



---

## Highlights

*Solve business problems faster with analytical techniques that deliver deeper insight.*

- Easily access, prepare and integrate structured data and text, web and survey data.
  - Quickly identify and extract concepts and sentiments from text in more than 30 languages.
  - Improve the coherence and consistency of current data with entity analytics.
  - Transform information about relationships to show the social behavior of individuals and groups with social network analysis.
  - Perform in-database analytics with leading databases.
- 

# IBM SPSS Modeler Premium

*Improve model accuracy with structured and unstructured data, entity analytics and social network analysis*

IBM® SPSS® Modeler Premium is a high-performance predictive and text analytics workbench that gives you unprecedented insight from your data. Your organization can achieve superior outcomes and ultimately, a positive return on investment because this workbench lets you base your critical business decisions on predictive intelligence.

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

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.

But what happens if the majority of your data is trapped in unstructured or textual form—in comments, files or on the web? Modeling with structured data alone may provide an incomplete view into your business processes and outcomes.



SPSS Modeler Premium empowers you to tap into the predictive intelligence contained in all forms of data. It goes beyond the analysis of structured numerical data and includes information from unstructured data, such as web activity, blog content, customer feedback, emails, articles and more to discover relationships between concepts and sentiments and create the most accurate predictive models possible.

Its entity analytics capabilities enable you to improve the coherence and consistency of current data by resolving identity conflicts within the records themselves. Identity resolution is vital in customer relationship management, fraud detection, anti-money laundering and security.

The social network capabilities in SPSS Modeler Premium transform information about relationships into key performance indicators that show the social behavior of individuals and groups. You can then identify social leaders who influence the behavior of others in the network. Combine these results with other measures and you can create comprehensive profiles of individuals to base your predictive models.

### **Boost productivity, quality and accuracy**

SPSS Modeler Premium allows you to perform both text analysis and data mining within an interactive, visual environment. The intuitive graphical interface makes it easy to see every step of the data mining process as part of a “stream.” Text analytics is straightforward and efficient, with interactive graphs to help you explore and display text data and patterns for instant analysis, as well as powerful classification and categorization techniques that transform text into an analytical asset.

From the visual interface, you can easily access and integrate data from many sources, including data in virtually any type of database, spreadsheet or flat file (such as IBM SPSS Statistics, SAS and Microsoft Excel files), textual data and data from Web 2.0 sources (such as RSS feeds), IBM SPSS Data Collection and IBM Cognos® Business Intelligence and legacy systems with IBM Classic Federation Server and zDB2 for z/OS support. No other data mining solution offers this versatility.

When combined with SPSS Modeler Premium Server, there is no need to move data from large databases as the analytics and mining take place in-database. This results in 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. SPSS Scoring Adapters allow models to be quickly scored in-database resulting in quicker decisions and better ROI. Additionally, algorithms from leading databases can be accessed directly with the user-friendly SPSS Modeler interface and built and scored as part of an SPSS Modeler stream—supported with IBM InfoSphere®, Microsoft SQL Server, Oracle and IBM Netezza.

With powerful automation tools such as automated data preparation and auto modeling, SPSS Modeler Premium makes it easy to prepare data for analysis, find the best model to use and quickly produce consistent and accurate results. SPSS Modeler supports the Cross-Industry Standard Process for Data Mining (CRISP-DM), which enables analysts to focus on solving business problems, rather than on programming, and individual projects can be efficiently organized using the CRISP-DM project manager.

## Take a more in-depth look at IBM SPSS Modeler Premium

With SPSS Modeler Premium, you can be more focused and agile in your planning and daily decision making because of its unique combination of capabilities. It enables you to have a better understanding of your organization, the environment in which you operate, your customers and other stakeholders. These capabilities include:

- **Entity analytics.** Organizations often combine multiple data sources. But what happens when there is no clear match between records? SPSS Modeler Premium's entity analytics capabilities let you discover these relationships and deliver non-obvious identity and relationship awareness. It allows you to consolidate records where appropriate—or keep them separate. Entity analytics is critical for border security, detecting fraud and properly identifying criminal suspects. But it is also enormously useful for businesses wanting to avoid presenting different offers to the same person in a marketing campaign or to ensure that you are building accurate models.
- **Social network analysis.** Discover the relationships among social entities and the resulting implications of these relationships on an individual's behavior. SPSS Modeler Premium's social network analysis capabilities are particularly useful for those in telecommunications and other industries concerned about churn. You can identify groups, group leaders and whether others will churn based on their influence. It consists of two capabilities that greatly enhance models to predict churn:
  - Group analysis—Identify the groups in the data and who are their leaders.
  - Diffusion analysis—Use existing churn information to determine who the current churners are and what will influence them to leave.

- **Big data analytics.** IBM SPSS Modeler Premium easily integrates with IBM and non-IBM databases, allowing models to be deployed and scored with greater speed and efficiency. Data can be scored within the database and in real time against transactional data such as high volume sales, customer service and claims transactions.
- **IBM Cognos software integration.** Analysts can access structured data directly through Cognos Business Intelligence software to quickly and reliably evaluate the likelihood of specific outcomes. You can make predictive intelligence gleaned from a customer view that combines structured and unstructured data available to business users and all stakeholders who rely on Cognos dashboards.
- **IBM Netezza functionality.** Perform in-database analytics within the Netezza appliance to build and deploy analytic applications that scale to sizes only addressable by supercomputers today.

Additionally, SPSS Modeler Premium can extract insights from text data, enabling organizations to:

- Receive faster, more accurate results when analyzing banking, insurance or advertising text, emoticons and slang with new industry-specific text analysis packages and templates.
- Create hierarchical categorization structures to organize concepts more logically and in greater detail.
- Import pre-defined categories, including hierarchical categories, annotations and keyword descriptors and export them to Excel.
- Save hierarchical categories for reuse with an enhanced semantic network grouping technique for category building.
- Extract text faster and more accurately, especially when working with large datasets, by leveraging new industry sensitive semantic networks and more efficient use of hardware.
- Define and test rules on sample text before applying them to your data, using the enhanced text link rule editor.

## Explore a broader range of data

Only with SPSS Modeler Premium can you directly access text, web and survey data and integrate these additional types of data into your predictive models for more useful recommendations and improved outcomes—and with entity analytics make sure you are using the best data for modeling. Through a proven natural language processing (NLP) linguistic extraction process, SPSS Modeler Premium pulls key concepts from many types of unstructured data and groups them into categories. The interactive text mining workbench saves you time and effort while helping you gain a competitive advantage. Unlike other text analytics tools, you do not need a linguistic background to use it. You can easily customize concept dictionaries for a particular domain area by using the Resource Editor, an integrated resource for managing the text extraction process. This enables you to find relevant concepts and associations faster. Extracted concepts, opinions and categories are then combined with structured data and applied to predictive models to uncover valuable insights into actions, behaviors, patterns and associations.

SPSS Modeler Premium includes customizable templates and libraries for specific business applications. The linguistic resources support a range of industries and applications, including sentiment analysis, CRM, security and intelligence, market intelligence, life sciences (genomics and MESH) and IT. It also supports multiple languages, including Dutch, English, French, German, Italian, Japanese, Portuguese and Spanish text. The English-language version also includes an interface that supports third-party translation options.

## Choose from an unparalleled breadth of techniques

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

## Support data mining across the enterprise

SPSS Modeler Premium 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 leverage IBM SPSS Modeler Premium Server. Using client/server architecture, SPSS Modeler Premium Server makes it possible for many analysts to work simultaneously without straining computing resources. The enterprise version supports in-database mining on leading information platforms and efficiently processes large amounts of data.

SPSS Modeler Premium Server also offers additional deployment options to help you extend the benefits of predictive analytics across geographic or functional lines and put results in the hands of decision makers quickly. For example, to help manage your analytical assets and automate analytical processes, use SPSS Modeler Premium with IBM SPSS Collaboration and Deployment Services.

IBM SPSS Modeler Premium features	
<b>Data understanding</b>	<ul style="list-style-type: none"> <li>• Create a wide range of interactive graphs with automatic assistance</li> <li>• Use visual link analysis to see associations in your data</li> <li>• Interact with data by selecting regions or items on a graph and viewing the selected information; or select key data for use in analysis</li> <li>• Access SPSS Statistics graphs and reporting tools directly from the SPSS Modeler Premium interface</li> </ul>
<b>Data preparation</b>	<ul style="list-style-type: none"> <li>• 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</li> <li>• Import delimited and fixed-width text files, SPSS Statistics files, SAS, SPSS Data Collection data sources or XML</li> <li>• Choose from the multiple data-cleaning options available in SPSS Modeler Premium to remove or replace invalid data, automatically impute missing values and mitigate outliers and extremes</li> <li>• Use entity analytics to combine or separate records resulting in cleaner data for modeling</li> <li>• Apply automatic data preparation to interrogate and condition data for analysis in a single step</li> <li>• Access data management and transformations performed in SPSS Statistics directly from SPSS Modeler</li> <li>• Use field filtering, naming, derivation, binning, re-categorization, value replacement and field reordering</li> <li>• Apply record selection, sampling (including clustered and stratified sampling), merging (including inner joins, full outer joins, partial outer joins, and anti-joins), and concatenation; sorting, aggregation and balancing</li> <li>• Choose from options for data restructuring, partitioning and transposition</li> <li>• Select from extensive string functions: string creation, substitution, search and matching, whitespace removal and truncation</li> <li>• Apply RFM (Recency, Frequency, and Monetary) scoring: aggregate customer transactions to provide value scores and combine these to produce a complete RFM analysis</li> <li>• Identify the groups in the data and who are considered the leaders of them with Group Analysis</li> <li>• Utilize churn information to determine who else is likely to be influenced by the churner to also churn with Diffusion analysis</li> <li>• Export data to databases, IBM Cognos Business Intelligence packages, SPSS Statistics, SPSS Data Collection, delimited text files, Excel, SAS, or XML</li> </ul>
<b>Modeling algorithms included</b>	<ul style="list-style-type: none"> <li>• Anomaly Detection—Detect unusual records through the use of a cluster-based algorithm</li> <li>• Apriori—Popular association discovery algorithm with advanced evaluation functions</li> <li>• Bayesian Networks—Graphical probabilistic models</li> <li>• C&amp;RT, C5.0, CHAID and QUEST—Decision tree algorithms including interactive tree building</li> <li>• CARMA—Association algorithm which supports multiple consequents</li> <li>• Cox regression—Calculate likely time to an event</li> <li>• Decision List—Interactive rule-building algorithm</li> <li>• Factor/PCA, Feature Selection—Data reduction algorithms</li> <li>• In-database mining algorithms for IBM InfoSphere*: Association, Clustering, Decision Tree, Logistic Regression, Naive Bayes, Regression, Sequence, Time Series</li> <li>• 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</li> <li>• In-database mining algorithms for Microsoft SQL Server*: Association Rules, Clustering, Decision Tree, Linear Regression, Naive Bayes, Neural Network, Sequence Clustering, Time-Series</li> <li>• 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</li> <li>• K-Means, Kohonen, Two Step, Discriminant, Support Vector Machine (SVM)—Clustering and segmentation algorithms</li> <li>• KNN—Nearest neighbor modeling and scoring algorithm</li> <li>• Logistic Regression—For binary outcomes</li> <li>• Neural Networks—Multi-layer perceptrons with back-propagation learning, and radial basis function networks</li> <li>• Regression, Linear, GenLin (GLM), Generalized Linear Mixed Models (GLMM)—Linear equation modeling</li> <li>• Self-learning response model (SLRM)—Bayesian model with incremental learning</li> <li>• Sequence—Sequential association algorithm for order-sensitive analyses</li> <li>• Support Vector Machine—Advanced algorithm with accurate performance for wide datasets</li> <li>• Time-series—Generate and automatically select time-series forecasting models</li> </ul>

\*Capability only with IBM SPSS Modeler Premium Server

IBM SPSS Modeler Premium features (continued)	
<b>Modeling and evaluation</b>	<ul style="list-style-type: none"> <li>• Employ a wide range of data mining algorithms with many advanced features to get the best possible results from your data.</li> <li>• Use automatic classification (binary and numeric) and clustering in place of selecting individual algorithms</li> <li>• Use interactive model and equation browsers and view advanced statistical output</li> <li>• Show relative impact of data attributes on predicted outcomes with variable importance graphs</li> <li>• Visualize the analytic results on geographic maps</li> <li>• Combine multiple models (ensemble modeling) or use one model to analyze a second model</li> <li>• Use the SPSS Modeler Component-Level Extension Framework (CLEF) to integrate custom algorithms</li> <li>• Use the R statistical programming language to extend analysis options, through the integration of SPSS Statistics</li> </ul>
<b>Deployment</b>	<ul style="list-style-type: none"> <li>• Export models using SQL or PMML (the XML-based standard format for predictive models)</li> <li>• Use IBM SPSS Collaboration and Deployment Services for innovative analytics management, process automation and deployment capabilities</li> </ul>
<b>Text-specific understanding and preparation features</b>	<ul style="list-style-type: none"> <li>• Extract text data from files, operational databases and RSS feeds (i.e., blogs, web feeds)</li> <li>• Select native language extractor options for Dutch, English, French, German, Italian, Portuguese, Spanish or Japanese or translate virtually any language using third-party translation software</li> <li>• Extract domain-specific concepts such as uniterms, expressions, abbreviations, acronyms and more</li> <li>• Calculate synonyms using sophisticated linguistic algorithms and embedded or user-specified linguistic resources</li> <li>• Name concepts by person, organization, term, product, location and other user-defined types</li> <li>• Extract non-linguistic entities such as address, currency, time, phone number and Social Security number</li> <li>• Use and customize pre-built templates and libraries for sentiment analysis, CRM, security and intelligence, market intelligence, life sciences and IT</li> <li>• Leverage pre-packaged Text Analytics Packages (TAPs) for the most common business applications, or create your own</li> <li>• Create clusters based on term co-occurrence using concept clustering algorithms, which provide an at-a-glance view of main topics and the way in which they are related</li> <li>• Intelligently group text documents and records based on content, using text classification algorithms</li> <li>• Enable advanced concept selection and deselection for use in predictive modeling</li> <li>• Use text-based and visual reports to interrogate concept relationship, occurrence, frequency and type</li> </ul>
<b>Text link analysis</b>	<ul style="list-style-type: none"> <li>• Identify and extract sentiments (for example, likes and dislikes) from text in Dutch, English, French, German and Spanish</li> <li>• Identify links and associations between, for example, people and events or diseases and genes</li> <li>• Identify and extract content from URLs within blogs</li> <li>• Include opinions, semantic relationships and linked events in deployable predictive models</li> <li>• Reveal complex relationships through interactive graphs that show multiple semantic links between two concepts</li> </ul>
<b>SPSS Modeler server*</b>	<ul style="list-style-type: none"> <li>• Use in-database mining to build models in the database using leading database technologies and leverage high-performance database implementations</li> <li>• Use SQL-pushback to push data transformations and select modeling algorithms directly into your operational databases</li> <li>• Score data within the database, resulting in significant performance improvements via IBM SPSS Modeler Server Scoring Adapters</li> <li>• 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</li> <li>• Transmit sensitive data securely between SPSS Modeler Client and SPSS Modeler Server through secure sockets layer (SSL) encryption</li> </ul>

\*Capability only with IBM SPSS Modeler Premium Server