



# IBM SPSS Modeler专业版

## 通过预测智能制定更佳决策

---

### 要点

- 通过评估趋势和可能的结果，制定更有效的战略。
- 借助这个直观、可视化的数据挖掘工作台，可轻松访问、准备和建模结构化数据。
- 使用最高级的统计和机器学习技术，快速建立和验证模型。
- 高效部署洞察力和预测模型。
- 直接在IBM® Cognos® Business Intelligence中无缝部署结果。
- 使用地图跨区域直观展示分析结果。
- 借助领先的数据库呈现数据库内分析。

通过分析历史数据，数据挖掘可让企业更清晰地查看当前环境并获取对未来事件更深入的洞察力。IBM SPSS® Modeler专业版是一个用来分析结构化数据并能创建预测智能的数据挖掘工作台。这种智能可使企业根据数据而非“直觉”做出业务决策

各类企业都发现他们可以使用IBM SPSS Modeler的预测性智能吸引客户，强化其忠诚度，根据成本效益地减少客户流失，减少欺诈和降低风险。公共部门组织可从SPSS Modeler的使用中大受裨益。他们可预测员工量能，前瞻性地应对公共安全问题，管理学生生命周期，改善课堂表现和应对许多其他的运营挑战。

预测智能可让您放眼未来。通过使用历史数据，SPSS Modeler可使您自信地预测结果，理解您的数据中隐藏的关系。有了这种理解力，您可使用强大可靠的分析技术，更深入地洞察您的客户、学生或选民，更快解决任何业务问题。



## 简化数据挖掘流程

SPSS Modeler专业版深受分析人员和业务用户的喜爱。其自动化数据准备和建模功能使毫无专业技能的非分析人员也能快速、简单地创建准确的模型。同时，专业分析人员可充分利用软件的高级数据准备和预测建模功能，建立最复杂的数据流。

## 企业优势

使用SPSS Modeler专业版，各类企业都能执行更深入的分析，获得洞察力，取得更大的成功。

- 企业可吸引客户，强化其忠诚度，更具成本效益地减少客户流失，识别和预防欺诈并降低风险。

- 政府部门可预测员工量能，评估方案成效并前瞻性地应对公共安全威胁。
- 教育机构可管理学生生命周期，改善课堂表现和应对许多其他有关学生和运营的挑战。
- 工业作业可完善维护规划，预防非计划停机。

用户可在直观的图形界面中轻松查看“数据流”中数据挖掘流程的每个步骤。通过与这些数据流进行交互，分析人员和业务用户可协作将业务知识和专业知识加入数据挖掘流程，可使他们专注于获取洞察力而非编写代码等技术工作。他们也能进行“思路”分析，更深入地探索数据，揭示对企业有用的其他关系。

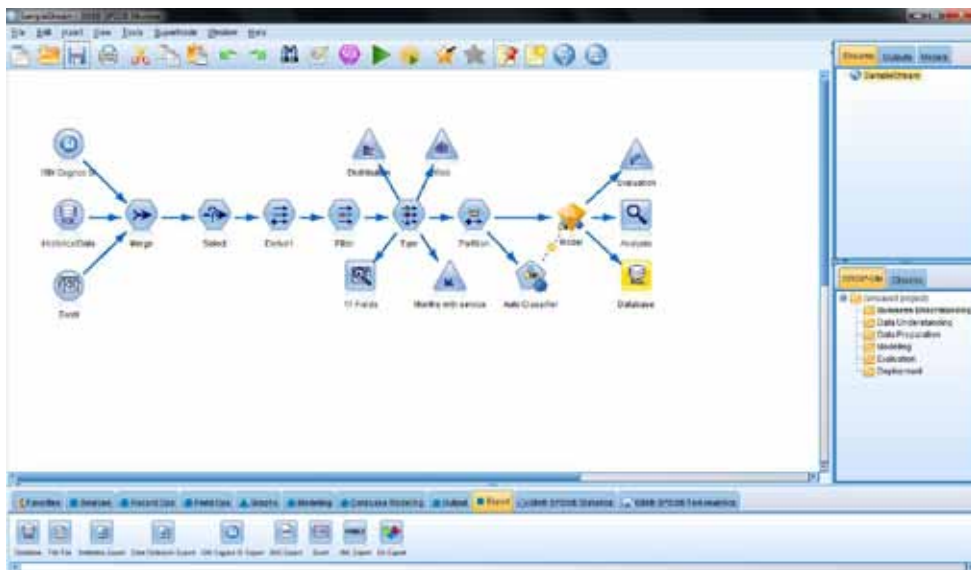


图1: SPSS Modeler的可视化界面可使用户快速、直观地建立预测模型而无需编程。

## 轻松访问和集成数据

您可从可视化界面中轻松地访问和集成几乎任何类型数据库、电子表格或平面文件中的数据,包括IBM SPSS Statistics、IBM SPSS Data Collection、Cognos Business Intelligence、SAS和Microsoft Excel文件。

和SPSS Modeler专业版服务器共同使用时,无需从巨大的数据库中移动数据,在数据库内就能执行分析和挖掘,这显著地提高了分析绩效。

SQL Pushback支持在数据库内进行数据转换和准备工作,用户不必编写任何SQL语句或做任何编程。此外,可直接从用户友好的SPSS Modeler界面访问领先数据库的算法,将这些算法作为SPSS Modeler数据流的一部分加以构建和评分——也支持IBM InfoSphere<sup>®</sup>、Microsoft SQL Server、Oracle和IBM Netezza。SPSS Scoring Adapters支持在数据库内快速对数据进行评分,使您更快做出决策并取得更优投资回报率(ROI)。

## 与IBM Cognos软件集成

分析人员可直接在SPSS Modeler界面内访问来自Cognos Business Intelligence环境的数据。IBM Cognos软件组织可提供用于企业级决策制定的完整且一致的信息视图。通过添加SPSS

Modeler专业版的分析功能,企业可使用其熟悉的企业级数据视图快速可靠地判断具体结果的可能性。

也正因为SPSS Modeler可将结果写入Cognos Business Intelligence——他们可以向业务用户和所有以Cognos作为企业分析信息门户的信息相关者提供预测智能。

## 选择一种无与伦比的技术

SPSS Modeler可提供一系列高级数据挖掘技术,专为满足各种数据挖掘应用程序所需而设计,包括如下所有的算法。

- **分类算法**——使用决策树(Decision Tree)、神经网络(Neural Networks)、逻辑回归(Logistic Regression)、支持向量机(Support Vector Machines)、多因素(Cox)回归和广义线性混合模型(GLMM)等技术,根据历史数据进行预测和预报。将自动分类模型同时用于二进制和数值型结果,简化模型创建。
- **分割算法**——凭借自动聚类、异常检测和聚类神经网络技术,将人员分类或检测不寻常模式。使用自动分类可在一个步骤中使用多种算法,不必在选择合适技术时猜来猜去。
- **关联算法**——使用Apriori、CARMA和序列关联发现关联、链接或序列。
- **时间序列与预测**——使用统计建模技术对一个或多个时间序列做出预测。

## 优化您当前的信息技术

SPSS Modeler凭借其开放且可伸缩的架构,可充分利用您现有的IT基础架构。它可在访问数据和部署结果时集成您的现有系统,所以您不必将数据以一种专有格式移入和移出。此外,数据库内挖掘、SQL pushback、多线程、服务器集群和数据库内评分等技术有助于保护资源、更快提供结果并减少IT总成本。

## 遵循一种可靠、可重复的流程

在数据挖掘流程的每个阶段, SPSS Modeler都支持实际行业标准,即跨行业数据挖掘标准流程(CRISP-DM)这意味着用户可通过数据挖掘专注于解决业务问题,而非为每个项目都彻底改造出一种全新流程。使用CRISP-DM项目管理器,可有效组织各个Modeler项目。

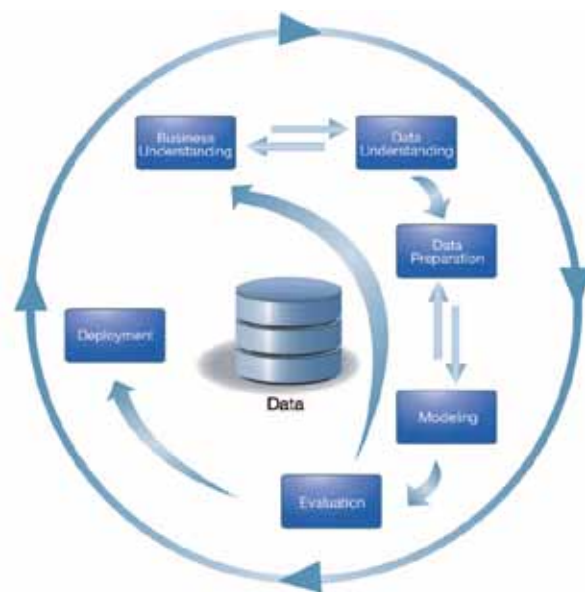


图2: 如图所示, CRISP-DM流程可使数据挖掘者有效实施数据挖掘项目,得出可衡量的业务结果。

## 跨企业部署预测模型

SPSS Modeler可高效分析大量来自中小型企业的典型数据。需要挖掘大容量或复杂数据的企业可充分利用IBM SPSS Modeler服务器提供的额外功能。通过使用客户端/服务器架构, SPSS Modeler服务器可在不消耗计算资源的情况下, 让多个数据分析人员同时工作。您可在领先的平台上充分利用数据库内挖掘和数据库内评分, 高效处理大量数据。SPSS Modeler服务器也提供了额外的部署选项, 帮助您跨地域和功能进行数据挖掘, 更快将结果提交至决策者手中。

IBM SPSS Modeler专业版功能	
数据理解	<ul style="list-style-type: none"> <li>• 凭借自动辅助创建大量互动图表</li> <li>• 使用可视链接分析查看您数据中的关联信息</li> <li>• 通过在图表中选择地区和项目及查看所选信息, 进行数据交互; 或选择关键信息进行分析</li> <li>• 可直接从SPSS Modeler界面访问SPSS Statistics图表和报告工具</li> </ul>
数据准备	<ul style="list-style-type: none"> <li>• 可访问来自Cognos Business Intelligence、IBM DB2<sup>®</sup>、Oracle、Microsoft SQL Server、IBM Informix<sup>®</sup>、IBM Netezza、MySQL (Oracle)和Teradata的业务数据, 也可通过zDB2和IBM Classic Federation Server支持访问大型机数据</li> <li>• 可导入分割文本文件和固定宽度文本文件、SPSS Statistics文件、SPSS Data Collection数据源、Excel、SAS或XML</li> <li>• 可从SPSS Modeler中选择多种数据清理选项, 删除或替换无效数据, 自动填充缺失值并减少离群值和极端值</li> <li>• 将自动数据准备用于查询和限制数据, 以便在一个单独步骤中进行分析</li> <li>• 从SPSS Modeler直接访问SPSS Statistics中执行的数据管理和转换</li> <li>• 使用字段过滤、命名、推导、合并、再范畴化、价值重置和字段重新排序</li> <li>• 应用记录选择、抽样(包括整群和分层抽样)、合并(包括内连接、完全外连接、部分外连接和反对连接)、排序、聚合和平衡</li> <li>• 可选择数据再构成、分割和对换选项</li> <li>• 可选择大量的字符串函数: 字符串创建、替代、查找和匹配, 空格移除和截断</li> <li>• 应用RFM (近因、频数和货币值)评分: 合计客户交易, 提供RFM价值数并将其结合进行完整的RFM分析</li> <li>• 可将数据导出至数据库、IBM Cognos Business Intelligence包、SPSS Statistics、SPSS Data Collection, 可导出的格式为分割文本文件、Excel、SAS或XML</li> </ul>

IBM SPSS Modeler专业版功能(续)	
建模算法	<ul style="list-style-type: none"> <li>异常检测——通过使用基于集群的算法检测不寻常的记录</li> <li>Apriori——借助高级判断函数的流行关联发现算法</li> <li>Bayesian Networks——图形概率模型</li> <li>C&amp;RT, C5.0, CHAID和QUEST——决策树算法, 包括交互式树构建</li> <li>CARMA——支持多结果的关联算法</li> <li>Cox回归——计算事件可能发生的时间</li> <li>Decision List——交互式规则建立算法</li> <li>Factor/PCA, Feature Selection——数据压缩算法</li> <li>IBM InfoSphere*数据库内挖掘算法: 关联(Association)、集群(Clustering)、决策树(Decision Tree)、逻辑回归(Logistic Regression)、朴素贝叶斯(Naive Bayes)、回归(Regression)、序列(Sequence)、时间序列(Time Series)</li> <li>IBM Netezza*数据库内挖掘算法: 贝叶斯网(Bayes Net)、决策树(Decision Trees)、分离集群(Divisive Clustering)、广义线性(Generalized Linear)、K均方聚类(K-Means)、KNN、线性回归(Linear Regression)、朴素贝叶斯(Naive Bayes)、PCA、回归树(Regression Tree)、时间序列(Time Series)</li> <li>Microsoft SQL Server*数据库内挖掘算法: 关联规则(Association Rules)、集群(Clustering)、决策树(Decision Tree)、线性回归(Linear Regression)、朴素贝叶斯(Naive Bayes)、神经网络(Neural Network)、序列集群(Sequence Clustering)、时间序列(Time-Series)</li> <li>Oracle*数据库内挖掘算法: 自适应贝叶斯(Adaptive Bayes)、Apriori、人工智能(AI)、决策树(Decision Tree)、一般线性模型(GLM)、K均方聚类(KMeans)、最小描述长度(MDL)、朴素贝叶斯(Naive Bayes)、非负矩阵分解(Non-Negative Matrix Factorization)、正交分区集群(Orthogonal Partitioning Clustering)、支持向量机(Support Vector Machine)</li> <li>K均方聚类(K-Means)、Kohonen、Two Step、Discriminant、支持向量机(SVM)——集群和分割算法</li> <li>KNN——最近邻建模和评分算法</li> <li>逻辑回归(Logistic Regression)——面向二进制结果</li> <li>神经网络(Neural Network)——具有反向传播学习和径向基函数网络的多层感知器</li> <li>回归(Regression)、线性(Linear)、一般线性模型(GLM)、广义线性混合模型(GLMM)——线性方程模型</li> <li>自学响应模型(SLRM)——具有增量学习的贝叶斯模型</li> <li>序列(Sequence)——面向序列敏感分析的序列关联算法</li> <li>支持向量机(Support Vector Machine)——面向广泛数据集的精准高级算法</li> <li>时间序列(Time-Series)——生成和自动选择时间序列预测模型</li> </ul>
建模和评估	<ul style="list-style-type: none"> <li>结合使用大量数据挖掘算法和高级功能, 从您的数据中获取最佳结果</li> <li>使用自动(二进制和数值型)分类和集群, 取代选择单一算法</li> <li>使用交互模型和方程浏览器, 查看高级统计输出</li> <li>借助变量重要性图表, 在预测的结果中展示数据属性的相关影响</li> <li>在地图中可视化分析结果</li> <li>结合多种模型(集成建模)或使用一种模型分析另一种模型</li> <li>使用SPSS Modeler Component-Level Framework (CLEF)集成自定义算法</li> <li>通过SPSS Statistics的集成, 使用R统计编程语言扩展分析选项</li> </ul>
部署	<ul style="list-style-type: none"> <li>使用SQL或PMML (预测模型的基于XML的标准格式)导出模型</li> <li>使用IBM SPSS Collaboration and Deployment Services的创新分析管理、流程自动化和部署功能</li> </ul>
<ul style="list-style-type: none"> <li>Modeler服务器*</li> </ul>	<ul style="list-style-type: none"> <li>使用领先的数据库技术并利用高性能的数据库实现, 使用数据库内挖掘在数据库内建立模型</li> <li>使用SQL-pushback推动数据转换并选择建模算法直接用于您的操作数据库</li> <li>通过并行执行数据流和多种模型, 利用包括IBM System z大型机在内的高性能硬件, 可更快创建解决方案并取得更佳投资回报率(ROI)</li> <li>通过安全套接字层(SSL)加密, 可在SPSS Modeler Client和SPSS Modeler Server间安全传输敏感数据</li> </ul> <p>*需要Modeler专业版服务器</p>