

SPSS Statistics

전 세계에서 가장 많은 사용자를 보유하고 있는 글로벌 통계소프트웨어로서, 데이터 입력 및 관리, 집계 통계 분석 등의 작업을 수행하며, 분석 결과를 표와 그래프로 나타낼 수 있는 전문적인 통계분석 도구입니다.

SPSS Statistics 구성 Module

● Standard (4 Modules)

- SPSS Statistics Base
- SPSS Advanced Statistics
- SPSS Regression
- SPSS Custom Tables

● Professional (8 Modules)

- SPSS Statistics Standard (4 Modules)
- SPSS Missing Values
- SPSS Categories
- SPSS Forecasting
- SPSS Decision Trees

● Premium (14 Modules)

- SPSS Statistics Professional (8 Modules)
- SPSS Neural Networks
- SPSS Direct Marketing
- SPSS Complex Samples
- SPSS Conjoint
- SPSS Exact Tests
- SPSS Amos

1 Linked Program

KoreaPlus Statistics

SPSS Statistics에 데이터솔루션만의 Value Add Component와 서비스가 추가된 KoreaPlus 패키지입니다. (Embedded on IBM SPSS)

- 데이터솔루션 기술지원팀의 전문적인 기술 서비스와 다양한 사용자 교육 및 전문 통계 교육을 제공 받으실 수 있습니다.

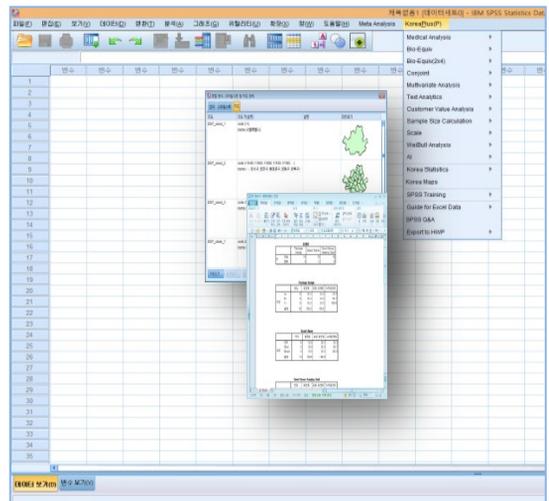
KoreaPlus Statistics 구성

+ Value-add Component

- KoreaPlus
- R을 이용한 고급 분석 기법 등

+ (주)데이터솔루션 기술지원

+ (주)데이터솔루션 교육



SPSS Statistics

- 데이터 분석으로 어려운 비즈니스 및 연구 과제를 해결

SPSS Statistics는 전 세계적으로 앞서가는 통계 툴입니다.

쉽고 빠르게 귀하의 데이터에서 인사이트를 찾아낼 수 있도록 도와드립니다.

용이한 결과 전달

더 심층적인 통찰력을 발견하고 프레젠테이션 준비 보고서, 시각화 및 지리적 분석을 통해 개선된 신뢰구간을 제공합니다.

고급 사용자를 위한 강력한 기능

명령어 구문 및 외부 프로그래밍 언어 (예: R, Python 등)와 같은 프로그래밍 옵션을 통해 분석 역량, 유연성 및 생산성을 극적으로 높입니다

SPSS Statistics 특징

- 다른 통계 툴에 비해 사용이 쉽고 비용이 적게 들며 기획에서 적용까지 분석 라이프사이클 전체를 지원
- 강력한 분석 기술과 시간을 절약할 수 있는 기능이 어떤 종류의 데이터셋이라도 빠르게 분석을 진행할 수 있도록 함
- 시간이 오래 걸리는 대용량 데이터의 준비과정을 단축
- 지리학적, 시계열 데이터를 통합하여 모델링 하고, 숨어있는 상관관계와 트렌드 발견 가능
- 불확실한 결과에 대해 분석할 시 몬테카를로 시뮬레이션을 이용하여 모델링 가능
- 피벗 기술을 활용해 표, 지도, 그래프를 작성 가능
- 메뉴기반의 유저인터페이스로 사용이 용이하고 Command Syntax와 프로그래밍도 활용 가능
- R, Python, JAVA, .Net 등 다양한 프로그래밍 옵션을 활용한 프로세스 자동화 가능
- 통합된 R 개발환경에서 R 프로그램을 개발 및 테스트 가능
- SPSS Statistics는 모든 데이터 형식, 다양한 프로그래밍 언어와 OS를 지원
- 다른 툴과 시스템 연계 용이
- Output(결과물)을 스마트폰이나 태블릿 등 다양한 디바이스에서 확인 가능
- 워드, 파워포인트, 엑셀, Cognos BI, TM1 등 다양한 형태로 Export 가능

Professional

다양한 변수를 이용하여 심층적인 모형 수립이 가능하며 생존 데이터 분석, 시계열 분석, 범주형 데이터 간의 관계 등 복잡한 관계를 파악할 수 있는 고급 통계 분석을 제공하여 데이터 품질, 데이터 복잡성, 자동화 및 예측 문제를 처리합니다

구성모듈

- SPSS Statistics Base
- SPSS Advanced Statistics
- SPSS Regression
- SPSS Custom Tables



- SPSS Missing Values
- SPSS Categories
- SPSS Forecasting
- SPSS Decision Trees

Standard

주요기능

- 선형 모델로 복잡한 관계를 설명하는 데이터의 고유한 특성에 맞게 다양한 회귀 및 고급 통계 프로시저를 제공합니다.
- 의사결정 나무로 그룹을 쉽게 식별하고 그룹 간의 관계를 발견하여 예측할 수 있도록 도와줍니다.
- 데이터 유효성 및 누락된 값을 통계적으로 확실한 결과를 얻을 수 있도록 지원합니다.
- 자동 모형 생성 기능을 이용하여 초보자들도 쉽게 최적의 시계열 모형 도출 가능 그룹 간의 관계를 파악하여 미래의 현상을 예측
- 지리적 분석을 사용하여 위치 및 시간 데이터를 통합, 탐색, 모델링을 할 수 있도록 도와줍니다.
- 계층 구조의 데이터, 생존 데이터 등 보다 다양한 형태의 데이터를 이용하여 심층적이고 정교한 분석 가능

활용산업군

- 의학, 제조, 제약 분야 : 생존 데이터 분석
- 제조: 제품 수요, 판매량, 생산량 예측
- 공공: 세금, 교통량 예측
- 마케팅: 광고 시청률 예측
- 기타 : 다양한 산업 군에서 예측 업무

분석기능강화

- Data Preparation 기능
- Bootstrapping 기능

R 연동 기능

- R 분석을 GUI 방식으로 쉽고 편리하게 수행

현재 오픈소스인 R을 전문 통계 패키지인 SPSS에서 만나볼 수 있습니다.

프로그래밍 기능 및 제품 확장성을 통해 별도의 R프로그램 활용 없이 Statistics 내에서 손쉽게 사용할 수 있도록 지원하고 있습니다.

분석

- 일반화 부스팅 회귀 예측
- 일반화 부스팅 회귀
- 이질적 상관관계
- 정준상관
- 회귀상대적중요도
- 회귀불연속
- tobit 회귀
- 로버스트 회귀
- Apriori
- 문항반응모형
- 등급반응모형
- TURF 분석
- 2개 변수, 집단 Q-Q도표
- 수정 P 값 계산
- 범주순서 정의
- 요약 독립표본 T검정
- 영과잉 계수 모형
- 비율 회귀 예측
- 비율 회귀
- 확장 rasch
- GARCH 모형
- 모수적 회귀분석
- 분위수 회귀
- Firth 로지스틱 회귀
- 방정식 시스템
- 잠재계층분석
- 서포트벡터머신(SVM)
- 밀도 기반 군집
- 밀도 군집에 의한 예측
- 밀도 실루엣
- COX 회귀 확장
- Ranfor 예측
- Ranfor 추정(Random Forest)

시스템기능

- R 작업공간가져오기
- 변수정보수집
- 데이터파일검색
- 파일 전체의 문자열 너비조정
- 레이크 가중값
- 성향 점수 매칭
- 케이스 대조 매칭
- 파일로 분할
- Programmability 변환
- 더미변수 작성

유틸리티

- Weibull 도표
- 부집단 비교
- 회귀 변수 도표
- 피벗표를 사용한 계산
- 변수 매크로 정의
- 중도절단표
- 텍스트 출력 작성
- 프로세스 데이터파일

Professional 세부모듈 소개

- SPSS Statistics Base

변수 이름, 변수 레이블, 값 레이블, 결측값과 같은 사전 정보와 모든 지정된 변수의 요약 통계 및 다중 응답 세트를 보고 합니다. 여러 종류의 변수를 설명하는데 유용한 그래픽 표시와 통계를 사용할 수 있고 이원배치표와 다원배치표를 만들어 다양한 검정과 연관 측도를 사용할 수 있습니다. 이외에 다중응답분석, 비모수검정, 신뢰성분석 등을 통해 전문적인 결과를 도출할 수 있습니다.

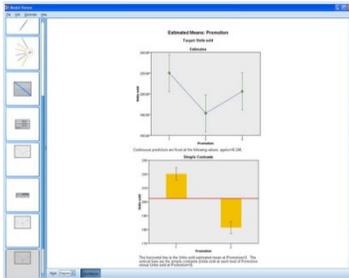
분석기능

- 보고서
 - OLAP 큐브
 - 케이스요약
 - 코드북
 - 행별요약보고서
 - 열별요약보고서
- 빈도분석
- 기술통계
- 데이터탐색
- 교차분석표
- 평균
 - 합계, 중위수, 그룹 중위수
 - 최소값, 최대값
 - 평균의 표준오차 등
- T검정
- 일원분산분석
- 일원배치 반복측도 분산분석
- GLM일변량분석
- 이변량상관계수
- 편상관계수
- 거리측도(상이성측도)
 - 선형모형
 - 선형회귀분석
 - 순서회귀분석
 - 곡선추정
 - 일부최소제곱 회귀분석
 - 최근접이웃분석
 - 판별분석
 - 요인분석
 - 군집화프로시저선택
 - 이단계군집분석
 - 계층적 군집분석
 - K-평균 군집분석
 - 비모수 검정
 - 다중응답분석
 - 신뢰성분석
 - 다차원척도법
 - 비율통계(두척도변수간 비율 통계)
 - ROC곡선, ROC분석
 - 시뮬레이션
 - 지리공간모형화(공간과시간모형화)
 - 거듭제곱 분석
 - Cohen's Weighted Kappa

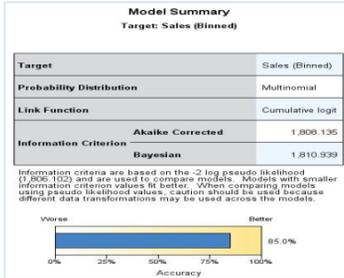
Professional 세부모듈 소개

- SPSS Advanced Statistics

GLM일반량에서 제공하는 일반선형모형을 확장하여 여러 종속 변수를 사용할 수 있도록 합니다. 종속 변수의 변동을 고정 및 무작위 성분으로 분해하는 도구를 제공합니다. 일반 선형 모형을 확장하여 데이터가 상호 관련되고 일정하지 않은 변동을 나타내어 데이터의 분산 및 공분산 모형을 만드는데 유연성을 제공합니다. GLM 모형을 제공하며 GLM의 확장인 일반화 추정 방정식도 제공합니다. 로그선형 분석을 통한 모형을 범주형 종속변수와 예측자 사이의 관계를 분석하기 위한 로그선형 모형. 층화 변수의 수준별로 개별 분석을 생성하는 Kaplan-Meier 생존분석, 공변량 값을 지정으로 지정 이벤트에 대한 시간을 모형화하는 Cox 회귀 분석을 제공합니다.



[선형혼합모델 평균추정치]



[일반화된 선형혼합모델]

Model Term	Coefficient	Std. Error	Z	Sig.	95% Confidence Interval Lower	95% Confidence Interval Upper
Intercept	47.288	3.283	14.529	.000	40.679	53.927
school_method-1	-1.616	2.369	-0.636	.526	-6.188	3.167
school_method-2	4.724	2.369	1.994	.046	0.076	9.370
school_method-3	.077	.077	1.000	.317	-.267	.413
school_method-4	-1.408	2.369	-0.594	.550	-6.036	3.219
school_method-5	.077	.077	1.000	.317	-.267	.413
teaching_method-6	-6.197	0.670	-9.260	.000	-7.510	-4.883
teaching_method-7	.077	.077	1.000	.317	-.267	.413
teaching_method-8	0.000	0.000	.000	.999	-0.000	0.000
gender-9	-0.312	0.125	-2.500	.012	-0.566	-0.057
gender-10	.077	.077	1.000	.317	-.267	.413
gender-11	-1.899	0.219	-8.673	.000	-2.329	-1.469
gender-12	.077	.077	1.000	.317	-.267	.413
gender	0.494	0.019	26.902	.000	0.457	0.530

[매개변수 추정치 테이블]

분석기능

- GLM 다변량분석
- GLM 반복 측도
- 분산 성분 추정 분석
- 선형 혼합 모형
 - 반복 공분산 유형
- 일반화 선형 모형
- 일반화 추정 방정식
- 일반화 선형 혼합 모형
- 모형선택 로그선형 분석
- 로그선형분석
- 일반화 로그선형분석
- 계층적 로그선형분석
- 로직 로그선형 분석
- 생명테이블
- Kaplan-Meier 생존분석
- Cox 회귀분석 분석
- 시간-종속 공변량 계산
- 생존분석
- 일반화 선형 모형
 - 고정효과 ANOVA
 - 공분산분석 (ANCOVA)
 - 다변량 공분산분석(MANCOVA)
 - 임의 또는 혼합효과 분산분석
 - 임의 또는 혼합효과 공분산분석
 - 다변량 분산분석(MANOVA)

Professional 세부모듈 소개

- SPSS Regression

공분산 패턴 수가 전체 케이스 수보다 작는지 여부와 데이터 입력방법과 관련 없이 모든 예측, 잔차, 영향력 통계, 각 개별 케이스 수준의 데이터를 사용하여 적합도 검정을 생성합니다. 예측변수 세트의 값을 기준으로 결정되는 특성이나 결과가 있는지 여부를 예측하려는 상황에서 유리합니다. 모형의 각 독립변수에 대한 승산비를 추정하는데 사용될 수 있습니다. 또한, 예측 변수 값에 따라 대상을 분류할 때에도 유용하며 종속변수의 범주를 제한하지 않고 사용할 수 있습니다. 자극의 강도와 자극에 대해 특정 반응을 나타내는 케이스 비율 사이의 관계를 측정할 때에도 유용하며, 종속변수와 독립변수 간의 비선형 관계 모형 탐색 시에도 유용합니다. 그리고 가중 변환의 범위를 검정하여 데이터에 가장 적합한 범위를 제공합니다.

범주 결과 예측

- 다수의 예측 요인의 경우 Score 및 Wald 방식을 사용하여 결과를 빠르게 얻을 수 있습니다.
- AIC(Akaike information criterion) 및 BIC(Bayesian information criterion)를 사용하여 모델 적합도를 평가합니다.

비선형 모델의 매개변수 추정

- CNLR을 사용하여 매개변수의 모든 조합에서 선형 및 비선형 제한조건을 사용합니다.
- 평활 손실 함수(목적 함수)를 최소화하여 매개변수를 추정하고 매개변수 표준 오류 및 상관관계(correlation)의 부트스트랩 추정을 계산합니다.

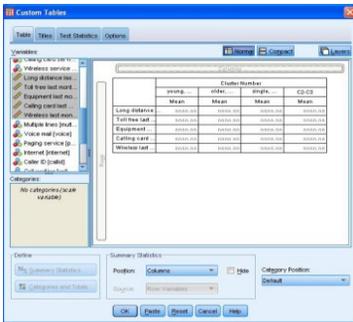
분석기능

- 이분형 로지스틱 회귀모형
- 로지스틱 회귀분석
- 다항 로지스틱 회귀분석
- 프로빗 회귀분석
- 4분위수
- 비제약된 비선형 회귀분석
- 제약된 비선형 회귀분석
- 가중 추정
- 가중치 최소 제곱법
- 2-단계 최소제곱 회귀분석

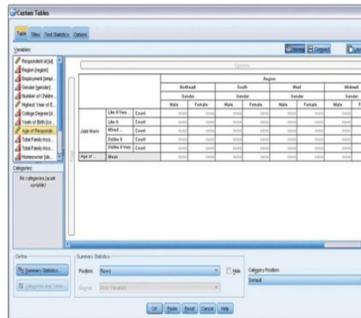
Professional 세부모듈 소개

- SPSS Custom Tables

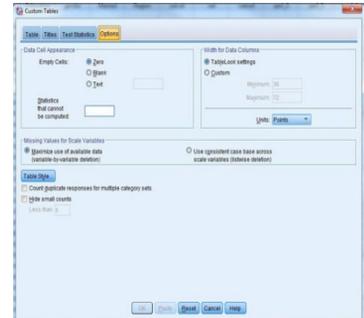
테이블 사용자 정의를 통해 단순 끌어다 놓기 테이블 작성기 인터페이스로 변수와 옵션을 선택하기만 해도 테이블을 미리 볼 수 있습니다. 변수의 단수 개수부터 척도변수의 평균 오차와 같은 산포 측정까지 모든 것을 포함할 수 있습니다. 또한 세 종류의 유의성 검증을 사용하여 행 변수와 열 변수 간의 관계를 연구할 수 있습니다.



[끌어다 놓기 테이블 빌드]



[그래픽 사용자 인터페이스]



[사용자정의가 가능한 테이블 형식]

분석기능

- 테이블 작성기 인터페이스
- 범주형 변수에 대한 단순 통계표
- 범주형 변수를 사용한 누적, 중첩 및 레이어
- 범주형 변수에 대한 총계 및 소계
- 범주형 변수의 계산된 범주
- 공유범주가 있는 변수 테이블
- 요약 통계
- 척도변수 요약
- 신뢰구간
- 검정통계량
- 다중 응답 세트
- 결측값
- 표 형식 지정 및 사용자 정의
- 표본파일

Professional 세부모듈 소개

- SPSS Missing Values

목록별 삭제가 적절한 방법인지 여부를 판단하는데 도움을 주며, 목록별 삭제가 없을 경우 결측값을 처리하는 방법을 제공합니다. 특히, 결측값 분석과 다중 대체는 결측 데이터의 패턴을 분석하며, 완벽한 폴링 결과를 생성하여 정확한 결과를 낼 수 있습니다.

누락 데이터 대체 문제점을 빠르게 진단

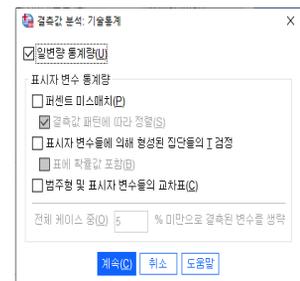
- 6개의 진단 보고서를 사용하여 다양한 각도에서 데이터를 검토합니다.
- 데이터에 대한 사례별 개요를 제공하는 데이터 패턴 보고서를 통해 누락 데이터를 진단합니다.
- 누락 데이터의 범위와 각 사례에 대해 극한 값을 판별합니다.

누락 데이터 대체 문제점을 빠르게 진단

- 데이터 세트에서 누락 패턴을 파악하고 누락 값을 타당한 추정치로 대체합니다.
- 데이터의 특성을 기준으로 가장 적합한 방식을 선택하는 자동 대체 모델을 활용하거나 대체 모델을 사용자 정의합니다.
- 각 데이터 세트에 대해 매개변수를 생성하도록 선형 회귀 또는 예측 최대 알고리즘과 같은 기법을 사용하여 작성된 개별 데이터 세트를 모델링합니다.
- 대체 내 및 대체 사이에서 변동을 고려한 추론 통계를 계산하고 추정치를 폴링하여 최종 매개변수 추정치를 획득합니다.



[결측값 분석 대화상자]



[결측값 분석 기술통계]

분석기능

- 결측값 분석
- 다중대체 (Multiple Imputation)
- 평균 추정 혹은 회귀에서 결측값 대체
- 결측 데이터에 대한 다수의 결측값 대체
- Listwise and Pairwise 통계

Professional 세부모듈 소개

- SPSS Categories

표준 통계로는 분석하기 어렵거나 분석이 불가능한 데이터를 분석하는데 도움을 줄 수 있습니다.

예측 분석, 통계 학습, 인지 지도, 환경 설정 스케일링과 같은 고급 기법을 사용하여 어떠한 특성이 사용자의 제품 또는 브랜드와 가장 밀접한 관계를 갖는지를 이해하고 고객이 다른 제품과 관련하여 사용자의 제품을 인지하는 방식을 학습할 수 있습니다.

범주에 숫자 값을 할당하여 범주형 데이터를 수량화함으로써 변환된 변수에 대한 선형 회귀 방정식을 작성할 수 있습니다. 데이터 차원을 축소하는 동시에 범주형 변수를 수량화하며, 두 명목변수 간의 관계를 설명하고 각 변수의 범주 관계를 설명합니다.

범주형 데이터에서 결과를 예측하고 관계를 발견

- 다변량 데이터와 해당 관계를 더욱 완벽하게 해석하고 쉽게 분석합니다.
- 범주형 데이터에 대한 추가 통계 조작을 수행하여 질적 변수를 양적 변수로 변환합니다.
- 시장 부문, 의료 진단, 정당 또는 생물학 종 등 연구하는 모든 유형의 범주에서 기반 관계를 그래픽으로 표시합니다.



[PROXSCAL 대화상자]



[PROXSCAL 주메뉴]

분석기능

- 범주형 회귀분석(CATREG)
- 범주형 주성분분석(CATPCA)
- 비선형 정준 상관 분석(OVERALS)
- 척도화 분석
- 대응일치분석
- 다중 대응일치분석
- 다차원척도법(PROXSCAL)
- 다차원확장(PREFSCAL)
- 신뢰도 분석
 - 플레이스 카파(Fleiss Kappa)

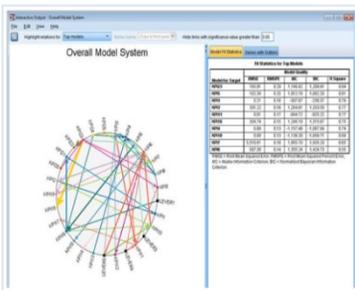
Professional 세부모듈 소개

- SPSS Forecasting

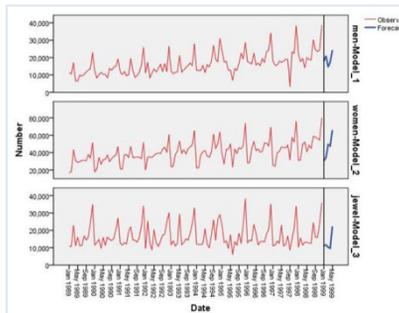
관측값을 일정한 시간 주기에서 단일 변수를 측정하여 얻을 수 있고, 이 모듈을 통해 계열의 미래값 예측을 시도하며, 계열의 모형이 증가 또는 감소할 것인지 여부를 예측할 수 있습니다.

이러한 예측은 모든 사업 또는 과학분야에 중요합니다. 지수평활, ARIMA 및 다변량 ARIMA등 모형을 추정하고 예측하며,주기동작을 식별합니다.

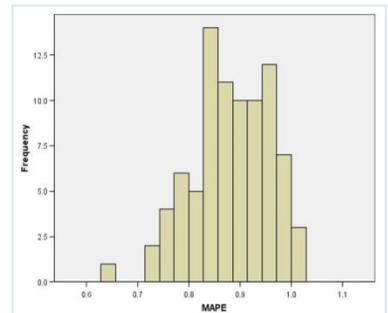
시계열 예측의 예로는 콜 센터에 매일 필요한 직원 수 예측, 특정 제품 또는 서비스에 대한 수요 예측 등이 있습니다.



[시계열 일반모델]



[예측차트]



[평균절대비율오차]

분석기능

- ARIMA
 - 정확한 최대우도법
 - 자기회귀 이동평균
 - 반적 혹은 제한적인 모형
 - Intervention 모형 분석
- 시계열 모델러
- 시계열 모형 적용
- 계절분해
 - 계열을 계절성분, 결합된 추세 및 순환 성분, 오차 성분으로 분해
 - 승법모형, 가법모형
- 스펙트럼 도표
- 시간 인과 모형
- 시간 인과 모형 적용
- 정확도 측도
- 이상값 유형
- ACF/PACF 도표
- GARCH 모형
- 지수평활법
 - 단순, 승법모형, 가법모형, Holt 모형
 - 모형에 사용된 주기, 초기계절요인, 초기 일반적 추세값 설정

Professional 세부모듈 소개

- SPSS Decision Trees

그룹을 더욱 잘 식별하고 이들 사이의 관계를 발견하며 향후 이벤트를 예측하도록 도와줍니다.

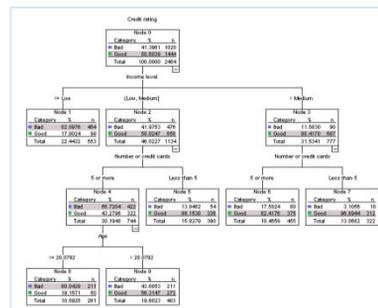
높은 시각적 분류와 의사결정 트리를 제공하여 직관적 방식으로 범주 결과를 제공할 수 있어 기술적 배경이 없는 대상에서 범주분석을 명확히 제공합니다. 또한 독립 변수값을 기반으로 종속 변수값을 예측하거나 케이스를 집단으로 분류합니다. 이 모듈은 탐색 및 확인 분류분석을 위한 검증 도구를 제공합니다.

Risk			
Estimate	Std. Error		
.205	.008		
Growing Method: CHAID			
Dependent Variable: Credit rating			

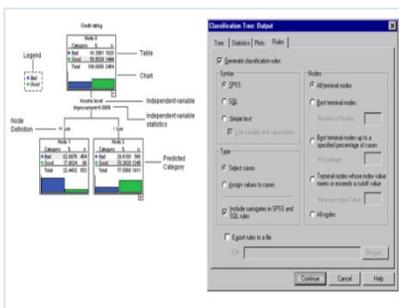
Observed	Predicted		
	Bad	Good	Percent Correct
Bad	665	355	65.2%
Good	149	1295	89.7%
Overall Percentage	33.0%	67.0%	79.5%

Growing Method: CHAID
Dependent Variable: Credit rating

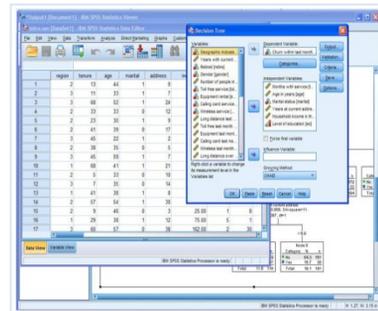
[위험 및 분류 테이블]



[CHAID 방법]



[트리 기반 분류모델]



[대화 상자 열기]

분석기능

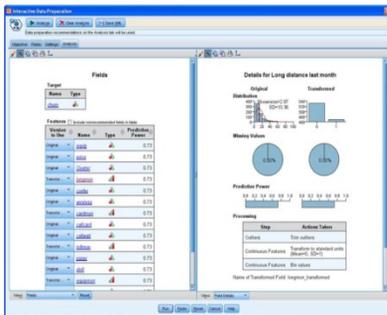
- 의사결정나무 작성
- 나무 편집기
- C&RT
- CHAID
- Exhaustive CHAID
- QUEST

Professional 분석기능 소개

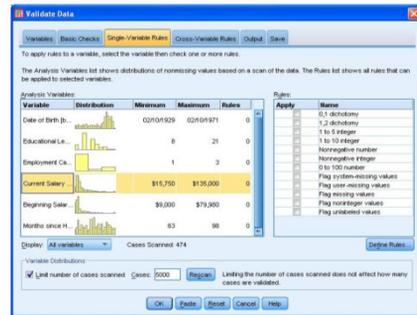
- SPSS Data Preparation

분석이전의 데이터 준비 단계로서 분석에 집중할 수 있도록 준비 단계를 간소화하기 위한 고급기술을 제공합니다.

규칙을 적용하여 각 변수의 측정 레벨에 따라 데이터 유효성 검사를 간소화하고 수동적인 검사를 제거합니다. 자동 감지를 통해 이상치 탐지를 빠르게 할 수 있습니다. 지도학습, 비지도학습 및 하이브리드(Hybrid) 방법으로 최적화 구간화를 실시합니다. 자동 데이터 준비를 통해 오류를 감지하여 수정하고 결측값을 대체합니다.



[자동화된 데이터 준비]



[데이터 유효성 검증]

	Occurrence or Reason		Variable Impact Statistics			
	Frequency	Percent	Minimum	Maximum	Mean	Std. Deviation
Dead on arrival	0	0%				
India Rankin score	0	0%				
CAT score result	0	0%				
Code-breaking drugs	0	0%				
Dead in hospital	0	0%				
Treatment result	0	0%				
Post-event preventative surgery	0	0%				
Post-event rehabilitation	0	0%				
Rankin score at 1 month	0	0%				
Rankin score at 3 months	0	0%				
Rankin score at 6 months	0	0%				
Barthel index at 1 month	13	61.9%	064	124	100	021
Barthel index at 3 months	1	4.8%	000	000	000	
Barthel index at 6 months	1	4.8%	000	000	000	
Recorded Barthel index at 1 month	0	0%				
Recorded Barthel index at 3 months	0	0%				
Recorded Barthel index at 6 months	0	0%				
Stroke between release and 1 month	0	0%				
Stroke between 1 and 3 months	0	0%				
Stroke between 3 and 6 months	0	0%				
Length of stay for rehabilitation	0	0%				
Total treatment and rehabilitation costs in thousands	6	28.6%	120	411	202	112
Missing Proportion	21	100.0%				
Overall			064	411	127	076

[이상 항목 감지 이유 요약]

Case	Single-Variables ^a		Cross-Variables	Receipt	Identifier	pvalue
	Variable	Value				
20	DEATH	1	DEATH	DEATH	102070026	255184
49	DEATH	1	DEATH	DEATH	871782252	237418
128	DEATH	1	DEATH	DEATH	680183205	215041
139	DEATH	1	DEATH	DEATH	1205050599	698521
162	DEATH	1	DEATH	DEATH	8548895939	125384
176	DEATH	1	DEATH	DEATH	0333204686	893285
274	DEATH	1	DEATH	DEATH	1138840485	112264
310	DEATH	1	DEATH	DEATH	20800200204	863285
414	DEATH	1	DEATH	DEATH	3351107142	462020
437	DEATH	1	DEATH	DEATH	2348729626	723384
466	DEATH	1	DEATH	DEATH	7183481292	619540
467	DEATH	1	DEATH	DEATH	9168894176	652070
462	DEATH	1	DEATH	DEATH	2137520354	723384
537	DEATH	1	DEATH	DEATH	5248522206	920076
544	DEATH	1	DEATH	DEATH	1609587482	508108
620	DEATH	1	DEATH	DEATH	6141688866	620764
620	DEATH	1	DEATH	DEATH	3397891810	539412
630	DEATH	1	DEATH	DEATH	3397891810	539412
639	DEATH	1	DEATH	DEATH	3962622031	327422
644	DEATH	1	DEATH	DEATH	4271782383	748432
649	DEATH	1	DEATH	DEATH	0650867550	618089
653	DEATH	1	DEATH	DEATH	0650867550	601448
722	DEATH	1	DEATH	DEATH	0418125580	877354
748	DEATH	1	DEATH	DEATH	8744721360	539412
752	DEATH	1	DEATH	DEATH	4403307441	620764
860	DEATH	1	DEATH	DEATH	9714872452	237418
881	DEATH	1	DEATH	DEATH	6613279456	874276
916	DEATH	1	DEATH	DEATH	2876783702	601318
933	DEATH	1	DEATH	DEATH	2087437472	680283
1010	DEATH	1	DEATH	DEATH	5284808838	657630
1028	DEATH	1	DEATH	DEATH	8021897483	185703
1054	DEATH	1	DEATH	DEATH	0365897644	287630
1173	DEATH	1	DEATH	DEATH	8737861990	185707

a. The number of variables that violated the rule follows each rule.

[데이터 유효성 검증 사례]

분석기능

- 검증 규칙
- 데이터 검증, 타당성 검사 (Validation Data)
- 자동 데이터 준비 (Automated Data Preparation)
- 특수 케이스 식별 (Identify Unusual Cases)
- 최적화 구간화
- 케이스 대조 매칭(Case-Control Matching)

Professional 분석기능 소개

- SPSS Bootstrapping

평균값, 중앙값, 비율, 오즈비, 상관계수 또는 회귀계수와 같은 추정 표준 오차와 신뢰 구간의 강력한 추정값을 도출하기 위한 방법이며, 가설 검정을 구축하는데도 사용이 가능합니다. 부트스트래핑은 이분산성 잔차의 회귀 모형에 적합한 샘플 수가 작은 경우 등과 같이 불확실한 경우나, 중앙값, 사분위수, 등의 신뢰구간을 계산하는 경우 등 매우 복잡한 공식을 필요로 하는 경우에 매개변수적 추정의 대안으로 가장 유용합니다.

- 데이터로부터 반복적으로 표본 재추출(Re-sampling)을 수행하므로 분포를 쉽고 빠르게 추정할 수 있습니다.
- 더욱 정확한 분석을 위해 수천 개의 다른 버전(version)의 데이터 셋을 생성하여 Bootstrapping을 수행할 수 있습니다.

분석기능

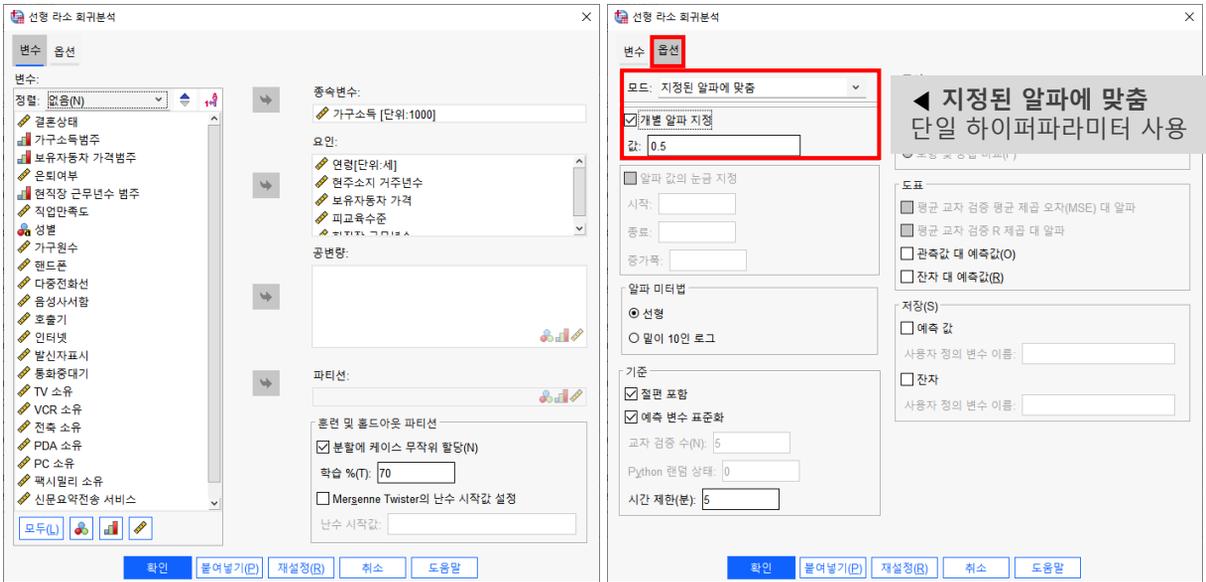
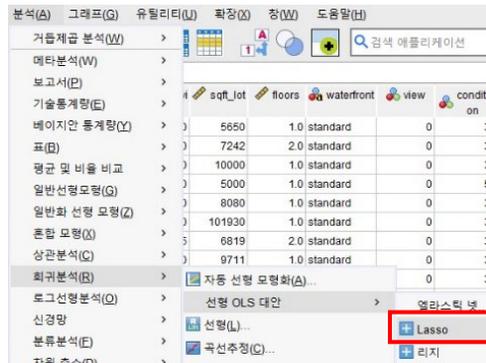
- 기술통계
 - ✓ 상관관계/비모수 상관관계
 - ✓ 교차분석
 - ✓ 기술통계
 - ✓ 실험
 - ✓ 빈도
 - ✓ 평균
 - ✓ 편상관
 - ✓ T 검정
- 모델링
 - ✓ Cox 회귀
 - ✓ 판별식
 - ✓ 일반화된 선형모형(GENLIN)
 - ✓ 일반선형모형(GLM)
 - ✓ 선형복합모형
 - ✓ 로지스틱 회귀
 - ✓ 명목 회귀
 - ✓ 일원배치
 - ✓ 이원 다항로지스틱 서수 회귀
 - ✓ 회귀분석
 - ✓ UniAnova

New SPSS Statistics

- 라쏘(Lasso) 회귀 분석 기능이 추가되었습니다.

라쏘(Lasso) 회귀 분석이란 모델의 과대적합을 방지하기 위해 L1 손실(가중치의 절대값)에 대해 정규화(규제)된 선형 회귀 모델을 추정하는 방법입니다.

29 버전에서는 하이퍼파라미터에 대한 트레이스 플롯, 교차 검증과 같은 다양한 하이퍼파라미터 최적화 방법을 제공합니다.



◀ 지정된 알파에 맞춤
단일 하이퍼파라미터 사용

		N	퍼센트
보기	학습	4516	70.6%
	검증	1884	29.4%
유효		6400	100.0%
제외됨		0	
전체		6400	

[케이스 처리 요약]

알파	학습 R 제곱	홀드아웃 R 제곱
.500	.817	.427

a. 종속변수: 가구소득
b. 모델: 연령, 거주년수, 자동차가격, 교육수준, 근무년수

[모델 요약]

알파	값 표준화 ^a			
	평균	표준 편차	표준화 계수	비표준화 계수
.500			69.188	64.093
	절편 ^b			
	[연령=18]	.002	.045	.000
	[연령=19]	.007	.081	.000
	[연령=20]	.010	.097	.000
	[연령=21]	.009	.094	.000
	[연령=22]	.011	.103	.000

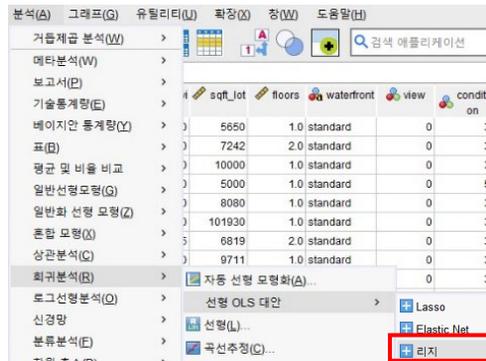
[라쏘 회귀 계수]

New SPSS Statistics

- 릿지(Ridge) 회귀 분석 기능이 추가되었습니다.

릿지(Ridge) 회귀 분석이란 모델의 과대적합을 방지하기 위해 L2 손실(가중치의 제곱합)에 대해 정규화(규제)된 선형 회귀 모델을 추정하는 방법입니다.

29 버전에서는 하이퍼파라미터에 대한 트레이스 플롯, 교차 검증과 같은 다양한 하이퍼파라미터 최적화 방법을 제공합니다.



표시 - 모델 비교
하이퍼파라미터 탐색 구간에서 산출된 모델 비교

교차검증을 통한 알파 선택
알파 값에 대한 탐색 구간을 설정하여 하이퍼파라미터 최적화

	N	퍼센트
보기 학습	4445	69.5%
검증	1955	30.5%
유효	6400	100.0%
제외됨	0	
전체	6400	

알파	교차 검증 중점 수	학습 R 제곱	평균 테스트 서브세트 R 제곱	홀드아웃 R 제곱
1.000	5.000	.904	.465	.434

a. 종속변수: 가구소득
b. 모델: 직업만족도, 연령, 거주년수, 자동차가격, 교육수준, 근무년수, 가구원수

알파	평균 테스트 서브세트 R 제곱	평균 테스트 서브세트 MSE
1.000	.465	3139.750
.900	.465	3139.918
.700	.465	3140.255
.500	.465	3140.593
.300	.465	3140.931
.100	.465	3141.269

[케이스 처리 요약]

[모델 요약]

[알파 값에 따른 모델 비교]

New SPSS Statistics

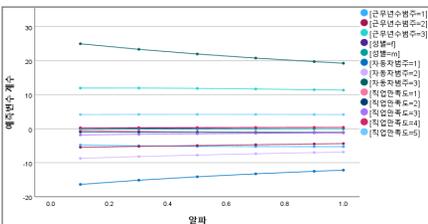
- 엘라스틱 넷(Elastic Net) 회귀 분석 기능이 추가되었습니다.

엘라스틱 넷(Elastic Net) 회귀 분석이란 모델의 과대적합을 방지하기 위해 라쏘(L1), 린지(L2) 회귀 분석의 정규화(규제) 방법을 혼합한 회귀 모델을 추정하는 방법입니다.

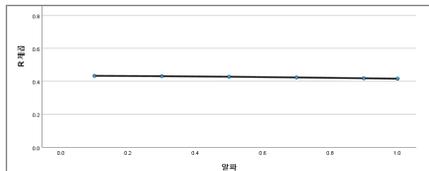
29 버전에서는 하이퍼파라미터에 대한 트레이스 플롯, 교차 검증과 같은 다양한 하이퍼파라미터 최적화 방법을 제공합니다.

SPSS Statistics 29 Professional의 Elastic Net 회귀 분석 인터페이스 스크린샷입니다. '회귀분석(R)' 메뉴에서 'Elastic Net'이 선택되어 있습니다. '선형 탄력적 순 회귀' 대화 상자에서 '변수' 목록과 '모드' (Trace Plot)가 설정되어 있으며, '개별 L1 비율 지정'과 '알파 값의 눈금 지정' 옵션이 활성화되어 있습니다. '알파 미터법'은 '선형'으로 설정되어 있습니다.

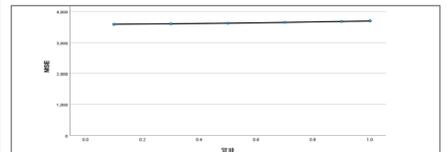
트레이스 플롯
알파 값에 대한 탐색 구간을
설정하여 트레이스 플롯 생성



[알파 값에 따른 회귀 계수]



[알파 값에 따른 R 제곱 변동]



[알파 값에 따른 MSE 변동]

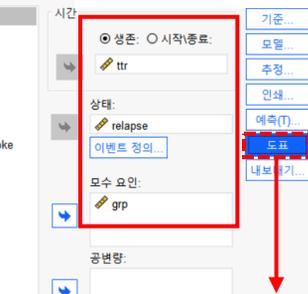
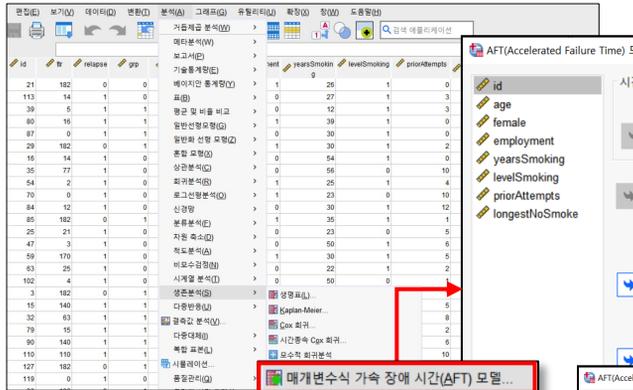
New SPSS Statistics

- 모수적 가속고장시간모형(Parametric Accelerated Failure Time) 이 추가되었습니다.

29 버전에서는 생존시간이 특정 분포를 따르며 공변량과의 선형 관계에 대한 가정을 기반으로 하는 생존자료 회귀 분석 기법인 모수적 가속고장시간(AFT) 모형을 새롭게 제공합니다.

▼변수 및 옵션 설정

시간: 생존 시간/시작 및 종료 시간
상태: 생존 상태



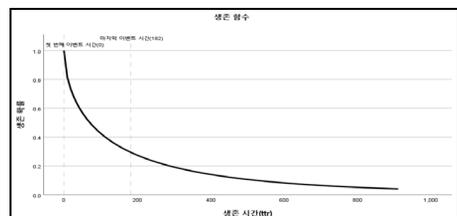
로그 가능도	-202.498
-2 로그 우도	404.996
Akaike 정보 기준(AIC)	410.996
Hurvich & Tsai 기준(AICC)	411.216
Schwarz Bayesian Criterion (BIC)	419.178

[Model Fit Statistics]

모수	계수	표준오차	카이제곱*	유의확률	95% 신뢰구간	지수 계수	Exp. 95% 신뢰구간		
				확률	상한	하한	상한		
Intercept	5.379	.296	330.700	<.001	4.799	5.958	216.737	121.387	386.986
(grp=0)의	-1.032	.390	7.025	.008	-1.796	-.269	.356	.166	.764
(grp=1)의	.000*	.000*							
(스케일)	1.665	.164			1.373	2.019			

a. 자유도 = 1
b. 이 모수는 종속되었으므로 0으로 설정됩니다.
c. 카이제곱 통계, p-값 및 지수 통계는 척도 모수에 대해 추정되지 않습니다.

[AFT 모델 회귀 분석 모수]

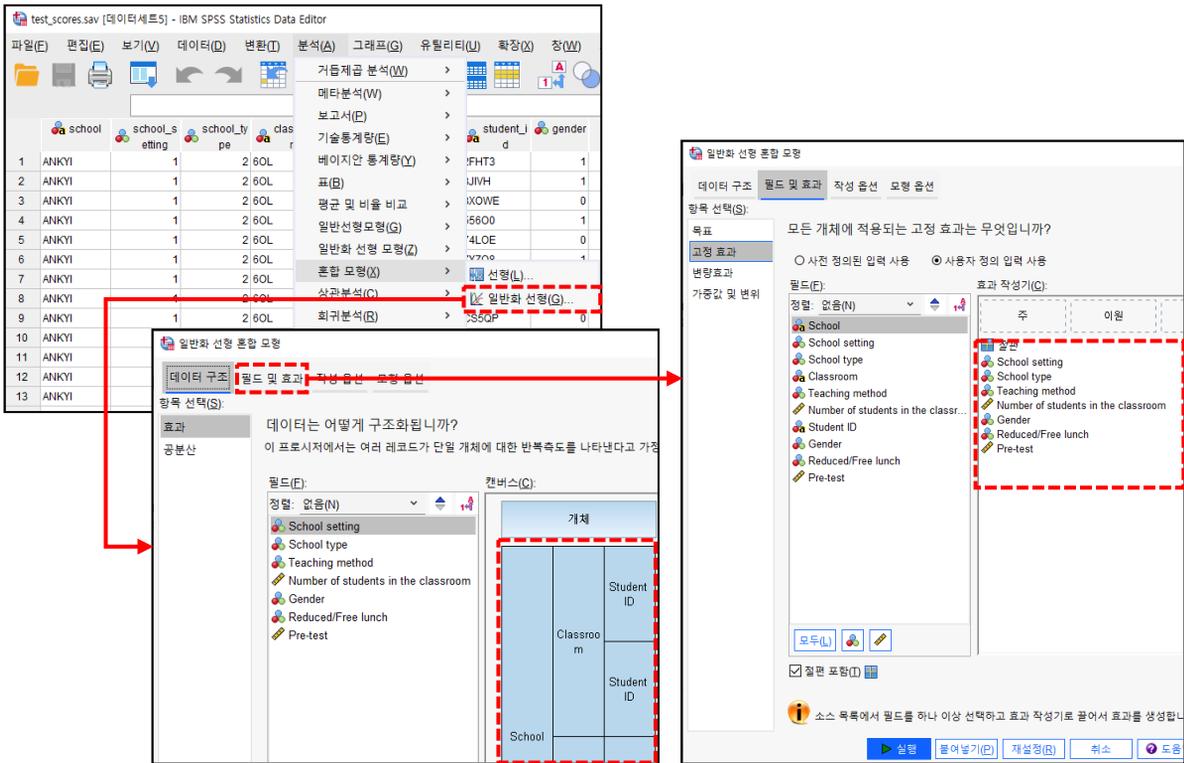


[생존 시간에 따른 생존확률 함수 출력]

New SPSS Statistics

- 의사-결정계수 척도(Pseudo-R-Squared measures)가 추가되었습니다.

29 버전에서는 선형 혼합 모델과 일반화 선형 혼합 모델 출력 결과에 의사-결정계수와 급내 상관계수(ICC; Intra-class Correlation Coefficient)를 새롭게 제공합니다.



유사-R 제곱 척도	주변	.767
	조건부	.944

[결정 계수]

▲ 결정계수

선형 모형으로 설명되는 분산의 비율로 일반적으로 사용되는 통계량

전반적인 ICCs	수정된	.762
	조건부	.178
SUBJECT = school ICC	수정된	.524
	조건부	.122
SUBJECT = school * classroom ICC	수정된	.237
	조건부	.055

[급내 상관계수]

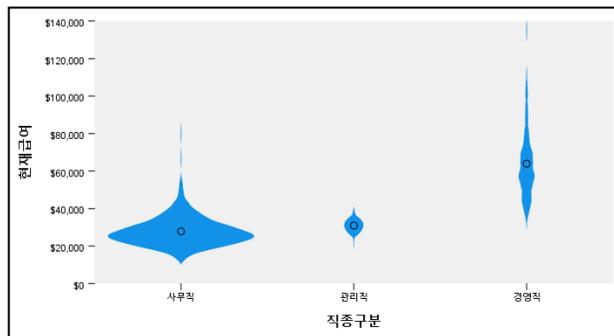
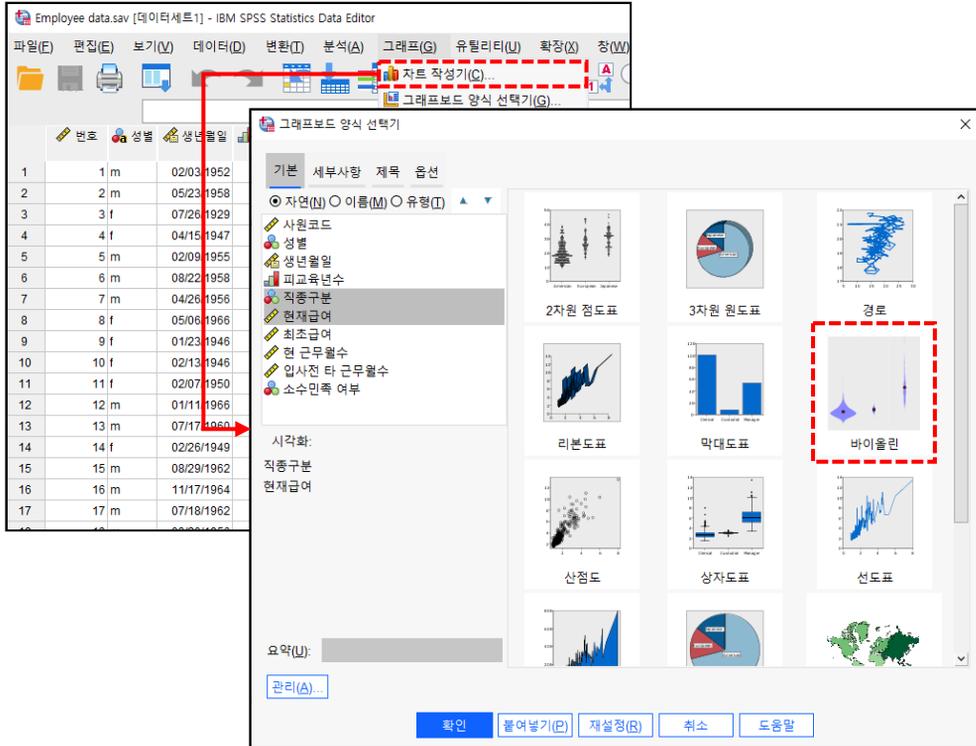
▲ 급내 상관계수

다층/계층형 데이터에서 임의의 그룹 요인에 의해 설명되는 분산 비율을 정량화 한 통계량

New SPSS Statistics

- 바이올린 플롯(Violin plots)이 추가되었습니다.

29 버전에서는 박스 플롯(Box Plot)과 커널밀도 플롯(Kernel Density Plots)를 합친 바이올린 플롯을 새롭게 제공합니다.



[바이올린 플롯]

▲ 바이올린 플롯

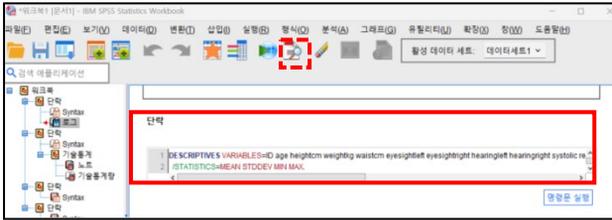
박스플롯과 같이 요약통계량을 알 수 있으며 커널 밀도 플롯과 같이 각 변수의 밀도를 확인할 수 있어 데이터 분포의 시각화하는데 사용

- Embedded on IBM SPSS Statistics 29 Professional

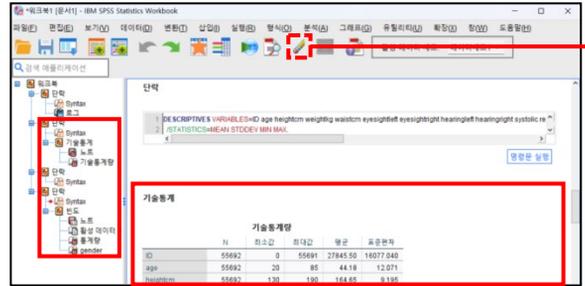
New SPSS Statistics

- 워크북 모드(Workbook Mode) 기능이 강화되었습니다.

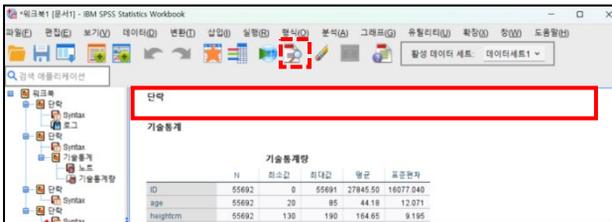
29 버전에서는 워크북 애플리케이션에 명령문 모두 숨기기 및 표시 도구와 모든 출력 지우기 도구가 새롭게 제공됩니다.



[명령문 표시]

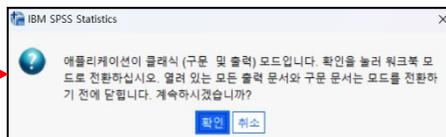
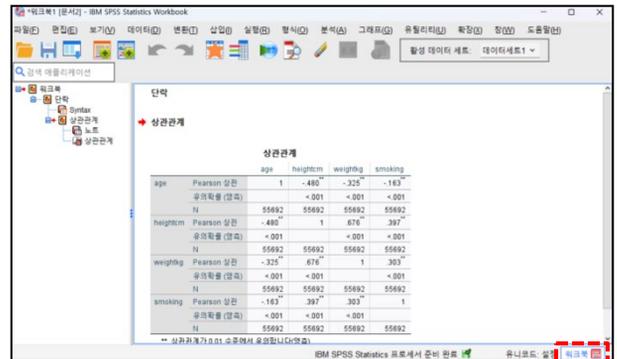
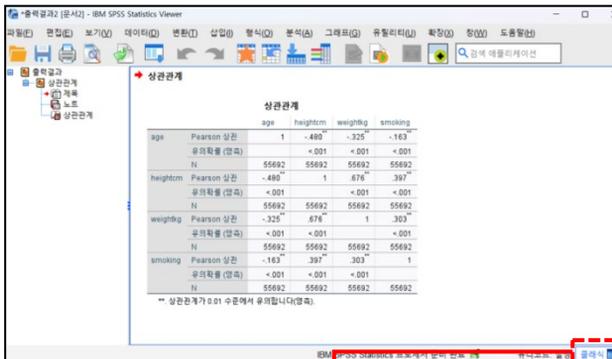


[출력 지우기]



[명령문 숨기기]

29 버전에서는 상태 표시줄에 클래식/워크북 전환 버튼을 새롭게 제공합니다.



▲ 출력 결과 필요 시 저장 필수

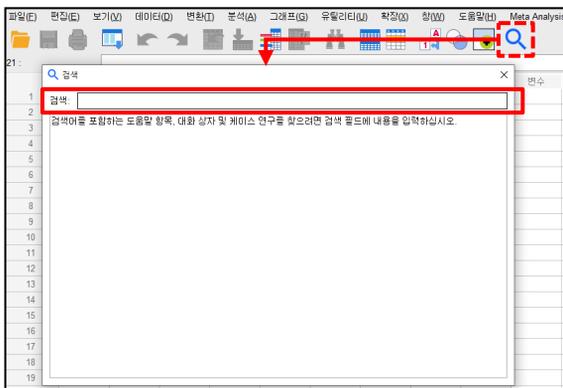
New SPSS Statistics

- 검색 기능이 강화되었습니다.

29 버전에서는 원하는 도구를 검색할 수 있는 톨 바가 즉시 검색하기 쉽게 개선되었습니다.

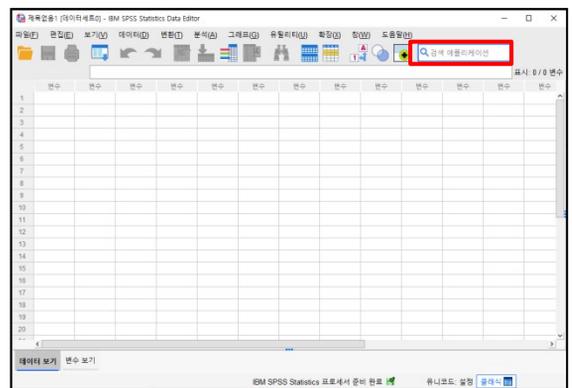
28 버전

- [돋보기 아이콘 클릭]-[검색]



29 버전

- 바로 검색 키워드 입력 가능

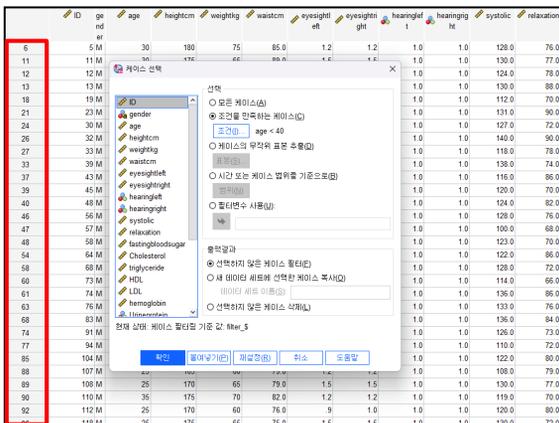


- 제외된 케이스 숨김 처리 기능을 제거했습니다.

29 버전에서는 [데이터]-[케이스 선택]에서 제외된 케이스를 표시하도록 변경되었습니다. 제외된 케이스를 보이지 않게 할 땐 옵션을 통해 케이스를 삭제할 수 있습니다.

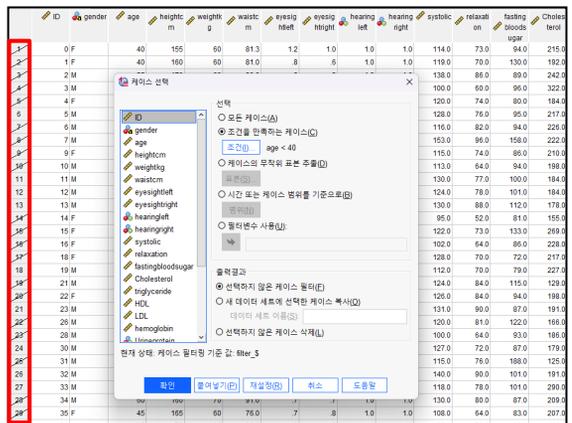
28 버전

- 제외 된 케이스가 숨김 처리가 되어서 어떤 케이스가 제외되었는지 알 수 없음



29 버전

- 케이스 선택 시 제외 된 케이스를 보여줌



KoreaPlus Statistics

- Embedded on IBM SPSS Statistics 29 Professional

New SPSS Statistics

- Python과 R 버전이 업데이트 되었습니다.

28 버전	29 버전
 <p>Python 3.8.x</p>  <p>R 4.0.x</p>	 <p>Python 3.10.4</p>  <p>R 4.2.0</p>

KoreaPlus Statistics

- Embedded on IBM SPSS Statistics 29



KoreaPlus Statistics 란 무엇입니까?

- ✓ KoreaPlus Statistics 란 SPSS Statistics 에 데이터솔루션만의 Value Add Component와 서비스가 추가된 KoreaPlus 패키지 입니다.

KoreaPlus Statistics 에는 어떤 기능들이 있습니까?

- ✓ 데이터 스테이션 제공(Data Station)
- ✓ AI 분석기법 제공 I (RNN 기반 시계열 학습 및 예측)
- ✓ AI 분석기법 제공 II (ML기반 연속형/범주형 데이터 학습 및 예측)
- ✓ TA 분석기법 제공 (Text Analytics)
- ✓ 의학 관련 분석기법 제공 (Medical Analysis, Subgroup Analysis)
- ✓ 추가 분석기법 제공 I (Conjoint, Multivariate Analysis)
- ✓ 추가 분석기법 제공 II (Customer Value Analysis)
- ✓ 추가 분석기법 제공 III (Sample Size Calculation, Weibull Analysis, Scale)
- ✓ 추가 분석 기법 제공 IV (PSM(Propensity Score Matching), Meta-analysis)
- ✓ 국가통계포털 링크 (Korea Statistics)
- ✓ Maps graph를 작성하기 위한 한국지도 제공 (Korea Maps)
- ✓ 초보자를 위한 SPSS 활용하기 방법 및 샘플 데이터 제공(SPSS Training)
- ✓ 엑셀 데이터 활용을 위한 도움말 기능 (Guide for Excel Data)
- ✓ 출력결과(Output)를 한글파일(HWP)로 바로 내보내기 (Export to HWP)

데이터 스테이션 제공(Data Station)

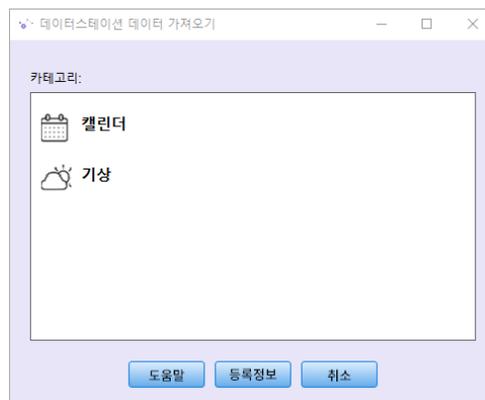
- 데이터솔루션의 자체 솔루션인 데이터 스테이션에서 분석에 필요한 데이터셋을 가져옵니다. 데이터 스테이션은 날짜, 기상 데이터를 수집, 정제하여 제공합니다.
- ✓ 캘린더 옵션은 양력 기준으로 음력, 국경일/공휴일, 24절기 정보를 가진 데이터셋을 가져옵니다,
- ✓ 기상 옵션은 일자별 지역별 날씨 정보와 대기질 정보를 가진 데이터셋을 가져옵니다.
- ✓ KoreaPlus의 Data Station은 KoreaPlus Statistics 27버전 이상을 지원합니다.
- ✓ Data Station 사이트에서 데이터를 가져오기 때문에 인터넷이 연결되어 있어야 합니다.

	sol_ymd	sol_year	sol_month	sol_day	sol_nday	sol_leapy	sol_week	lun_year
			h			ear		
1	20200101	2020	1	1	0	0	3	2019
2	20200102	2020	1	2	0	0	4	2019
3	20200103	2020	1	3	0	0	5	2019
4	20200104	2020	1	4	0	0	6	2019
5	20200105	2020	1	5	0	0	0	2019
6	20200106	2020	1	6	0	0	1	2019
7	20200107	2020	1	7	0	0	2	2019
8	20200108	2020	1	8	0	0	3	2019
9	20200109	2020	1	9	0	0	4	2019
10	20200110	2020	1	10	0	0	5	2019
11	20200111	2020	1	11	0	0	6	2019
12	20200112	2020	1	12	0	0	0	2019
13	20200113	2020	1	13	0	0	1	2019
14	20200114	2020	1	14	0	0	2	2019
15	20200115	2020	1	15	0	0	3	2019

[캘린더 데이터]

	main_name	sub_name	city_name	w_date	area_code	actual_yr	w_description	w_avg_ta
1	강원도	남부	태백시	2020-01-01	1003	1	맑음	-4.9
2	강원도	남부	태백시	2020-01-02	1003	1	맑음	-3.4
3	강원도	남부	태백시	2020-01-03	1003	1	맑음	-4
4	강원도	남부	태백시	2020-01-04	1003	1	맑음	-1.9
5	강원도	남부	태백시	2020-01-05	1003	1	맑음	-2.9
6	강원도	남부	태백시	2020-01-06	1003	1	흐림	0
7	강원도	남부	태백시	2020-01-07	1003	1	비	5.2
8	강원도	남부	태백시	2020-01-08	1003	1	비	2.7
9	강원도	남부	태백시	2020-01-09	1003	1	맑음	-2.6
10	강원도	남부	태백시	2020-01-10	1003	1	구름조금	-1.4
11	강원도	남부	태백시	2020-01-11	1003	1	맑음	0
12	강원도	남부	태백시	2020-01-12	1003	1	구름조금	-3.2
13	강원도	남부	태백시	2020-01-13	1003	1	맑음	-4.8
14	강원도	남부	태백시	2020-01-14	1003	1	맑음	-4.2
15	강원도	남부	태백시	2020-01-15	1003	1	맑음	-5.6

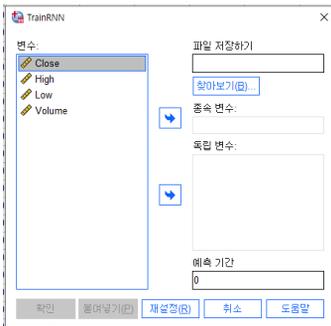
[기상 데이터]



[Data Station의 대화상자]

AI (RNN기반 시계열 학습 및 예측)

- RNN(Recurrent Neural Network)은 히든 노드가 방향을 가진 엷지로 연결되어 순환구조를 이루는 Neural Network의 한 종류.
- KoreaPlus의 RNN 시계열 학습 및 예측을 활용하면 통계적 분석에 익숙하지 않은 사용자들이 손쉽게 시계열 분석 수행 가능
- ✓ LSTM(Long Short Term Memory) 사용
- ✓ 분석하고자 하는 값과 시계열 분석에 영향을 줄 수 있는 요소들, 갖고 있는 데이터를 예측하여 연장할 기간 등을 설정하고 예측 된 값이 추가 된 데이터 시트를 얻음
- ✓ Train loss Value 정보를 출력하여 예측한 값과 실제 값의 차이 확인



[RNN시계열학습]



[RNN시계열학습-그래프]

	Volume	Close	Low	High	predictive_value
1	35718.00	12.00	11.00	16.00	35438.66
2	41439.00	15.00	11.00	23.00	43186.40
3	41331.00	15.00	14.00	23.00	47301.49
4	48956.00	17.00	14.00	17.00	47420.49
5	56286.00	17.00	15.00	21.00	53163.75
6	51469.00	18.00	16.00	21.00	58528.33
7	64720.00	18.00	13.00	21.00	55191.26
8	62290.00	20.00	20.00	24.00	64096.23
9	77150.00	22.00	22.00	33.00	63164.59
10	64280.00	22.00	21.00	25.00	73891.81

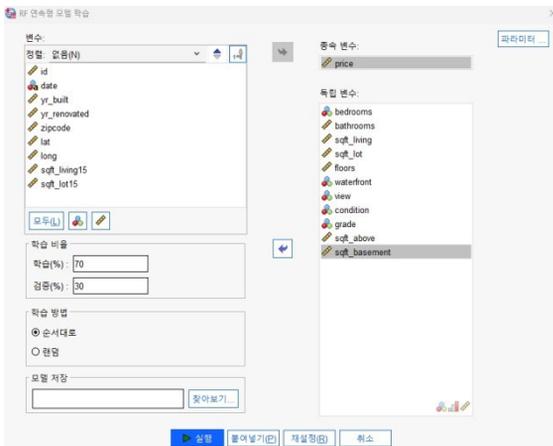
[결과-예측된 데이터시트]

[Check Point]

- * KoreaPlus의 AI는 KoreaPlus Statistics 25버전 이상부터 가능합니다.
- * 딥러닝 알고리즘은 많은 연산이 필요하기 때문에 CPU만으로는 학습에 많은 시간이 소요 됩니다.
속도를 향상시키기 위해 *GPU(Graphic Processing Unit)가 있는 환경에서 학습 하는 것을 권장합니다.
- * GPU 환경확인
 - ▶ PC의 종류가 64비트이면서 그래픽 카드가 **NVIDIA**일 경우 사용이 가능 (속도 최소10배 차이)
 - [Windows PC 종류 확인] 제어판 > 시스템 및 보안 > 시스템 > 시스템종류 : 64비트 운영체제
 - [Windows 그래픽 카드 확인] 제어판 > 장치 관리자 > 디스플레이 어댑터에 **NVIDIA**가 존재하는지 확인

AI (ML기반 연속형/범주형 데이터 학습 및 예측)

- 컴퓨터가 수집된 데이터를 학습하여 패턴 및 특징을 발견하고 다양한 머신러닝 알고리즘을 통해 이를 활용하는 기계적 학습 및 추론 방법
- AI챗봇, IoT, 스마트워치, 자율주행 자동차, 스마트 팩토리과 같이 제조, 금융, 유통, 통신 등 다양한 분야에서 적용 가능



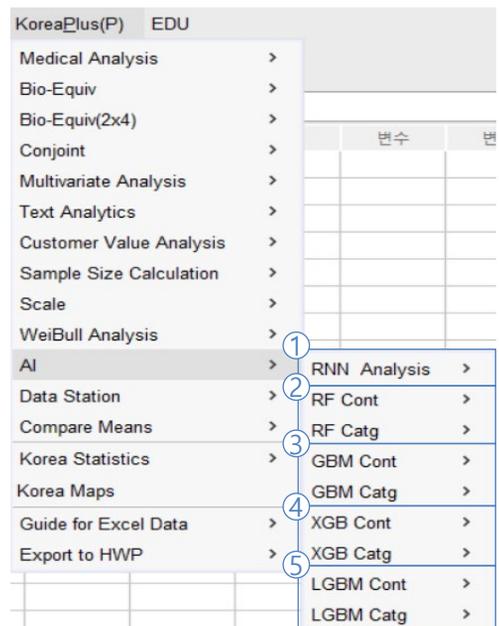
[RF 연속형 모델 학습]

	view	condition	grade	sqft_above	sqft_basement	predict_value
1	0	3	7	1180	0	259955.50
2	0	3	7	2170	400	545482.25
3	0	3	6	770	0	205822.00
4	0	5	7	1050	910	553779.73
5	0	3	8	1680	0	464550.00
6	0	3	11	3890	1530	1314247.5
7	0	3	7	1715	0	272871.00
8	0	3	7	1060	0	288205.00
9	0	3	7	1050	730	274617.00
10	0	3	7	1890	0	324559.50
11	0	3	8	1860	1700	677423.49
12	0	4	7	860	300	464553.50
13	0	4	7	1430	0	345696.50
14	0	4	7	1370	0	358940.50
15	0	3	7	1810	0	550681.00
16	3	3	9	1980	970	752101.50

[RF 연속형 모델 예측 결과]

✓ AI (Artificial Intelligence)

- RNN : 히든 노드가 방향을 가진 엷지로 연결되어 순환구조를 이루는 Neural Network의 한 종류
- RF : 배깅을 기반으로 동작하는 대표적인 알고리즘
- GBM : 부스팅을 기반으로 동작하는 가장 대표적인 알고리즘
- XGB : GBM에서 병렬구조를 적용하여 동시 다발적으로 학습을 수행하는 알고리즘
- LGBM : GBM에서 중요도가 낮은 데이터는 제거하고 서로 상관성이 높은 데이터를 하나로 통합하는 알고리즘



* Cont : 연속형 / Catg : 범주형

의학 관련 분석 기법 제공

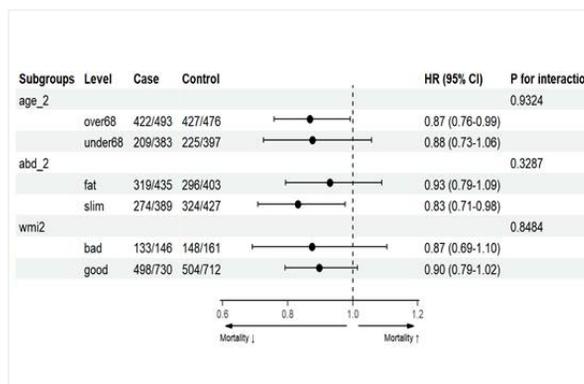
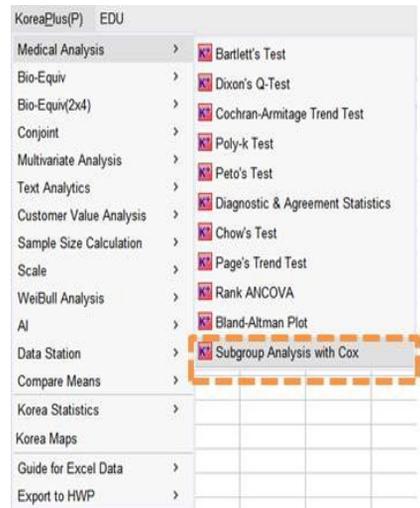
■ 의학에서 주로 사용되는 고급 분석기법을 제공

- ✓ SPSS 제품을 좀 더 편리하게 사용 가능하도록 대화상자 제공.
- ✓ Medical Analysis : 생존 관련 임상 시험에서 많이 사용되는 고급분석 기능 제공.



■ Subgroup Analysis with Cox

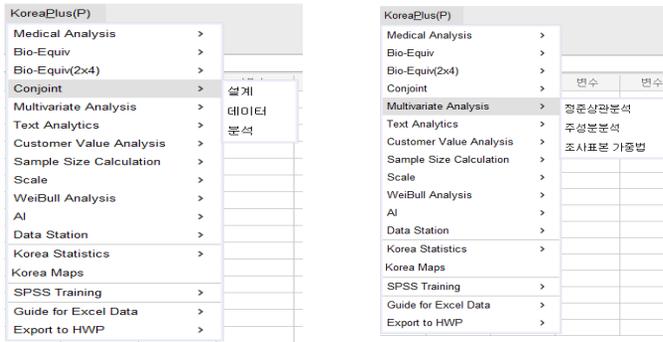
- ✓ 처치에 따른 효과를 서브그룹 별로 분석하는 통계 기법.
- ✓ Cox 비례위험 모형으로 생존 분석 실행.
- ✓ 처치 변수별 K-M 생존 곡선과 K-M 위험 곡선 출력.
- ✓ 서브그룹별 위험비 추정치와 상호작용 유의확률을 한번에 보여주는 Forest Plot 출력.



추가 분석기법 제공 I (※별도 Conjoint 모듈 구매 시, 이용 가능)

Conjoint, Multivariate Analysis 분석 제공

- ✓ 기존에 Syntax로만 구현되었던 분석을 대화상자로 분석할 수 있도록 메뉴 제공
- ✓ Conjoint 분석 : 제품/서비스의 여러 특성 가운데 어떤 특성요인이 소비자에게 영향을 미치는지 파악하는 마케팅 리서치 기법으로 신제품 컨셉 평가, 최적 가격 설정, 시장 세분화에 활용
- ✓ Multivariate Analysis : 주성분분석, 정준상관분석, 조사표본가중법과 같은 다변량 분석.

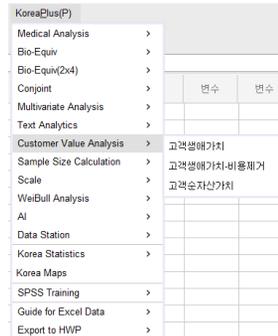


추가 분석기법 제공 II

고객 생애 가치 분석 (Customer Life Value : CLTV) 제공

- ✓ 고객생애가치 : 고객으로부터 미래의 일정기간 얻게 될 이익(=수입-비용)을 할인율에 따라 현재 가치로 환산한 재무적 가치
- ✓ 고객 가치 분석 : 고객 개인별로 전체 생애 동안의 총 기여가치를 측정하는 방법
- ✓ KoreaPlus에서의 고객 생애 가치 분석은 마케팅 비용을 고려하지 않고 단순히 관찰 기간에 나타날 공헌 마진을 현재 가치로 할인하여 측정하는 방법으로 제공
- ✓ 장점 : 측정 자체가 매우 단순하여 활용적인 측면이 강함
- ✓ 개인별 마케팅 비용의 차이가 크지 않은 산업이나 CRM 관점의 마케팅 비용을 측정하기 어려운 기업에서 활용 가능

고객ID	공헌마진	잔존생애 기간	할인율	CLTV
1	15.00	5	1.0	56.86
2	45.00	3	.5	55.95
3	24.00	6	.7	73.17
4	75.00	10	.7	322.59
5	9.00	8	.9	43.21
6	13.50	25	.4	49.02
7	36.00	3	.8	71.62
8	19.50	8	.5	52.02
9	69.00	11	.3	134.45
10	30.00	25	.2	54.46
11	96.00	4	.7	232.99
12	96.00	20	.4	326.92
13	75.00	18	.5	307.55
14	4.50	27	.6	24.94
15	19.50	30	.7	128.68
16	135.00	20	.5	574.67
17	111.00	7	.1	54.04
18	28.50	10	.9	157.61
19	40.50	14	.8	238.68
20	45.00	23	.9	359.77



추가 분석기법 제공 III

Sample Size Calculation과 Weibull Analysis 제공

- ✓ Sample Size Calculation(표본 크기 계산) : 최적의 표본 크기를 산출하는 기법. 이를 통해 통계적 검정력 확보는 물론 표본 산출을 위한 시간 및 비용 절약
- ✓ Weibull Analysis : 가장 적절한 수명분포를 결정하고 모수를 추정하는 접근법. 주로 부품의 수명 측정에 사용되며 질병으로 인한 사람의 수명이나 제품의 제조와 운반에 걸리는 시간, 날씨 예보 등에 활용

KoreaPlus(P)	
Medical Analysis	>
Bio-Equiv	>
Bio-Equiv(2x4)	>
Conjoint	>
Multivariate Analysis	>
Text Analytics	>
Customer Value Analysis	>
Sample Size Calculation	>
Scale	>
Weibull Analysis	>
AI	>
Data Station	>
Korea Statistics	>
Korea Maps	>
SPSS Training	>
Guide for Excel Data	>
Export to HWP	>

KoreaPlus(P)	
Medical Analysis	>
Bio-Equiv	>
Bio-Equiv(2x4)	>
Conjoint	>
Multivariate Analysis	>
Text Analytics	>
Customer Value Analysis	>
Sample Size Calculation	>
Scale	>
Weibull Analysis	>
AI	>
Data Station	>
Korea Statistics	>
Korea Maps	>
SPSS Training	>
Guide for Excel Data	>
Export to HWP	>

추가 분석기법 제공 IV

Scale 분석 기능 제공

- ✓ Kuder-Richardson 21 : 문항 간의 내적 일치도 정도를 나타내어주는 지수.
- ✓ Differential Item Functioning (차별기능문항) : 각 문항이 특정 집단에 편파적으로 적용되는지 판별하는 데 사용되는 기법
- ✓ Item Response Theory (문항반응이론) : 문항에 대한 피험자의 반응과 개인의 잠재 능력 간의 관계를 수리적 함수로 나타낸 측정 이론

KoreaPlus(P)	
Medical Analysis	>
Bio-Equiv	>
Bio-Equiv(2x4)	>
Conjoint	>
Multivariate Analysis	>
Text Analytics	>
Customer Value Analysis	>
Sample Size Calculation	>
Scale	>
Weibull Analysis	>
AI	>
Data Station	>
Korea Statistics	>
Korea Maps	>
SPSS Training	>
Guide for Excel Data	>
Export to HWP	>

Kuder-Richardson 21	2-모수 로지스틱 모형
Item Response Theory	>
Differential Item Functioning	>

Kuder-Richardson 21	로지스틱 회귀분석(LR)
Item Response Theory	>
Differential Item Functioning	>

추가 분석기법 제공 V

PSM(Propensity Score Matching) 기능 제공

- ✓ Felix Thoemmes교수가 개발한 SPSS 확장 번들로 다양한 성향점수 계산 및 매칭 기법을 제공하며, 매칭 결과를 판단할 수 있는 다양한 지표를 제공
- ✓ 별도의 기술지원 없이 사용자가 손쉽게 설치 할 수 있는 설치 파일 제공

The image shows the SPSS PSM (Propensity Score Matching) interface. On the left is the 'PS Matching' menu. In the center is the 'Propensity Score Matching' dialog box with various options like 'Estimation Algorithm' and 'Matching Algorithm'. On the right are four histograms comparing 'Unmatched Treated', 'Matched Treated', 'Unmatched Control', and 'Matched Control' propensity scores. Below these is a data table with columns for ID, Treatment, age, distance, income, ps, and psweight.

ID_1	Treatment_1	age_1	distance_1	income_1	ps_1	psweight_1	ID_c	Treatment_c	age_c	distance_c	income_c	ps_c	psweight_c	
1	1.00	1	24.00	6.00	5.50	350	1.000	24.00	0	22.00	6.00	4.30	0.017	1.000
2	2.00	1	23.00	3.00	6.10	854	28.00	0	22.00	2.00	3.00	2.16	1.000	
3	3.00	1	20.00	1.00	4.50	484	1.000	27.00	0	22.00	3.00	2.40	0.64	1.000
4	4.00	1	22.00	1.00	5.30	870	1.000	19.00	0	22.00	5.00	5.60	2.97	1.000
5	5.00	1	24.00	2.00	2.40	311	1.000	30.00	0	20.00	1.00	1.50	0.38	1.000
6	6.00	1	27.00	3.00	6.30	986	1.000	32.00	0	27.00	2.00	2.20	0.63	1.000
7	7.00	1	26.00	2.00	3.30	781	1.000	31.00	0	24.00	2.00	1.80	1.93	1.000
8	8.00	1	30.00	2.00	6.50	999	1.000	26.00	0	21.00	3.00	6.70	7.81	1.000

추가 분석기법 제공 VI

Meta-analysis 기능 제공

- ✓ 동일한 주제하에 수행된 여러 연구들로부터 도출된 다양한 연구결과를 체계적으로 종합 분석하는 계량적 접근방법.
- ✓ Effect Size, Moderator Analysis, Publication Bias Analysis, Cumulative Meta Analysis, Sensitivity Analysis 의 연속형 데이터, 이분형 데이터, 상관관계 데이터 결과 값 제공

The image shows the SPSS Meta-analysis interface. On the left is the 'Meta Analysis' menu with options for '연속형 데이터', '이분형 데이터', and '상관관계 데이터'. In the center is the '효과크기분석 : 연속형 데이터' dialog box with various options like 'Effect measure' and 'Model'. On the right is a forest plot showing Odds Ratio for five studies (study_1 to study_5) and a pooled effect size of 1.76 [1.10; 2.81].

Study	Experimental Events	Control Events	Total	Odds Ratio	OR	95% CI (W/random)	
study_1	16	65	12	65	1.44	[0.62; 3.35]	15.9%
study_2	10	40	8	40	1.33	[0.46; 3.83]	12.3%
study_3	19	80	14	80	1.57	[0.88; 3.18]	17.4%
study_4	80	400	25	400	3.75	[2.34; 6.02]	24.7%
study_5	11	40	9	40	1.52	[0.54; 4.29]	12.6%
study_5	18	65	16	65	1.17	[0.54; 2.57]	17.1%
Random effects model	690	690			1.76	[1.10; 2.81]	100%

국가통계포털 링크

■ 국가통계포털 링크가 추가

- ✓ 국가 통계 데이터를 이용하기 위한 국가 통계 제공 사이트 바로 가기
- ✓ Korea Statistics 메뉴를 통해 통계 관련 사이트(정부·공공)로 바로 이동 가능.

KoreaPlus(P)	
Medical Analysis	>
Bio-Equiv	>
Bio-Equiv(2x4)	>
Conjoint	>
Multivariate Analysis	>
Text Analytics	>
Customer Value Analysis	>
Sample Size Calculation	>
Scale	>
WeiBull Analysis	>
AI	>
Data Station	>
Korea Statistics	>
Korea Maps	>
SPSS Training	>
Guide for Excel Data	>
Export to HWP	>

01. 국가통계포털
02. 마이크로데이터 서비스시스템(MDIS)
03. 공공데이터 포털
04. 통계지리정보(SGIS)
05. 서울시 통계지도(GIS)
06. 한국은행 경제통계시스템
07. 금융감독원 - 금융통계정보시스템
08. 국가교통DB센터
09. 국가에너지통계종합정보시스템
10. 고용노동통계
11. 보건복지데이터포털
12. ITSTAT (지식경제부 통계포털)
13. ISTATS (산업연구원 산업통계 포털)
14. 서울시 열린 데이터 광장

한국지도 제공

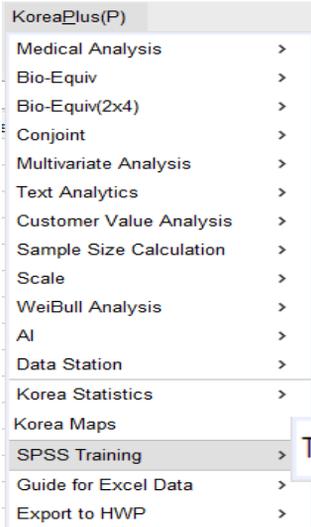
■ KoreaPlus Statistics Maps에 한국 지도 템플릿 추가

- ✓ Korea Maps 메뉴 클릭만으로 한국지도 파일생성
- ✓ 사용자가 직접 한국 지도를 다운받지 않아도 통계지리정보사이트에서 제공하는 다양한 한국 지도 활용 가능

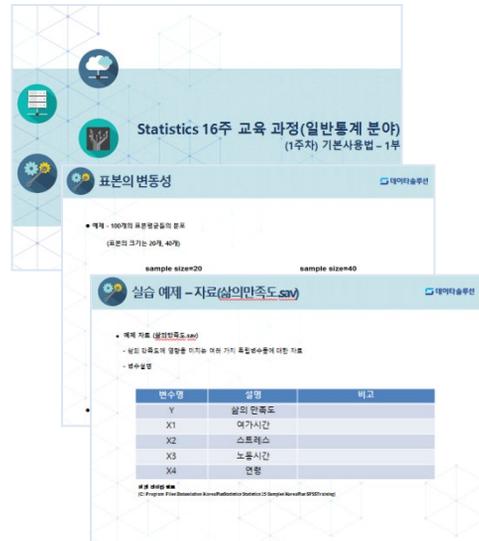
KoreaPlus(P)	
Medical Analysis	>
Bio-Equiv	>
Bio-Equiv(2x4)	>
Conjoint	>
Multivariate Analysis	>
Text Analytics	>
Customer Value Analysis	>
Sample Size Calculation	>
Scale	>
WeiBull Analysis	>
AI	>
Korea Statistics	>
Korea Maps	>

SPSS Training

- 실습 자료 및 샘플데이터 제공
 - 초보자를 위한 16주 과정 SPSS 활용 실습 파일, 샘플 데이터 제공
 - 8주차, 16주차에 Test 및 답안 제공



Training 바로가기

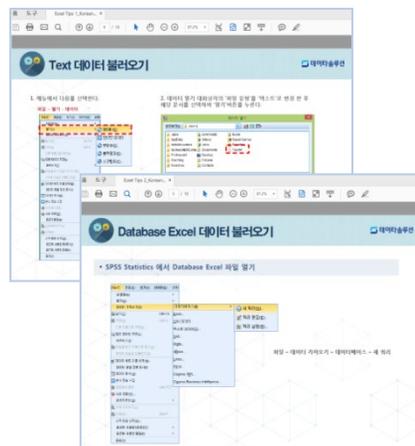


엑셀 데이터 불러오기 Tips

- 엑셀 데이터 불러오기 도움말 추가
 - 사용자가 쉽게 엑셀파일을 활용할 수 있는 Tip 제공
 - 분석 대상 데이터로 많이 쓰여지는 엑셀 데이터를 편리하고 빠르게 활용할 수 있도록 [엑셀 파일 불러오기]와 [데이터베이스에 연결된 엑셀 파일 불러오기]에 대한 도움말 제공

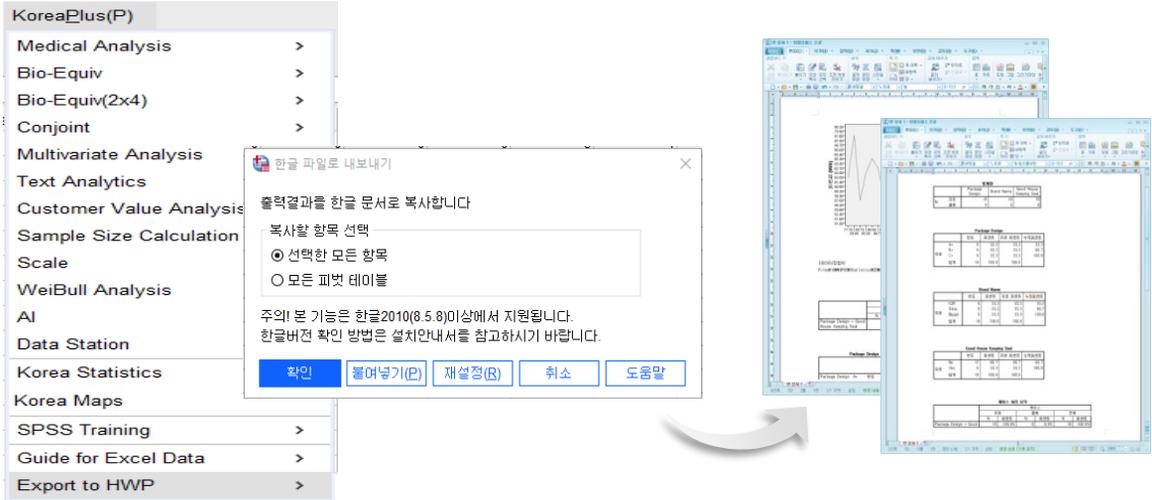


Importing Excel Data 바로가기



아래한글 내보내기

- 출력결과(Output)를 아래한글(hwp)로 바로 내보내기 가능
 - ✓ [한글 파일 내보내기] 기능으로 엑셀 및 워드 파일을 포함한 강력한 보고서 작성 가능
 - ✓ 결과값을 바로 한글 파일로 내보내어 보고서 작성 시간을 단축



기타혜택

- SW material 제공
- 한글 매뉴얼 제공
- KoreaPlus Statistics 무료교육



데이터솔루션

KoreaPlus Statistics (SPSS 29) > ...

이름	소유자
Administration	데이터솔루션
Documentation	데이터솔루션
InstallationGuide	데이터솔루션
KoreaPlusStatistics	데이터솔루션
Sample	데이터솔루션

데이터솔루션

KoreaPlus Statistics 문서파일 > ...

이름	소유자
SPSS Statistics 29	데이터솔루션
SPSS Statistics 29 사용자 가이드	데이터솔루션
SPSS Statistics 29 설치 및 업그레이드 가이드	데이터솔루션
SPSS Statistics 29 사용자 가이드	데이터솔루션
SPSS Statistics 29 사용자 가이드	데이터솔루션
SPSS Statistics 29 사용자 가이드	데이터솔루션
SPSS Statistics 29 사용자 가이드	데이터솔루션
SPSS Statistics 29 사용자 가이드	데이터솔루션
SPSS Statistics 29 사용자 가이드	데이터솔루션
SPSS Statistics 29 사용자 가이드	데이터솔루션



KoreaPlus Statistics : 1편
Text Analytics

무료
수강신청



KoreaPlus Statistics : 2편
PSM

무료
수강신청

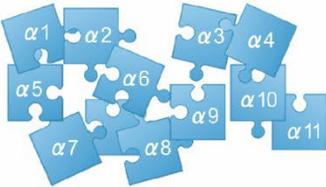
KoreaPlus Statistics의 확장 유틸리티 솔루션

KoreaPlus Statistics Utility



KoreaPlus Statistics Utility는
2023년 4월 GS인증(Good Software) 1등급을 취득하였습니다!

데이터솔루션의 전문 기술팀이 개발한 KoreaPlus Statistics Utility는 KoreaPlus Statistics의 확장 유틸리티 솔루션입니다. KoreaPlus Statistics Utility는 그래픽 기반 사용자 환경(GUI) 구현과 위젯을 통해 데이터 수집을 빠르게 실행할 수 있도록 지원합니다. 주요 특징으로는 캘린더 및 기상 데이터를 수집하고 분석할 수 있는 '데이터 스테이션(Data Station)', 분석가들이 스스로 데이터 활용 역량을 점검할 수 있는 'DQ(Data Quotient)' 서비스, KoreaPlus 제품군의 라이선스를 갱신, 관리할 수 있는 기능과 데이터솔루션의 빅데이터러닝센터 교육 과정을 수강할 수 있는 혜택 및 학습자료 등을 제공합니다. KoreaPlus Statistics Utility는 KoreaPlus Statistics 제품 구매 시 함께 제공됩니다.

<p>데이터솔루션 고유 부가기능 (Value Add)</p>		<ul style="list-style-type: none"> · 메타분석 / 고객가치분석 · 한글 내보내기 한국지도 · AMOS 아웃풋 내보내기 · Text분석 등 최신 기법
<p>편의 기능</p>		<ul style="list-style-type: none"> · DataStation(Calendar, Weather Data 수집) · DQ(데이터 역량진단 서비스) · Edu Training · KoreaPlus Statistics 라이선스 관리 · 국가 통계포털 링크
<p>KoreaPlus Statistics Utility</p>		



IBM에서 소개한 KoreaPlus Statistics

SPSS Statistics에 Value Add Component가 추가 된 KoreaPlus Statistics를 IBM Partner Plus Directory에서 소개하고 있습니다.

웹사이트 주소 : <https://www.ibm.com/partnerplus/directory/solution/0036>

IBM Partner Plus Directory / Solutions /

KoreaPlus Statistics

KoreaPlus Statistics was developed as advanced statistical analysis tool based on IBM SPSS statistics, which is optimized for Korean users.

Provided by



Datasolution, Inc.
IBM Platinum Partner

View Partner → Contact Partner

Overview

Benefits

Overview

KoreaPlus Statistics provides an easy-to-use GUI, specialized expertise with advanced statistical analysis functions which are not provided in SPSS, analytical solutions for optimized functions in the field from test design to hypothesis verification, statistics after analysis, and visualization. Also, it provides appropriate analytical methods and statistics according to the data characteristics of the industry.



Company Overview

KoreaPlus Statistics provides an easy-to-use GUI, specialized expertise with advanced statistical analysis functions which are not provided in SPSS, analytical solutions for optimized functions in the field from test design to hypothesis verification, statistics after analysis, and visualization. Also, it provides appropriate analytical methods and statistics according to the data characteristics of the industry.

KoreaPlus Statistics

소개영상 바로보기



전문 기술지원 서비스 제공

■ 데이터솔루션 기술지원 역량

- ✓ 1990년부터 대한민국에서 SPSS 공급 및 서비스 제공
- ✓ 30여년간 축적된 기술지원 DB 보유
- ✓ 원활한 기술지원을 위한 전담 기술지원팀(Technical Support)을 보유
- ✓ 다양한 기술지원 채널(전화, 이메일, 각종 홈페이지, 블로그)을 보유
- ✓ 국내 유일의 SPSS 제품 기술지원
 - ① 최초 설치 지원
 - ② 재인증, 재설치 지원
 - ③ 인터넷 망분리 설치 지원
 - ④ 각종 오류 지원
 - ⑤ Patch 제공

번호	이슈 ID	이슈 수	분류	상태	보고자	이슈 발생	보고 날짜	입력여부	요청
1	0004473	2	Statistics	해결된 이슈 (완료)	정인진	이동진	2017-01-19	2017-01-19	비밀 오류 문제
2	0004472	1	Statistics	해결된 이슈 (완료)	이동진	이동진	2017-01-18	2017-01-18	Statistics 구동시 특정 프로그램 작동이 안되는 현상
3	0004471	2	Statistics	발생 (완료)	이동진	이동진	2016-12-22	2017-01-18	SPSS Statistics 23ver. 출력결과값 내보내기 오류
4	0004470	2	Statistics	해결된 이슈 (완료)	정인진	이동진	2017-01-18	2017-01-18	내보내기 라이선스 인증

전문적인 기술지원 DB 보유

1	0004468	1	Amos	해결된 이슈 (완료)	신영진	이동진	2017-01-17	2017-01-17	라이선스 인증 오류
2	0004466	1	Statistics	해결된 이슈 (완료)	홍정희	이동진	2017-01-17	2017-01-17	내보내기 오류
3	0004465	1	Statistics	해결된 이슈 (완료)	변정희	이동진	2017-01-17	2017-01-17	설치지원
4	0004464	2	Statistics	해결된 이슈 (완료)	이동진	이동진	2017-01-16	2017-01-16	내보내기 라이선스 갱신 요청
5	0004460	1	Amos	해결된 이슈 (완료)	이동진	이동진	2017-01-16	2017-01-16	amos 설치 오류
6	0004457	1	Statistics	해결된 이슈 (완료)	이동진	이동진	2017-01-13	2017-01-14	라이선스 매니저 프로그램 재설치 요청 미 건
7	0004454	1	Amos	해결된 이슈 (완료)	고인홍	이동진	2017-01-12	2017-01-13	설치 및 인증
8	0004454	1	Statistics	해결된 이슈 (완료)	정인진	이동진	2017-01-13	2017-01-13	인증코드 획득 관련
9	0004452	2	Amos	발생 (완료)	이동진	이동진	2017-01-12	2017-01-13	부호소문자형 결과 오류

■ 장애 발생 시

- ✓ 1차: 유선 기술지원
- ✓ 2차: 원격 기술지원
- ✓ 3차: 방문 기술지원

빅데이터 러닝센터 전문 교육 프로그램 제공

		SPSS	Amos	오픈소스R	개설예정
코스	학술연구자 과정	구조방정식모형 분석 과정		메타분석 과정	
기초 과정	SPSS 기초통계분석	김계수 교수의 Amos구조 방정식 모델링과 분석		메타분석의 이해와 활용: 초급	
	리서치와 통계분석 I (SPSS 기초통계편)	우종필 교수의 Amos구조 방정식 모형 분석			
	3일 만에 익히는 R 기초통계분석	석박사학위 및 학술논문 작성 중심의 Amos 구조방정식 모델링			
중·고급 과정	SPSS 범주형 자료분석	김계수 교수의 Amos를 활용한 종단자료분석		메타분석의 이해와 활용: 중급	
	SPSS 중급통계분석1: 분산분석과 선형모형	우종필 교수의 Amos에서 통제변수, 매개변수, 조절변수의 사용		메타분석 고급1: 메타분석 논문 작성법	
	SPSS 중급통계분석2: 회귀분석			메타분석 고급2: 다변량 및 다중메타분석	
	SPSS 다변량 분석	우종필 교수의 Amos 잠재성장곡선 모델		메타분석 고급3: 네트워크 메타분석 및 메타구조 방정식	
	리서치와 통계분석 II (SPSS 중급통계편)	우종필 교수의 Amos 다중집단분석			
	2일 만에 익히는 R 중급통계분석				
실전·응용 과정	SPSS를 활용한 Kaggle 활용		SPSS 설문지 설계 및 분석 과정		
	SCI, SSCI 논문 게재를 위한 효과적인 투고 요령 및 해결책		SPSS를 활용한 논문을 위한 통계분석		
의학 보건학	SPSS 의학보건학 심화과정 1 : 일반선형모형		SPSS 의학보건학 심화과정 4 : 생존분석		
	SPSS 의학보건학 심화과정 2 : 일반화선형모형		SPSS 의학보건학 심화과정 5 : 결측자료분석		
	SPSS 의학보건학 심화과정 3 : 반복측정자료분석		SPSS-R 연동기능을 이용한 성향점수(PSM) 분석방법		

교육 및 정보제공 서비스

- 사용자 교육을 위한 자체적인 교육장, 실습 PC 보유
- 전문적이고 체계적인 위탁교육 프로그램 제공
- 계약 체결 시 정기/비정기 교육쿠폰 무상제공
- 빅데이터러닝센터 정기/비정기 교육 20% 할인
- 데이터솔루션 세미나 및 Promotion 우선 참여기회 부여



빅데이터러닝센터

DataSolution User Portal

- 데이터솔루션은 SW관련 다양한 기술지원 및 교육, 세미나 지원 등 사용자 편의를 위한 SPSS User Portal을 운영하고 있습니다.
- KoreaPlus Statistics 제품을 구매하신 고객은 자유롭게 이용이 가능합니다.

웹사이트 주소 : <http://spss.datasolution.kr>



제품

KoreaPlus Statistics >

- Embedded on IBM SPSS Statistics
- Embedded on IBM SPSS Amos
- Embedded on IBM SPSS Modeler

IBM SPSS

IBM SW

Biz SW

KoreaPlus Statistics (Embedded on IBM SPSS Statistics)



KoreaPlus Statistics (Embedded on IBM SPSS Statistics)

다양한 분야에서 활용되고 있는 SPSS 제품 고급분석과정을 대화상자(GUI)로 구현한 확장 모듈을 통해 더 편리하고 신속하게 분석을 수행할 수 있습니다.

[바로구매](#)

개요	특장점	제품	모듈&유틸리티	KoreaPlus Utility
--------------------	---------------------	--------------------	-----------------------------	-----------------------------------

KoreaPlus Statistics는 교육기관, 공공기관, 일반기업(금융, 제조, 컨설팅) 등의 다양한 산업군에서 많이 사용되는 최적화된 통계분석 패키지입니다. KoreaPlus Statistics(Embedded on IBM SPSS)는 SPSS 프로그램 자체에서 제공하는 모든 기능 외에 국내 사용자들에게 유용하게 쓰이는 기능들을 추가로 제공하고 있으며, 이러한 Value Add된 KoreaPlus기능으로는 출력결과를 한글파일로 내보내기, 메타분석, PSM, AI 분석기법, TA(Text Analytics) 기능 등이 있습니다.