



## 주요 특징

- 이 직관적인 시각적 데이터 마이닝 워크벤치에서 구조적 데이터를 손쉽게 액세스, 준비하고 모델링합니다.
- Cognos Business Intelligence 소프트웨어 및 InfoSphere™ Warehouse와의 통합을 통해 비즈니스 분석의 이점을 더 폭넓게 누립니다.
- IBM Classic Federation Server 및 zDB2 지원으로 레거시 시스템에 저장된 데이터를 분석합니다.
- 가장 앞선 통계 및 기계 학습 기술로 신속하게 모델을 작성하고 검증합니다.
- 통찰 및 예측 모델을 효율적으로 배치합니다.
- zLinux, SuSE Linux® Enterprise Server를 통해 그리고 IBM Smart Analytics System for Power에 포함되는 형태로 다양한 배치 옵션을 제공합니다.

## 예측 인텔리전스로 더 우수한 의사 결정

데이터 마이닝을 통해 조직은 현재의 상황을 더 정확하게 파악하고 미래의 상황을 더 심도 있게 예측할 수 있습니다.

IBM SPSS Modeler Professional\*은 구조적 수치 데이터를 분석하여 결과를 모델링하고 비즈니스 의사 결정에 예측 인텔리전스를 제공하는 예측 자료를 생성하는 데이터 마이닝 워크벤치입니다. 예측 인텔리전스를 통해 더 효과적인 전략을 마련할 수 있습니다. 벤치마킹 결과, 계획, 실적과 같은 추세를 평가할 뿐 아니라 미래를 내다보는 것도 가능합니다. 즉 가능한 결과를 평가하고 요인 간 상호 작용이 그 결과에 어떤 영향을 미치는지 파악할 수 있습니다.

IBM SPSS Modeler Professional 최신 버전에서는 IBM Cognos 8 Business Intelligence 및 InfoSphere Warehouse 솔루션과의 통합을 더욱 강화했습니다. 따라서 사용자는 단일 벤더에서 공급하는 최고의 솔루션으로 데이터 저장 및 관리부터 의사 결정권자를 위한 예측 인텔리전스 배치까지 모두 해결할 수 있습니다. 이제 Cognos 애플리케이션을 기반으로 주요 지표를 모니터링하는 전 세계 수천여 개 기업들이 리포트 및 대시보드에 예측 인텔리전스를 결합하여 더 통찰력 있게 미래를 예측하고 새로운 기회를 심분 활용할 수 있게 되었습니다.

이제 IBM SPSS Modeler Professional 소프트웨어를 메인프레임 컴퓨팅 환경에서 호스팅하는 것이 가능하므로, 예측 인텔리전스의 범위가 더욱 확대되었습니다. 이 향상된 기능에 대한 자세한 내용은 2 페이지를 참조하십시오.

### 데이터 마이닝 프로세스 능률화

IBM SPSS Modeler는 전 세계의 많은 분석가 및 비즈니스 사용자들이 애용하는 솔루션입니다. 자동화된 데이터 준비 및 모델링 기능을 제공하므로, 고급 분석 기술이 없는 사용자도 정확한 모델을 손쉽게 그리고 신속하게 만들 수 있습니다. 또한 전문 분석가는 이 소프트웨어의 고급 데이터 준비 및 예측 모델링 기능을 심분 활용할 수 있습니다.



\* IBM SPSS Modeler Professional의 이전 명칭은 PASW Modeler입니다.

---

## 비즈니스 이점:

어떤 유형의 기업이든 IBM SPSS Modeler Professional의 이점을 누릴 수 있습니다.

- 더 효과적인 비용으로 고객을 유치하고 고객의 충성도를 강화하며 고객 이탈을 최소화하고 위험도 줄일 수 있습니다.
  - 공공 기관에서는 인력 수용력을 예측하고 프로그램의 실효성을 평가하며 공공 안전 문제에 미리 대처할 수 있습니다.
  - 교육 기관에서는 학생 라이프사이클을 관리하고 학업 성과를 높이며 다른 여러 운영상의 문제를 해결할 수 있습니다.
- 

IBM SPSS Modeler의 직관적인 그래픽 인터페이스에서는 데이터 마이닝 프로세스의 각 단계를 하나의 “스트림”에 포함된 형태로 손쉽게 확인할 수 있습니다. 분석가와 비즈니스 사용자는 이러한 스트림을 대화식으로 다루면서 협업을 통해 데이터 마이닝 프로세스에 비즈니스 지식을 추가할 수 있습니다. 데이터 마이닝 담당자는 코드 작성과 같은 기술적인 일보다 지식 탐색에 주력하면서 “사고의 맥락(train-of-thought)” 분석을 추구하고 더 심도 있게 데이터를 탐색하며 숨겨진 관계를 추가로 밝혀낼 수 있습니다.

이 시각적 인터페이스에서 IBM SPSS Data Collection\*제품을 비롯한 다양한 소스의 데이터 그리고 IBM SPSS Statistics, SAS 및 Microsoft® Excel 파일 등 모든 데이터베이스, 스프레드시트 또는 플랫폼 파일 유형의 데이터에 손쉽게 액세스하고 이를 통합할 수 있습니다. 그리고 앞서 언급한 대로 Cognos 8 Business Intelligence에서 곧바로 데이터에 액세스하는 것이 가능합니다.

다른 어떤 데이터 마이닝 솔루션도 이와 같은 다양한 기능을 제공하지 않습니다. IBM SPSS Modeler의 강력한 자동화 도구(자동 데이터 준비 및 자동모델링 등)를 통해 분석할 데이터를 손쉽게 준비하고, 데이터에 숨겨진 패턴에 따라 최상의 모델을 찾으며, 일관성 있고 정확한 결과를 신속하게 생성할 수 있습니다.

## Modeler Professional 14.1의 새로운 기능

이번 Modeler Professional 버전에서는 예측 인텔리전스를 더 다양한 조직에서 적용할 수 있도록 기능이 강화되었습니다. 즉 이 대상 조직들은 계획 및 일상적인 의사 결정 과정에서 더 우수한 집중력과 민첩성을 발휘하게 됩니다. 해당 기업, 업무 환경, 고객 및 기타 이해 관계자를 더 통찰력 있게 이해할 수 있기 때문입니다. 다음과 같은 기능을 새롭게 제공합니다.

- **Cognos와의 통합 연계성.** 이제 분석가는 Modeler 인터페이스에서 곧바로 Cognos 8 Business Intelligence 환경의 데이터에 액세스할 수 있습니다. Cognos 소프트웨어에서는 전사적 차원의 의사 결정을 지원할 완전하고 일관성 있는 정보 뷰를 작성하여 제공합니다. Modeler의 분석 기능이 추가된 덕분에 해당 기업에서 자주 사용하는 전사적 데이터 뷰를 통해 특정 결과의 확률을 신속하고 정확하게 평가할 수 있습니다. 그리고 Modeler에서 Cognos 8 Business Intelligence로 결과를 기록할 수 있으므로, Cognos를 기업 분석용 정보 포털로 이용하는 비즈니스 사용자 및 모든 정보 이해 관계자에게 예측 인텔리전스를 전달할 수 있습니다.

\*\* IBM SPSS Data Collection 및 IBM SPSS Statistics의 이전 명칭은 PASWData Collection 및 PASW Statistics입니다.

- **향상된 InfoSphere 및 DB2 통합.** InfoSphere Data Warehouse 를 사용하는 곳에서는 이제 데이터 마이닝 알고리즘의 전 범위를 Modeler 내부에서 사용할 수 있습니다. 따라서 사용자는 데이터 마이닝 성능 저하 없이 직관적인 그래픽 인터페이스를 이용할 수 있다는 장점이 있습니다. 지원되는 알고리즘에는 로지스틱 회귀 (logistic regression), Naïve Bayes, 시계열(time-series), RBF(radial basis function)가 추가되었습니다. 압축, 파티셔닝과 같은 추가적인 DB2 액세스 옵션으로 대규모 데이터 소스도 손쉽게, 효과적으로 활용할 수 있습니다.
- **메인프레임 데이터 마이닝.** 일반적으로 메인프레임 컴퓨터는 기업 경영 이력에 관한 방대한 데이터를 저장하고 있습니다. Modeler 에서 zDB2 및 IBM Classic Federation 서버를 지원하므로 과거의 사건이 미래에 어떤 영향을 미칠 것인지 파악하는 게 가능합니다. 따라서 현재의 활동을 더 명확하게 이해하고, 변화하는 비즈니스 현실을 인식하며, 믿음만한 예측 인텔리전스를 바탕으로 계획을 세울 수 있습니다.
- **System z 기반 Linux 지원.** 복잡한 정보 시스템을 간소화하면서 동시에 오늘날의 까다로운 보안, 투명성 및 비용 통제 요구 사항을 이행하려는 기업이라면 System z 에서 Linux를 지원하는 방법으로 많은 이점을 누릴 수 있습니다. IBM SPSS Modeler Professional Server는 이제 그러한 환경에서도 지원됩니다.

## 더 다양한 데이터 유형을 포함하여 더 우수한 결과 제공

고객들은 사용 가능한 모든 데이터 유형을 포함시키면 예측 모델의 정확도를 “끌어올려” 더 유용한 권장 내용과 더 나은 결과를 얻을 수 있다는 것을 알게 되었습니다.

대량의 텍스트 데이터를 수집하는 기업에서는 IBM SPSS Modeler Premium의 대화형 텍스트 마이닝 워크벤치에서 어떤 텍스트 유형으로부터도 개념과 의견을 추출할 수 있습니다. 즉 운영 소스에서 수집한 텍스트, 콜 센터 노트, 고객 이메일, 미디어 또는 저널 기사, 블로그, RSS 피드 등 다양한 데이터 유형이 지원됩니다. IBM SPSS Data Collection 제품의 설문 데이터에 직접 액세스하여 인구 통계, 태도 및 행동 정보를 손쉽게 모델에 포함함으로써 서비스 대상 집단 또는 조직에 대한 완전한 이해 기반을 구축할 수 있습니다.

## 가장 폭넓은 기술 범위에서 선택

IBM SPSS Modeler는 다음 알고리즘을 비롯하여 어떤 데이터 마이닝 애플리케이션의 필요 사항도 해결할 수 있도록 광범위한 고급 데이터 마이닝 기술을 제공합니다.

- 분류 알고리즘 - 의사결정나무, 신경망, 로지스틱 회귀, 시계열, SVM(Support Vector Machines), Cox 회귀 등의 기법을 적용해 이력 데이터를 바탕으로 예측합니다. 이진 및 수치 결과 모두에 대해 자동 분류 모델링을 적용할 수 있어 효율적으로 모델을 생성합니다.

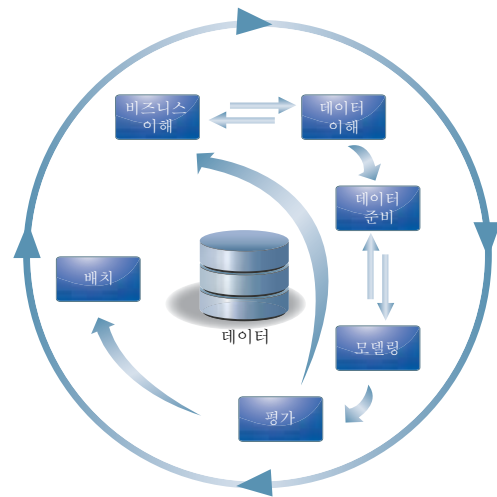
- 세분화(Segmentation) 알고리즘 - 자동 클러스터링, 변칙 탐지 및 클러스터링 신경망 기법을 적용해 사람을 그룹화하거나 이상 패턴을 탐지합니다. 자동 분류를 통해 하나의 단계로 여러 알고리즘을 적용합니다. 그리고 어림짐작 없이 확실하게 최적의 기법을 선택합니다.
- 연상(Association) 알고리즘 - Apriori, CARMA 및 순차적 연상을 적용하여 연관성, 링크 또는 순서를 찾아냅니다.

### 현재의 정보 및 기술 최적화

Modeler의 개방적이고 확장 가능한 아키텍처에서는 기존의 IT 인프라를 최대한 활용합니다. 데이터에 액세스할 때 그리고 결과를 배치할 때 기존 시스템과 통합되므로, 전용 데이터 형식과 다른 형식 사이를 오갈 필요가 없습니다. 그리고 인-데이터베이스(in-database) 마이닝, 멀티스레딩, 서버 클러스터링, SQL 푸시백과 같은 기법을 적용하여 자원을 절약하고 더 신속하게 결과를 전달하며 전반적인 IT 비용을 줄일 수 있습니다.

### 검증되고 반복 가능한 프로세스 적용

Modeler는 데이터 마이닝 프로세스의 각 단계에서 사실상의 산업 표준인 CRISP-DM(Cross-Industry Standard Process for Data Mining)을 지원합니다. 따라서 고객은 프로젝트마다 새 프로세스를 다시 마련할 필요 없이 데이터 마이닝을 통한 비즈니스 문제 해결에 주력할 수 있습니다. CRISP-DM 프로젝트 관리자를 통해 개별 Modeler 프로젝트를 효율적으로 체계화할 수 있습니다.



이 그림에서 소개하는 CRISP-DM 프로세스에서는 데이터 마이닝 담당자가 효율적인 데이터 마이닝 프로젝트를 수행하여 측정 가능한 비즈니스 결과를 내놓을 수 있습니다.

## 전사적 범위에서 예측 모델링 배치

IBM SPSS Modeler에서는 일반적인 중소기업 환경에서 생성되는 양의 데이터를 효율적으로 분석할 수 있습니다. 많은 양의 데이터 또는 복잡한 데이터에 대한 데이터 마이닝이 필요한 경우 IBM SPSS Modeler Server를 사용할 수 있습니다. 클라이언트/서버 아키텍처를 적용하는 Modeler Server에서는 여러 명의 분석가가 컴퓨팅 자원에 대한 부담 없이 동시에 작업할 수 있습니다. 대표적인 정보 플랫폼에서 인-데이터베이스 마이닝을 수행하면서 많은 양의 데이터도 효율적으로 처리할 수 있습니다. Modeler Server에서는 데이터 마이닝의 이점을 지리적 또는 기능적 경계 너머로 확장하고 의사 결정권자에게 신속하게 결과를 전달할 수 있도록 추가적인 배치 옵션도 제공합니다.

## IBM Business Analytics 정보

IBM Business Analytics 소프트웨어는 의사 결정권자가 비즈니스 성과를 개선하는 데 도움이 되는 완벽하고 일관되며 정확한 정보를 제공합니다. 비즈니스 인텔리전스, 예측 분석, 재무 실적 및 전략 관리 그리고 분석 애플리케이션으로 구성된 포괄적인 포트폴리오를 통해 현재의 성과에 대해 명확하고 즉각적이며 실천 가능한 통찰을 얻고 향후의 결과를 예측할 수 있습니다. 어떤 규모의 기업에서든 산업 전문 솔루션, 검증된 사례 및 전문 서비스와 연계함으로써 업무 생산성을 극대화하고 의사 결정을 안전하게 자동화하며 더 나은 결과를 거둘 수 있습니다.

이 포트폴리오에 포함된 IBM SPSS Predictive Analytics 소프트웨어는 향후 상황을 예측하고 그 통찰을 바탕으로 능동적인 사전 조치를 취해 더 우수한 비즈니스 성과를 거둘 수 있도록 지원합니다. 전 세계의 기업, 정부/공공 기관 및 교육 기관에서 IBM SPSS 기술을 활용하여 고객을 유치, 유지, 증대하고 부정 행위 및 위험 요소를 최소화하면서 경쟁력을 강화하고 있습니다. 일상 업무에서 IBM SPSS 소프트웨어를 활용한다면 예측형 기업으로 거듭날 수 있습니다. 즉 비즈니스 목표 달성을 위해 의사 결정의 방향을 정하고 이를 자동화하며 측정 가능한 경쟁 우위를 달성할 수 있습니다. 자세한 정보를 얻거나 담당자와 연락하시려면 다음 사이트를 참조하십시오.

[ibm.com/spss](http://ibm.com/spss)

## 기능 데이터 이해

- 자동 지원 기능으로 다양한 대화형 그래프 작성
- 시각적 링크 분석을 통해 데이터 내 연관성 파악
- 그래프에서 지역 또는 항목을 선택하고 해당 정보를 보는 대화식 방법으로 데이터 사용 또는 분석에 사용할 핵심 데이터 선택
- Modeler에서 곧바로 Statistics 그래프 및 리포팅 도구에 액세스

## 데이터 준비

- Cognos Business Intelligence, IBM DB2, Oracle, Microsoft SQL Server™, Informix, Neoview, Netezza, mySQL (Sun) 및 Teradata 데이터 소스 그리고 메인프레임 데이터(zDB2 및 IBM Classic Federation Server 지원)로부터의 운영 데이터에 액세스
- 구분형(delimited) 및 고정폭 텍스트 파일, Statistics 파일, SAS, Data Collection 데이터 소스 또는 XML 가져오기
- Modeler의 다양한 데이터 클리닝 옵션 - 잘못된 데이터 제거 또는 대체하거나, 자동으로 결측치를 대체하고 이상치 또는 극단적인 데이터를 완화
- 자동 데이터 준비 적용 - 분석할 데이터 질의 및 조건을 하나의 단계로 해결
- 구분형 텍스트 파일, Excel, Statistics, SAS, Cognos Business Intelligence 패키지 및 운영 데이터베이스로 데이터 내보내기

- 구분형 텍스트 파일, Excel, Statistics, SAS 및 운영 데이터베이스로 데이터 내보내기
- 필드 필터링, 이름 지정, 파생, 비닝(binning), 재카테고리화, 값 대체 및 필드 순서 재지정
- 레코드 선택, 샘플링, 병합 및 연결, 정렬, 집합 및 조정
- 데이터 재구성, 파티셔닝 및 전치
- 다양한 문자열 기능: 문자열 생성, 대체, 검색, 일치, 화이트스페이스 제거, 잘라내기
- Modeler에서 곧바로 Statistics의 데이터 관리 및 변환 기능에 액세스
- RFM 스코어링: 고객 트랜잭션을 통합하여 Recency, Frequency 및 Monetary 값 스코어를 제공하고 이들을 조합하여 전체 RFM 분석 실시

## 모델링 및 평가

- 고급 데이터 마이닝 알고리즘으로 데이터에서 최상의 결과 도출
- 대화형 모델 및 수식 브라우저 사용, 고급 통계 출력 조회
- 예측 결과에 대한 데이터 속성의 상대적 영향 표시 - 변수 중요도 그래프
- 여러 모델 조합(양상블 모델) 또는 하나의 모델을 사용하여 제2 모델 분석
- 개별 알고리즘을 선택하지 않고 자동 분류(이진 및 수치) 및 자동 클러스터링 사용
- Modeler의 CLEF (Component-Level Extension Framework)를 사용하여 사용자 정의 알고리즘 통합
- Statistics와의 통합으로 R을 사용하여 분석 옵션 확장

## 모델링 알고리즘 포함

- C&RT, C5.0, CHAID & QUEST - 의사결정나무 알고리즘(대화형 트리 작성 포함)
- 의사결정 리스트 - 대화형 규칙 작성 알고리즘
- K-Means, Kohonen, Two Step, Discriminant, SVM(Support Vector Machine) - 클러스터링 및 세분화 알고리즘
- Factor/PCA, Feature Selection - 데이터 차원감소(DataReduction) 알고리즘
- Regression, Linear, GenLin (GLM) - 선형 수식 모델링
- SLRM(Self-learning Response Model) - 점진적 학습이 포함된 베이지안(Bayesian) 모델
- 시계열 - 시계열 예측 모델 생성 및 자동 선택
- 신경망 - back-propagation learning 및 RBF (Radial Basis Function) network 기법을 이용한 Multi-layer perceptron 알고리즘
- SVM - 넓은 데이터 집합을 다루는 고급 알고리즘
- 베이지안 네트워크 - 시각적 확률 모델
- Cox 회귀 - 어떤 사건이 일어날 때까지의 가능한 시간 계산
- 이상 탐지 - 클러스터 기반 알고리즘으로 비정상적인 결과 탐지
- KNN - Nearest neighbor 모델링 및 스코어링 알고리즘
- Apriori - 널리 쓰이는 연관성 검색 알고리즘, 고급 평가 기능 포함
- CARMA - 복수 후항을 지원하는 연관성 알고리즘
- Sequence - 순서가 중요한 분석을 위한 순차적 연관성 알고리즘
- IBM InfoSphere 인-데이터베이스 마이닝 알고리즘 지원: 의사결정나무,

- 연관성, Sequence, 회귀, 로지스틱 회귀, 클러스터링, Naïve Bayes, 시계열 및 RBF
- Microsoft SQL Server 인-데이터베이스 마이닝 알고리즘 지원: 의사결정나무, 연관성 규칙, 선형 회귀, 클러스터링, 시퀀스 클러스터링, Naïve Bayes, 시계열 및 신경망
  - Oracle 인-데이터베이스 마이닝 알고리즘 지원: 의사결정나무, GLM, O-Cluster (Orthogonal Partitioning Clustering), KMeans, Apriori, MDL (Minimum Description Length), SVM, Naïve Bayes, Adaptive Bayes, NMF (Non-Negative Matrix Factorization) 및 AI (Artificial Intelligence)

### 배치

- SQL 또는 PMML을 사용하는 모델 내보내기(예측 모델에 대한 XML 기반 표준 형식)
- 혁신적인 분석 관리, 프로세스 자동화 및 배치 기능을 제공하는 IBM SPSS 협업 및 배치 서비스 활용

### Modeler 서버(옵션)

- 인-데이터베이스 마이닝 사용 - 대표적인 데이터베이스 기술을 이용해 데이터베이스에 모델을 생성하고 고성능 데이터베이스 구현 활용
- SQL-pushback을 이용하여 데이터 변환 적용, 운영 데이터베이스에 대해 곧바로 모델링 알고리즘 선택
- IBM System z 시스템을 비롯한 고성능 하드웨어 활용 - 문제 해결 시간 단축, 스트림 및 여러 모델의 병렬 실행으로 ROI 향상
- Modeler 클라이언트와 Modeler 서버 간에 SSL 암호화를 통해 기밀 정보를 안전하게 전송



---

**IBM Corporation**

135-270 서울특별시 강남구 도곡동 467-12  
군인공제회관빌딩 한국 아이비엠 주식회사 고객만족센터

IBM 홈 페이지는 다음 URL에서 확인할 수 있습니다.

**ibm.com**

IBM, IBM 로고, ibm.com, Cognos, InfoSphere 및 WebSphere는 미국 또는 기타 국가에서 사용되는 International Business Machines Corporation의 상표입니다. 이러한 상표 및 기타 IBM 상표가 상표 기호(® 또는 ™)와 함께 이 정보에서 처음 표시되어 있는 경우 이 기호는 이 정보가 출판되었을 때 IBM이 보유한 미국 등록 상표 또는 보통법상 상표임을 나타냅니다. 이러한 상표는 다른 국가에서도 등록 상표 또는 보통법상 상표일 수 있습니다. 현재 IBM의 상표 목록은 "저작권 및 상표 정보"([ibm.com/legal/copytrade.shtml](#))에 있습니다.

[ibm.com/legal/copytrade.shtml](#))에 있습니다.

Linux는 미국 또는 기타 국가에서 사용되는 Linus Torvalds의 상표입니다.

Microsoft, Windows, Windows NT 및 Windows 로고는 미국 또는 기타 국가에서 사용되는 Microsoft Corporation의 상표입니다.

기타 회사, 제품 및 서비스 이름은 타사의 상표 또는 서비스표입니다.

미국 정부 사용자의 제한된 권한 - IBM Corp.과의 GSA ADP 스케줄 계약에 따라 사용, 복제 또는 공개가 제한됩니다.

SPSS는 IBM 회사인 SPSS, Inc.의 상표이며, 전 세계 여러 지역에 등록되어 있습니다.

Produced in the United States of America  
05-10

© Copyright IBM Corporation 2010  
All Rights Reserved.

