that copyright notice and this permission notice appear in supporting documentation.

  NEIL HODGSON DISCLAIMS ALL WARRANTIES WITH REGARD TO THIS SOFTWARE, INCLUDING ALL IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS, IN NO EVENT SHALL NEIL HODGSON BE LIABLE FOR ANY SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR OTHER TORTIOUS ACTION, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THIS SOFTWARE.

- Progressë® Telerik® UI for WPF: Copyright © 2008-2019 Progress Software Corporation. All rights reserved. Usage of the included Progress® Telerik® UI for WPF outside of JMP is not permitted.
- ZLIB Compression Library - Copyright © 1995-2005, Jean-Loup Gailly and Mark Adler.
- Made with Natural Earth. Free vector and raster map data @ naturalearthdata.com.
- Packages - Copyright © 2009-2010, Stéphane Sudre (s.sudre.free.fr). All rights reserved.

  Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

  Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

  Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

  Neither the name of the WhiteBox nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

  THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES

(INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

- iODBC software - Copyright © 1995-2006, OpenLink Software Inc and Ke Jin (www.iodbc.org). All rights reserved.

  Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

  – Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
  – Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
  – Neither the name of OpenLink Software Inc. nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

  THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL OPENLINK OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

- This program, "bzip2", the associated library "libbzip2", and all documentation, are Copyright © 996-2019 Julian R Seward. All rights reserved.

  Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

  1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

  2. The origin of this software must not be misrepresented; you must not claim that you wrote the original software. If you use this software in a product, an acknowledgment in the product documentation would be appreciated but is not required.

  3. Altered source versions must be plainly marked as such, and must not be misrepresented as being the original software.

4. The name of the author may not be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE AUTHOR "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE AUTHOR BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Julian Seward, jseward@acm.org

bzip2/libbzip2 version 1.0.8 of 13 July 2019

- R software is Copyright © 1999-2012, R Foundation for Statistical Computing.
- MATLAB software is Copyright © 1984-2012, The MathWorks, Inc. Protected by U.S. and international patents. See www.mathworks.com/patents. MATLAB and Simulink are registered trademarks of The MathWorks, Inc. See www.mathworks.com/trademarks for a list of additional trademarks. Other product or brand names may be trademarks or registered trademarks of their respective holders.
- libopc is Copyright © 2011, Florian Reuter. All rights reserved.

  Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

  – Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
  – Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and / or other materials provided with the distribution.
  – Neither the name of Florian Reuter nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

  THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF

USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

- libxml2 - Except where otherwise noted in the source code (e.g. the files hash.c, list.c and the trio files, which are covered by a similar license but with different Copyright notices) all the files are:

  Copyright © 1998 - 2003 Daniel Veillard. All Rights Reserved.

  Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

  The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

  THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL DANIEL VEILLARD BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

  Except as contained in this notice, the name of Daniel Veillard shall not be used in advertising or otherwise to promote the sale, use or other dealings in this Software without prior written authorization from him.

- Regarding the decompression algorithm used for UNIX files:

  Copyright © 1985, 1986, 1992, 1993

  The Regents of the University of California. All rights reserved.

  THIS SOFTWARE IS PROVIDED BY THE REGENTS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE REGENTS OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

  1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

3. Neither the name of the University nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

- Snowball - Copyright © 2001, Dr Martin Porter, Copyright © 2002, Richard Boulton.

  All rights reserved.

  Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

  1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

  2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

  3. Neither the name of the copyright holder nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

  THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS \"AS IS\" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

- Pako - Copyright © 2014–2017 by Vitaly Puzrin and Andrei Tuputcyn.

  Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the “Software”), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

  The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED “AS IS”, WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

- HDF5 (Hierarchical Data Format 5) Software Library and Utilities Copyright 2006 –2015 by The HDF Group. NCSA HDF5 (Hierarchical Data Format 5) Software Library and Utilities Copyright 1998-2006 by the Board of Trustees of the University of Illinois. All rights reserved. DISCLAIMER: THIS SOFTWARE IS PROVIDED BY THE HDF GROUP AND THE CONTRIBUTORS “AS IS” WITH NO WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED. In no event shall The HDF Group or the Contributors be liable for any damages suffered by the users arising out of the use of this software, even if advised of the possibility of such damage.
- agl-aglfn technology is Copyright © 2002, 2010, 2015 by Adobe Systems Incorporated. All Rights Reserved.

  Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

  – Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
  – Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
  – Neither the name of Adobe Systems Incorporated nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

  THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS “AS IS” AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

- dmlc/xgboost is Copyright © 2019 SAS Institute.

  Licensed under the Apache License, Version 2.0 (the “License”); you may not use this file except in compliance with the License. You may obtain a copy of the License at

  http://www.apache.org/licenses/LICENSE-2.0

  Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an “AS IS” BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

- libzip is Copyright (C) 1999-2019 Dieter Baron and Thomas Klausner.

  This file is part of libzip, a library to manipulate ZIP archives. The authors can be contacted at <libzip@nih.at>.

  Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

  1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

  2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

  3. The names of the authors may not be used to endorse or promote products derived from this software without specific prior written permission.

  THIS SOFTWARE IS PROVIDED BY THE AUTHORS ``AS IS@AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE AUTHORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

- OpenNLP 1.5.3, the pre-trained model (version 1.5 of en-parser-chunking.bin), and dmlc/xgboost Version .90 are licensed under the Apache License 2.0 are Copyright(C) January 2004 by Apache.org.

  You may reproduce and distribute copies of the Work or Derivative Works thereof in any medium, with or without modifications, and in Source or Object form, provided that You meet the following conditions:

- You must give any other recipients of the Work or Derivative Works a copy of this License; and
- You must cause any modified files to carry prominent notices stating that You changed the files; and
- You must retain, in the Source form of any Derivative Works that You distribute, all copyright, patent, trademark, and attribution notices from the Source form of the Work, excluding those notices that do not pertain to any part of the Derivative Works; and
- If the Work includes a “NOTICE” text file as part of its distribution, then any Derivative Works that You distribute must include a readable copy of the attribution notices contained within such NOTICE file, excluding those notices that do not pertain to any part of the Derivative Works, in at least one of the following places: within a NOTICE text file distributed as part of the Derivative Works; within the Source form or documentation, if provided along with the Derivative Works; or, within a display generated by the Derivative Works, if and wherever such third-party notices normally appear. The contents of the NOTICE file are for informational purposes only and do not modify the License. You may add Your own attribution notices within Derivative Works that You distribute, alongside or as an addendum to the NOTICE text from the Work, provided that such additional attribution notices cannot be construed as modifying the License.
- You may add Your own copyright statement to Your modifications and may provide additional or different license terms and conditions for use, reproduction, or distribution of Your modifications, or for any such Derivative Works as a whole, provided Your use, reproduction, and distribution of the Work otherwise complies with the conditions stated in this License.