IBM SPSS Modeler Professional

Make better decisions through predictive intelligence

Highlights

Create more effective strategies by evaluating trends and likely outcomes.

- Easily access, prepare and model structured data with this intuitive, visual data mining workbench.

- Rapidly build and validate models using the most advanced statistical and machine-learning techniques available.

- Efficiently deploy insight and predictive models.

- Seamlessly deploy results directly into IBM® Cognos® Business Intelligence.

- Present analytic results across regions visually with geographical mapping.

- Perform in-database analytics with leading databases.

Data mining provides organizations with a clearer view of current conditions and deeper insight into future events by analyzing historical data. IBM SPSS® Modeler Professional is a data mining workbench used to analyze structured data to create predictive intelligence. This intelligence allows business decisions to be made based on the data, rather than "gut feel".

Businesses of all types have found that they can use the predictive intelligence of IBM SPSS Modeler to attract customers, strengthen their loyalty, reduce customer attrition or "churn" more cost effectively, decrease fraud and minimize risk. Public sector organizations have benefited from using Modeler to predict workforce capacity, proactively respond to public safety issues, manage the student lifecycle, improve classroom performance and address many other operational challenges.

Predictive intelligence enables you to look into the future. Using historical data, SPSS Modeler lets you confidently predict outcomes and gain an understanding about the relationships hidden in your data. This understanding allows you to solve any business problem faster using powerful, proven analytical techniques that deliver deeper insight into your customers, students or constituents.

Streamline the data mining process

SPSS Modeler Professional is popular with analysts and business users alike. Its automated data preparation and modeling features enable non-analysts to produce accurate models quickly and easily without specialized skills. While professional analysts can take advantage of the software's advanced data preparation and predictive modeling capabilities to create the most sophisticated of streams.

Organization benefits

Using SPSS Modeler Professional, all types of organizations are performing more in-depth analyses that help them gain insight to achieve greater success.

- Businesses attract customers, strengthen their loyalty, reduce customer attrition more cost effectively, identify and prevent fraud and reduce risk.
- Public sector organizations predict workforce capacity, evaluate program effectiveness and proactively respond to public safety issues.

- Educational institutions manage the student lifecycle, improve classroom performance and address many other student and operational challenges.
- Industrial operations refine maintenance planning to prevent unscheduled downtime.

The intuitive graphical interface of SPSS Modeler makes it easy for users to visualize every step of the data mining process as part of a "stream." By interacting with these streams, analysts and business users can collaborate – adding business knowledge and domain expertise to the data mining process. It allows them to focus on discovering insights rather than on technical tasks like writing code. They can also pursue "train-of-thought" analysis, explore the data more deeply; uncovering additional relationships that make sense to the organization.
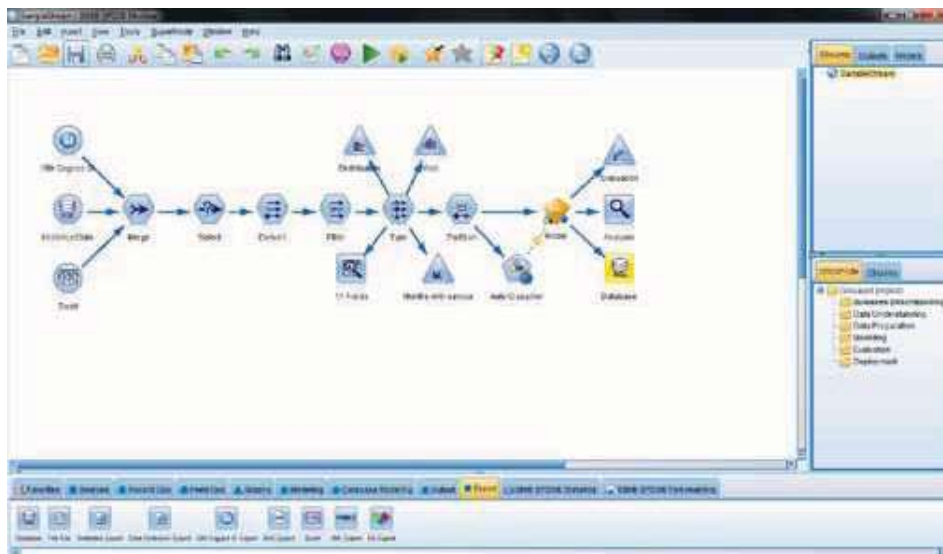

Figure 1: The SPSS Modeler visual interface enables users to build predictive models quickly and intuitively, without the need for programming.

Easily access and integrate data

From this visual interface, you can easily access and integrate data in virtually any type of database, spreadsheet or flat file, including IBM SPSS Statistics, IBM SPSS Data Collection, Cognos Business Intelligence, SAS and Microsoft Excel files.

When combined with SPSS Modeler Professional Server, there is no need to move data from large databases, since the analytics and mining take place in-database. The result is a significant improvement in analytical performance.

SQL Pushback allows for data transformation and preparation tasks to be performed within the database without the user having to write any SQL or do any programming. Additionally, algorithms from leading databases can be accessed directly with the user-friendly SPSS Modeler interface and built and scored as part of a Modeler stream – supported with IBM InfoSphere®, Microsoft SQL Server, Oracle and IBM Netezza. SPSS Scoring Adapters allow the data to be scored in-database, resulting in quicker decisions and better ROI.

Integration with IBM Cognos software

Analysts can access data from their Cognos Business Intelligence environment directly within the SPSS Modeler interface. IBM Cognos software organizes and delivers a complete and consistent view of information for enterprise-wide decision making. By adding the analytic capabilities of SPSS Modeler Professional, organizations can quickly and reliably evaluate the likelihood of specific outcomes, using their familiar enterprise-wide data view.

Also, since SPSS Modeler can write results into Cognos Business Intelligence – they can make predictive intelligence available to business users and all information stakeholders who rely on Cognos as their information portal into enterprise analytics.

Choose from an unparalleled breadth of techniques

SPSS Modeler offers an array of advanced data mining techniques that are designed to meet the needs of every data mining application, including all of the following algorithms.

- **Classification algorithms**. Make predictions or forecasts based on historical data using techniques such as Decision Trees, Neural Networks, Logistic Regression, Support Vector Machines, Cox regression, Generalized Linear Mixed Models (GLMM) and more. Use automatic classification modeling for both binary and numeric outcomes to streamline model creation.
- **Segmentation algorithms**. Group people or detect unusual patterns with automatic clustering, anomaly detection and clustering neural network techniques. Use automatic classification to apply multiple algorithms with a single step and take the guesswork out of selecting the right technique.
- **Association algorithms**. Discover associations, links or sequences using Apriori, CARMA and sequential association.
- **Time series and forecasting**. Generate forecasts for one or more series over time using statistical modeling techniques.

Optimize your current information technologies

With its open and scalable architecture, SPSS Modeler makes the best use of your existing IT infrastructure. It integrates with your existing systems, both when accessing data and when deploying results, so you don't need to move data into and out of a proprietary format. Additionally, techniques such as in-database mining, SQL pushback, multi-threading, server clustering and in-database scoring help conserve resources, deliver results faster and reduce overall IT costs.

Follow a proven, repeatable process

During every phase of the data mining process, SPSS Modeler supports the de facto industry standard, the CRoss-Industry Standard Process for Data Mining (CRISP-DM). This means users can focus on solving business problems through data mining, rather than on reinventing a new process for every project. Individual Modeler projects can be efficiently organized using the CRISP-DM project manager.
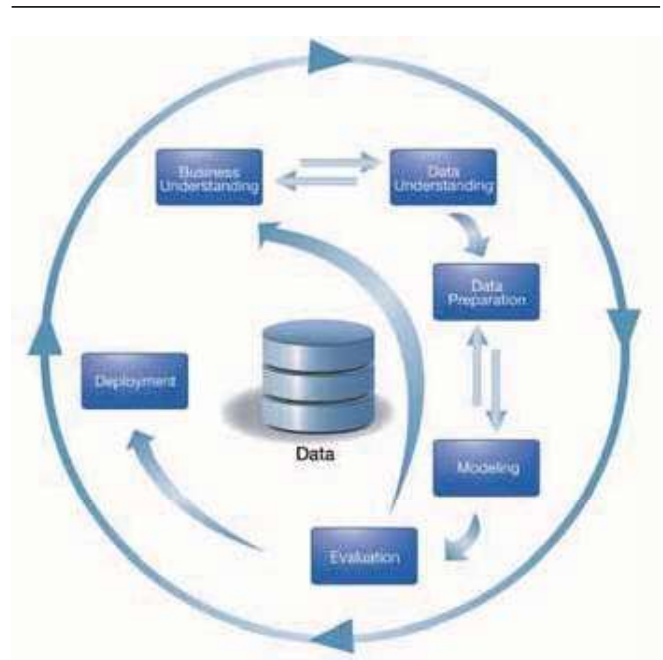

Figure 2: The CRISP-DM process, as shown in this diagram, enables data miners to implement efficient data mining projects that yield measurable business results.

Deploy predictive modeling across the enterprise

SPSS Modeler can efficiently analyze the amounts of data typically generated by small to mid-sized organizations. Organizations with high-volume or complex data mining requirements can take advantage of the additional power of SPSS Modeler Server. Using client/server architecture, SPSS Modeler Server allows many data analysts to work simultaneously without straining computing resources. You can take advantage of in-database mining and in-database scoring on leading platforms and efficiently process large amounts of data. SPSS Modeler Server also offers additional deployment options to help you extend the benefits of data mining across geographic or functional lines and put results in the hands of decision makers quickly.

IBM SPSS Modeler Professional features	
Data understanding	• Create a wide range of interactive graphs with automatic assistance • Use visual link analysis to see associations in your data • Interact with data by selecting regions or items on a graph and viewing the selected information; or select key data for use in analysis • Access SPSS Statistics graphs and reporting tools directly from the SPSS Modeler interface
Data preparation	• Access operational data from Cognos Business Intelligence, IBM DB2®, Oracle, Microsoft SQL Server, IBM Informix®, IBM Netezza, mySQL (Oracle) and Teradata data sources, as well as mainframe data through zDB2 and IBM Classic Federation Server support • Import delimited and fixed-width text files, SPSS Statistics files, SPSS Data Collection data sources, Excel, SAS or XML • Choose from the multiple data-cleaning options available in SPSS Modeler to remove or replace invalid data, automatically impute missing values and mitigate outliers and extremes • Apply automatic data preparation to interrogate and condition data for analysis in a single step • Access data management and transformations performed in SPSS Statistics directly from SPSS Modeler • Use field filtering, naming, derivation, binning, re-categorization, value replacement and field reordering • Apply record selection, sampling (including clustered and stratified sampling), merging (including inner joins, full outer joins, partial outer joins, and anti-joins), sorting, aggregation and balancing • Choose from options for data restructuring, partitioning and transposition • Select from extensive string functions: string creation, substitution, search and matching, whitespace removal and truncation • Apply RFM (Recency, Frequency, and Monetary) scoring: aggregate customer transactions to provide RFM value scores and combine these to produce a complete RFM analysis • Export data to databases, IBM Cognos Business Intelligence packages, SPSS Statistics, SPSS Data Collection, delimited text files, Excel, SAS, or XML

IBM SPSS Modeler Professional features (continued)	
Modeling algorithms included	• Anomaly Detection – Detect unusual records through the use of a cluster-based algorithm • Apriori – Popular association discovery algorithm with advanced evaluation functions • Bayesian Networks – Graphical probabilistic models • C&RT, C5.0, CHAID and QUEST – Decision tree algorithms including interactive tree building • CARMA – Association algorithm which supports multiple consequents • Cox regression – Calculate likely time to an event • Decision List – Interactive rule-building algorithm • Factor/PCA, Feature Selection – Data reduction algorithms • In-database mining algorithms for IBM InfoSphere*: Association, Clustering, Decision Tree, Logistic Regression, Naive Bayes, Regression, Sequence, Time Series • In-database mining algorithms for IBM Netezza*: Bayes Net, Decision Trees, Divisive Clustering, Generalized Linear, K-Means, KNN, Linear Regression, Naive Bayes, PCA, Regression Tree, Time Series • In-database mining algorithms for Microsoft SQL Server*: Association Rules, Clustering, Decision Tree, Linear Regression, Naive Bayes, Neural Network, Sequence Clustering, Time-Series • In-database mining algorithms for Oracle*: Adaptive Bayes, Apriori, Artificial Intelligence (AI), Decision Tree, General Linear Model (GLM), KMeans, Minimum Description Length (MDL), Naive Bayes, Non-Negative Matrix Factorization, O-Cluster (Orthogonal Partitioning Clustering), Support Vector Machine • K-Means, Kohonen, Two Step, Discriminant, Support Vector Machine (SVM) – Clustering and segmentation algorithms • KNN – Nearest neighbor modeling and scoring algorithm • Logistic Regression – For binary outcomes • Neural Networks – Multi-layer perceptrons with back-propagation learning, and radial basis function networks • Regression, Linear, GenLin (GLM), Generalized Linear Mixed Models (GLMM) – Linear equation modeling • Self-learning response model (SLRM) – Bayesian model with incremental learning • Sequence – Sequential association algorithm for order-sensitive analyses • Support Vector Machine – Advanced algorithm with accurate performance for wide datasets • Time-series – Generate and automatically select time-series forecasting models
Modeling and evaluation	• Employ a wide range of data mining algorithms with many advanced features to get the best possible results from your data. • Use automatic classification (binary and numeric) and clustering in place of selecting individual algorithms • Use interactive model and equation browsers and view advanced statistical output • Show relative impact of data attributes on predicted outcomes with variable importance graphs • Visualize the analytic results on geographic maps • Combine multiple models (ensemble modeling) or use one model to analyze a second model • Use the SPSS Modeler Component-Level Extension Framework (CLEF) to integrate custom algorithms • Use the R statistical programming language to extend analysis options, through the integration of SPSS Statistics
Deployment	• Export models using SQL or PMML (the XML-based standard format for predictive models) • Use IBM SPSS Collaboration and Deployment Services for innovative analytics management, process automation and deployment capabilities
Modeler server*	• Use in-database mining to build models in the database using leading database technologies and leverage high-performance database implementations • Use SQL-pushback to push data transformations and select modeling algorithms directly into your operational databases • Leverage high-performance hardware, including IBM System z® machines, to experience quicker time-to-solution and achieve greater ROI through parallel execution of streams and multiple models • Transmit sensitive data securely between SPSS Modeler Client and SPSS Modeler Server through secure socket layer (SSL) encryption

*Requires Modeler Professional Server