

IBM SPSS Modeler

Faster time to value
with visual data science

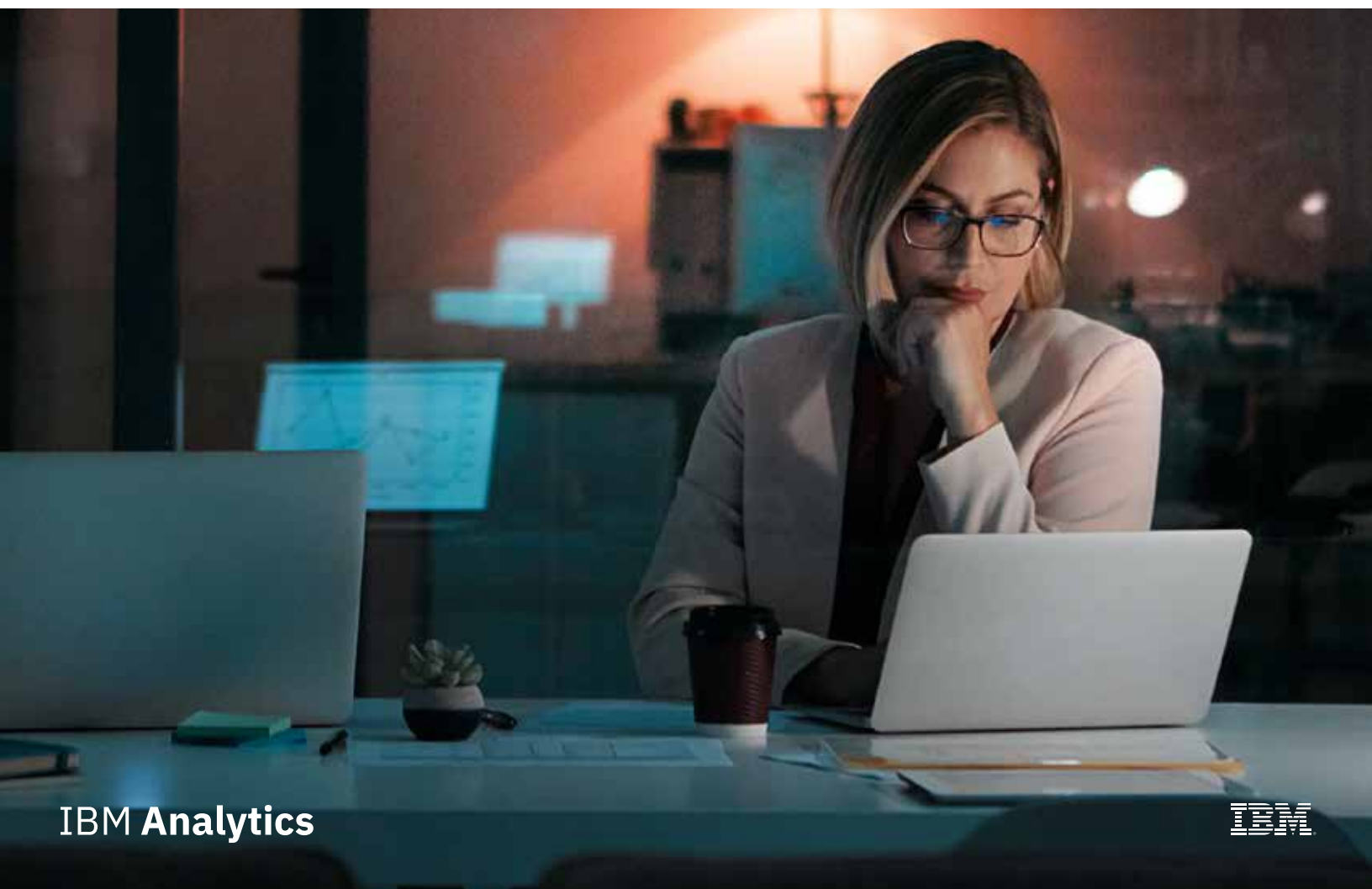


Table of contents

3 Introduction to data science

Follow the path to improved outcomes through data science and predictive analytics

4 IBM SPSS Modeler

Broader analytics, faster results, flexible deployment

6 IBM SPSS Modeler Subscription

Accessible, low-cost monthly pricing

7 IBM SPSS Modeler Professional

Make better decisions with desktop solution for a single or multi-user environment

8 IBM SPSS Modeler Premium

Enterprise-scale predictive intelligence deployed on premises or in the cloud

9 IBM SPSS Modeler Gold

See relationships in your data more clearly

10 IBM SPSS Collaboration and Deployment Service

Manage analytical assets, automate processes and efficiently share results

11 IBM SPSS Analytic Server

Make data science and predictive analytics easier for big data

Who uses IBM SPSS Modeler software?

Businesses use data science and predictive analytics:

Retail

- Marketing personalization
- Improving customer service
- Demand forecasting

Finance

- Fraud prevention
- Financial trade optimization
- Risk management
- Anti-money laundering

Telco

- Customer churn and retention
- Network performance

Utilities

- Energy usage pattern analysis
- Identify efficiency opportunities

Media and Entertainment

- Ad targeting
- Audience prediction

Manufacturing

- Sales and marketing forecasting and budgeting
- New product development
- Operational excellence
- Predictive maintenance

Colleges and universities use analytics for:

- Teaching and student assessment
- Administration
- Enrollment management
- Alumni development

Government agencies use analytics for:

- Human capital management
- Program evaluation
- Fighting crime and protecting public safety
- Promoting public health
- Preventing fraud, waste and abuse
- Compliance

Medical and healthcare organizations use analytics for:

- Evidence-based medicine
- Treatment outcome analysis
- Aided diagnosis
- Disease prevention

Follow the path to improved outcomes through data science and predictive analytics

CAPTURE information, PREDICT outcomes and ACT on insights

To get the answers you need for successful decision making, it's important to follow all the steps in the data analysis process—and using the right data analysis tools along the way can help you arrive at those decisions faster and more accurately.

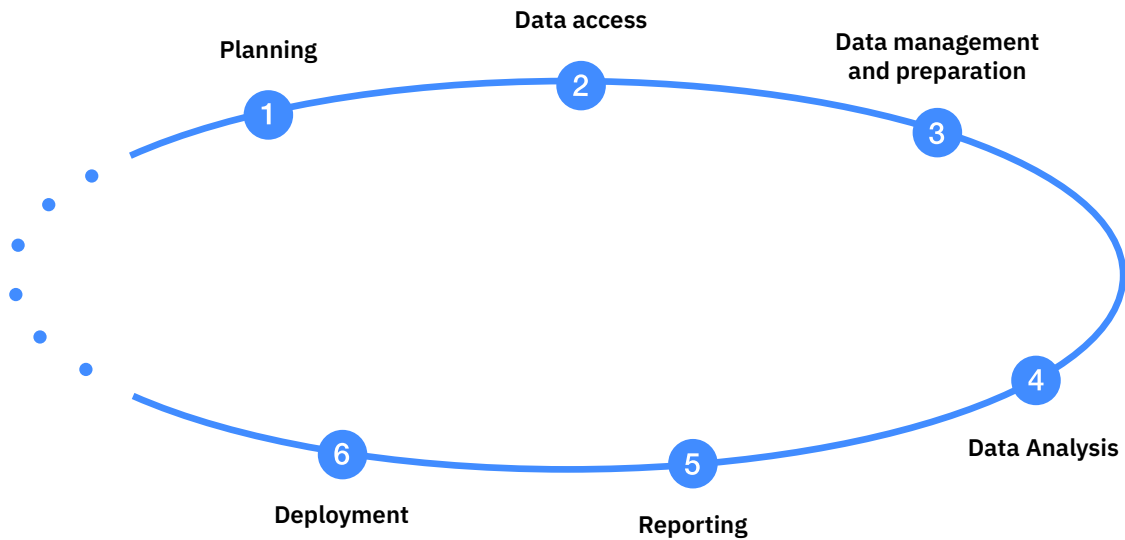
The process starts with planning. Beginning with the end result in mind, the first steps are to set objectives, identify data sources and carefully craft the process. The next step is data access, where data is brought in from available sources, using open database connectivity (ODBC) or direct file input. After that comes data management and data preparation, during which data is reviewed for suspicious, invalid or missing cases; variables; and data values. In the next step, data analysis, the data is examined, tested, explored and transformed. Patterns are identified, hypotheses are tested and information is extracted. Next comes reporting, where data is summarized, put in tables and charts, and ready for

consumption. The last step is deployment. Data, reports and procedures are distributed to users globally, with interaction and access managed centrally. At this point, the process starts all over again to make sure you maintain your competitive advantage.

When your organization follows this analytical path, you can benefit in a number of ways. You can gain a better understanding of your current situation, consider the most appropriate options, predict what is likely to happen next and take actions to improve outcomes.

Potential benefits to your business include the ability to acquire, grow and retain customers; reduce costs; minimize risk and fraud; and improve efficiency.

IBM® SPSS® Modeler products are offered in an open technology platform that offers rich options for both data science and business users. On the following pages, take a look at how you can turn your organization's data into a strategic asset that can give your business a competitive advantage.



IBM SPSS Modeler

Broader analytics, faster results, flexible deployment

With IBM SPSS Modeler software, you can:

- Access, manipulate and model virtually any type of data, from practically any source, from within a single, intuitive interface.
- Deployment options include using IBM SPSS Modeler as part of IBM Watson™ Studio as well as standalone deployment.
- Leverage open source scripting languages such as R and Python to further extend your capabilities.
- Proactively and repeatedly identify revenue opportunities and reduce costs.
- Boost productivity through collaboration and improve scalability and performance with server-based options.
- Accelerate your paths from model exploration to deployment.

IBM SPSS Modeler software is a powerful, versatile data and text mining workbench that helps you gain significant insight from your data. Its breadth and depth of techniques allows you to build predictive models more easily, efficiently and rapidly. Organizations use IBM SPSS Modeler software to help improve business outcomes in critical areas such as customer relationship management, marketing, fraud and risk mitigation, inventory management and resource planning.

Big data analytics

IBM SPSS Modeler software is designed to easily connect with nearly all databases, including Hadoop distributions, allowing models to be deployed and scored with greater speed and efficiency. Models have even been designed specifically with big data applications in mind to better support scalable analytics for organizations looking for solutions to their big data challenges.

Gain faster, more accurate insight

IBM SPSS Modeler software is designed with simplicity of execution in mind, so you don't need to know how to code to enjoy its benefits.

IBM SPSS Modeler software's intuitive graphical interface makes it easier to visualize the data science process and how it all fits together. This interface is designed to allow you to easily access your data, regardless of its format. Data can come from flat files and databases as well as big data environments such as Hadoop. You can also read from and write to an IBM Cognos® Analytics environment as well as

Automotive company uses predictive analytics to fulfill uptime commitments on its trucks

A leading global manufacturer of trucks, buses, construction equipment and marine and industrial engines, this company also provides complete solutions for financing and service. Haulage firms are a significant customer segment for the company.

An unplanned standstill is one of the most problematic issues that can affect a haulage firm. Apart from the inconvenience for the driver, it creates extra costs for repairs, lost transport revenue and, in the worst-case scenario, damage to customer reputation. One important prerequisite for reducing the number of unplanned standstills is to be able to predict maintenance needs and to tailor servicing for each individual truck.

The growing business need for predictive maintenance to fulfill the uptime commitments on company trucks led to the decision to invest in a new predictive analytics platform using IBM SPSS Modeler. "We have traditionally been very good in reactive reporting. Now, the big area that has been the trigger for analytics has been within the predictive maintenance area," says the business intelligence solution architect.

By monitoring a truck's usage and the current status of the vehicle's various key components, it is possible to predict component failure while truck is on the road or in the shop—before a breakdown can occur. The company applies machine learning techniques to automatically discover patterns and learn from the vast amount of data it collects. The company can now identify necessary parts and provide repair instructions even before the truck arrives for service, helping reduce diagnostic time by up to 70 percent and lowering repair time by more than 20 percent.

integrate with IBM InfoSphere® technology, so you can gain even more predictive intelligence and integrate it into your business processes.

IBM SPSS Modeler software's powerful automated data preparation helps you get your data ready for analysis in a single step, while automated modeling takes the guesswork out of choosing an analysis method by applying relevant techniques with a single click. Data preparation and modeling are presented in an interactive, highly visual viewer that makes models easier to understand and communicate. After models are created, you can apply them to new data to help identify cross-sell and up-sell opportunities proactively, manage customer retention strategically, and identify fraud before it happens.

 [Watch the video →](#)

 [Learn more →](#)

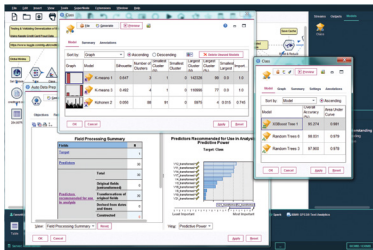
Server-based option for enterprise-level scalability

Sometimes you need to go beyond the desktop to analyze big data or deploy data mining as a part of enterprise architecture. Although the client provides the interface and the ease of use that makes modeling accessible to users of varied abilities, the optional server provides client/server architecture, in-database mining, SQL pushback and batch processing. In-database mining exposes modeling algorithms in operational databases, such as SQL, IBM DB2® and Oracle databases, and allows you to leverage database resources, vastly improving performance while enabling you to use the additional algorithms IBM SPSS Modeler software provides. SQL pushback allows you to push data transformations and some key models directly into your operational database, improving performance significantly by leveraging the power of the database for high-volume, mission-critical tasks.

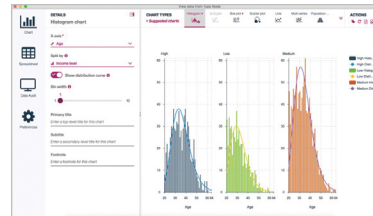
IBM SPSS Modeler software can scale from a single desktop to an enterprise deployment. Users can interact with the software on their desktop, use SPSS Modeler software's client and server architecture to apply predictive analytics to enterprise data, or integrate predictive analytics into business processes with an enterprise deployment. Organizations that lack the infrastructure or IT resources to implement their software locally have options as well. IBM SPSS Modeler software can also be deployed to the cloud, helping your organization achieve maximum flexibility.

Choose the IBM SPSS Modeler software that best meets your analysis needs:

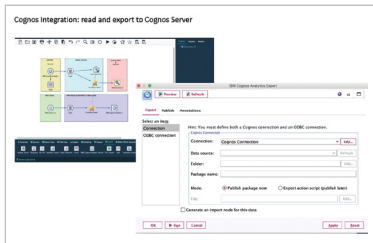
- **IBM SPSS Modeler Subscription** offers the capabilities you need to access several structured data sets, analyze data using a vast range of algorithms and deploy the results using your desktop. It offers a flexible monthly payment option without an annual commitment.
- **IBM SPSS Modeler Professional** software includes the tools needed to leverage your structured data—such as behaviors and interactions tracked in your CRM systems, demographics, purchasing behavior and sales data in a multi-user environment.
- **IBM SPSS Modeler Premium** software goes beyond the analysis of structured data to include unstructured data such as web activity, blog content, customer feedback, emails, articles and more to create more accurate predictive models.
- **IBM SPSS Modeler Gold** software includes all the capabilities of IBM SPSS Modeler Premium software with the added ability to build and deploy predictive models directly into business processes to enable people or systems to make the right decisions each time. It combines predictive analytics with rules, scoring and optimization within an organization's processes and operational systems to deliver recommended actions at the point of impact.



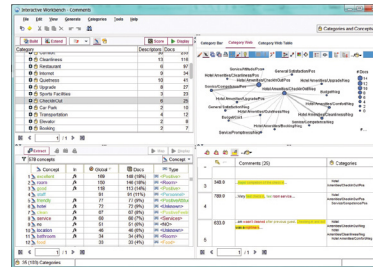
IBM SPSS Modeler software offers leading ease-of-use features such as automatic data preparation and automatic modeling, making it easier to build models that leverage single or multiple (ensemble) techniques.



IBM SPSS Modeler software includes advanced, interactive visualization for models that use a single technique or ensemble models that combine techniques, making modeling results easier to understand and communicate.



IBM SPSS Modeler software introduces the ability to read from and write to an IBM Cognos Analytics environment directly, helping ensure that business analysts can leverage the same view of data across their analytical work and can easily integrate predictive intelligence into their Cognos Analytics dashboards directly from the IBM SPSS Modeler interface.



IBM SPSS Modeler Premium software adds the ability to integrate unstructured data (text) into modeling efforts; analysts can extract concepts and associated sentiment, visualize relationships between concepts and sentiment, and structure unstructured data for use in predictive models from a single interface.

IBM SPSS Modeler Subscription

The powerful capabilities you need, with a flexible monthly subscription

With IBM SPSS Modeler Subscription, you can:

- Access and export data—without data size limits—from flat files, spreadsheets, major relational databases and Hadoop.
- Transform data automatically into the best format for modeling.
- Choose from 40+ base machine learning algorithms.
- Automate modeling, using a single run to test multiple methods.
- Use R, Python and Python scripting.
- Explore geographic data with geospatial analytics.
- Extend capabilities with SQL optimization to push back the processing of many common data processing steps to relational databases and Hadoop.

Whether you're a single user or a small company looking for an affordable way to put data science to work for you—or you're trying to introduce the power of predictive analytics to a larger organization in a scalable way—IBM SPSS Modeler Subscription can be the step to take. It's an ideal solution for individual projects and smaller teams, with flexible per-user, per month pricing and the scalability to expand to more users and business areas as your business grows.

The monthly subscription model means no annual commitments. Plus, you always have the latest updates, and your subscription is easy to renew. Getting started is easy. Simply sign up online for a free 30-day trial and download the fully functional desktop application. To keep using the software after the trial, just make your first subscription payment—there's no need to reinstall the software.

Increasing competitiveness by increasing efficiency and reducing costs

Analytics helps consumer products firm optimize warehouse stock levels

For businesses operating in saturated markets, cost-efficiency is a powerful tool to combat pressure on margins. This was true for a German producer of sparkling wines, spirits and a variety of other drinks under a number of brands. A market leader in Germany, the company wanted to maximize profitability by boosting working capital and cutting warehouse costs—but inefficiencies and poor accuracy in its sales forecasting process presented a challenge.

In particular, a largely manual approach to running rolling forecasts was not scaling. Says the company's Chief Information Officer (CIO), "To solve the challenge, we looked for a way to augment our people's expertise with predictive analytics. The aim was to enable them to build more accurate, granular rolling forecasts, and use those insights to produce the optimal quantity of each SKU."

The company started with a proof of concept project, extracting data from ERP and CRM systems and using IBM SPSS Modeler to construct models to forecast sales over a three-month period, analysing a minimum of 12 data points for each product. The initial work enabled the firm to refine models and learn how to treat price promotions in forecasting.

Today the company employs a hybrid approach that combines statistical models with existing manual methods to deliver the most accurate results for decision makers—in effect combining both art and science into the process. The result is a 22 percent increase in the accuracy of selected sales forecasts, enabling the company to reduce stock levels without increasing the risk of stockouts. That also boosts working capital by freeing up funds previously tied up in overstock.

IBM SPSS Modeler Professional

Make better decisions through predictive intelligence

IBM SPSS Modeler Professional software is designed to make it easy to discover insights in your data so that you can proactively take advantage of opportunities and mitigate risks. This intuitive, visual data-mining workbench enables business users and professional analysts alike to quickly and easily access, prepare and model structured data to solve business challenges. A graphical interface walks you through virtually every step of the modeling process. By interacting with streams, analysts and business users can collaborate in adding business knowledge to the data-mining process, explore data more deeply and achieve faster results.

Improve predictive insights with geospatial analytics

IBM SPSS Modeler Professional software offers advanced geospatial analytics tools for modeling and making predictions based on many types of location and time data, including geographic coordinates, addresses, zip codes and natural resources. Organizations in virtually every industry are turning to new data types and new forms of analysis to remain competitive. Geospatial data, also referred to as location data or spatial data, enables businesses to add the context of time and location to traditional data, so they can see changes over time and exactly where those changes are taking place. The geospatial analytics tools in IBM SPSS Modeler Professional software help improve predictive insights by accounting for time and space in predictive models, so businesses can accurately forecast events at a specific location for virtually any future point in time.

Streamline the data-mining process

SPSS Modeler Professional software 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.

Easily create and evaluate sophisticated models

Choose from an array of prebuilt algorithms to help you create models visually and intuitively.

Classification algorithms: Make predictions or forecasts using techniques such as decision tree, neural networks, logistic regression, time series, support vector machines and Cox regression. Leverage automatic classification modeling for binary and numeric outcomes.

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.

View models interactively and apply advanced analytical and visualization techniques to help you understand and communicate the results of your analysis. Then efficiently deploy insight and predictive models on a scheduled basis or in real time.

Optimize your current information technologies

With its open and scalable architecture, SPSS Modeler software is designed to make the best use of your existing IT infrastructure. It integrates with your existing systems 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, multithreading, server clustering and in-database scoring help conserve resources, deliver results faster and reduce overall IT costs.

IBM SPSS Modeler Premium

Help improve model accuracy with structured and unstructured data

Additionally, IBM SPSS Modeler Premium software 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 predefined 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 data sets, 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.

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

With IBM SPSS Modeler Premium software text analytics capabilities, you can have a better understanding of your organization, the environment in which you operate, your customers and other stakeholders. The interactive, visual environment of IBM SPSS Modeler software uses advanced linguistic technologies and natural language processing to rapidly process unstructured text data. From this text, it extracts and organizes the key concepts.

The customizable, industry-specific text analysis packages enable you to analyze relevant terms and phrases in addition to acronyms, emoticons and slang in the right context. Interactive graphs help you explore and display text data and patterns for instant analysis. You can create hierarchical categorization structures and include them as inputs to a predictive model to yield better and more focused decisions and results. Predefined categories, such as hierarchies, annotations and keyword descriptors can be imported to categorize initial unstructured data so you can organize concepts more logically and in greater detail.

Promoting safer workplaces and reducing risk for member cities

Analytics helps municipal self-insurance pool extract insights from historical data

A large municipal self-insurance pool, this organization provides liability, workers' compensation, environmental, property, and earthquake coverage, as well as risk management training and loss control services, to nearly 120 cities and other governmental agencies across the state.

The organization acts as a central resource for administering claims and providing risk management services, enabling members to gain cost savings and operational efficiencies. In fulfilling this role, the organization was collecting massive quantities of data that essentially sat on the shelf.

Thinking about new ways to enhance its ability to identify and mitigate risk, the organization realized that analytics presented a huge opportunity. It deployed IBM SPSS Modeler to explore its historical claims data, and reveal eye-opening insights.

"We knew that workers' compensation indemnity claims result in high payouts for our members, so we decided to look at those first," says one of the organization's financial analysts. "Using SPSS Modeler, we found that there was a high concentration of indemnity claims in one particular department within public works. Together with our members, we can now examine working practices in that department, highlighting and addressing high-risk behaviors to help reduce the number and cost of claims."

The solution also enables the organization to measure the effectiveness of training. Offering analytics as a service to help municipalities improve safety, reduce claims and save money, the organization is well positioned to win new members and act as a leader in improving working practices.

IBM SPSS Modeler Gold

Enterprise-scale predictive intelligence
deployed on premises or in the cloud

With IBM SPSS Modeler Gold, organizations can:

- Access a variety of data sources, such as data warehouses, databases, Hadoop distributions or flat files, to find hidden patterns in the data.
- Deliver predictive, resource-aware and strategically aligned decisions to people and systems at the point of impact almost instantly.
- Put analytics in the hands of whoever will benefit from it, regardless of their statistical or analytical background.
- Solve your business problems with a single platform that is designed to handle simple, descriptive analysis all the way to the most complex optimization problems.
- Analyze vast amounts of data in less time while fully using your existing IT investments with in-database performance and minimized data movement.
- Take advantage of an open platform that can be deployed in most environments and integrated with other IBM solutions to bridge the gap between analytics and action.

Take a more in-depth look at IBM SPSS Modeler Gold

IBM SPSS Modeler Gold software is a powerful predictive analytics platform that is designed to bring predictive intelligence to decisions made by individuals, groups, systems and your enterprise. It scales from desktop deployments to integration with operational systems to provide you with a range of advanced algorithms and techniques. Applying these techniques to decisions can result in rapid ROI and can enable organizations to proactively and repeatedly reduce costs while increasing productivity.

A versatile data mining workbench, IBM SPSS Modeler Gold software can help organizations build predictive models quickly and intuitively without the need for programming. It enables users to discover patterns and trends in structured and unstructured data using a visual interface supported by statistical algorithms. With IBM SPSS Modeler Gold software, users can take advantage of predictive analytics, rules and scoring to get the most out of virtually every customer interaction to grow revenue, increase customer loyalty, mitigate risk of fraud and lower costs.

IBM SPSS Modeler Gold software supports the interplay between analysts and business users by applying analytics to everyday business issues. It deploys operational systems to maximize the impact of predictive analytics across the organization by supporting:

- Data preparation and manipulation
- Automated data modeling
- Text analytics
- Integration with IBM technology

Enterprise benefits

- Enables users to discover patterns and trends in structured and unstructured data using a distinctive visual interface supported by statistical algorithms.
- Allows the ability to analyze structured (age, price, product, location) or unstructured (text, emails, social media) data.
- Offers an extensive range of analytics to address a variety of business problems, such as descriptive analytics and advanced algorithms, and includes automated modeling, text analytics, entity analytics, social network analysis, decision management and optimization.
- Uses predictive analytics, rules and scoring to get the most out of virtually every customer interaction to help increase revenues and loyalty, prevent threats, mitigate risk of fraud and reduce cost.
- Supports the interaction between expert analysts and business users.

A comprehensive predictive analytics platform

IBM SPSS Modeler Gold software is an extensive analytics platform that enables organizations to build predictive models to enhance decision making. IBM SPSS Modeler Gold software combines:

- IBM Collaboration and Deployment Services
Deployment Manager
- IBM SPSS Collaboration and Deployment Services
Base Server
- IBM SPSS Collaboration and Deployment Services—
Automation

IBM SPSS Modeler software serves as the analytic engine by allowing predictive models to be built on structured and unstructured data with its GUI. IBM SPSS Collaboration and Deployment Services software is the deployment and execution platform that provides analytical process governance for analytical assets, scoring of models and integration capabilities with other IBM products and into operational systems.

IBM SPSS Modeler Gold software uses predictive analytics, rules and scoring to get the most out of every customer interaction to help increase revenues and loyalty, prevent threats, mitigate risk of fraud and reduce cost. It helps support the interaction between expert analysts and business users and enables organizations to go beyond traditional data mining by combining and integrating a variety of analytic techniques into business systems.

“Instead of spending two weeks developing a new model ourselves, and then repeating much of that work every time we wanted to update it, we purchased IBM SPSS Modeler, which provides a wide range of prebuilt, proven models that we can apply immediately to whatever data-set we choose, and refresh at the touch of a button.”

– Financial Analyst, large municipal self-insurance pool

IBM SPSS Collaboration and Deployment Services

Manage analytical assets and share results efficiently and securely

IBM SPSS Collaboration and Deployment Services enables the deployment and sharing of predictive analytics across the enterprise. The solution provides centralized, secure storage of analytical assets, and advanced capabilities for management and control of predictive analytic processes. It also offers advanced mechanisms for delivering analytics results to users. Analysts can work together and share critical business information more easily with others. It protects the business by storing analytical assets in one place, and automatically tracking changes made to them. Analysts can easily publish information, allowing users easy access to it whenever they need it.

Share assets and enhance productivity

IBM SPSS Collaboration and Deployment Services provides a secure way to share, integrate and re-use analytical assets. It is an analytically aware, central, searchable repository specifically designed to work with the IBM SPSS portfolio. Models, statistical analyses, syntax or process steps, documented best practices and other assets can be stored in the repository, which analysts can access from within the tools they already use. They can then re-use relevant assets on other projects or in other parts of the organization.

Build strong protection for intellectual property

IBM SPSS Collaboration and Deployment Services provides safeguards against loss of, or damage to, analytical assets. Administrators can schedule appropriate backups of the central repository, control access and create a complete audit trail, showing who has modified specific assets, when and how. This helps to make the governance of analytical assets more systematic and reliable. Automatic version control capabilities ensure that only the current, approved versions of analytics are used in production. You can also integrate with existing authentication systems for authentication and single sign-on capabilities.

Publish results broadly and easily

IBM SPSS Collaboration and Deployment Services ensures that the right people have visibility into analytical results by making it easy for analysts to publish results. It also makes it easy for business users to access results formerly reserved only for analysts.

With IBM SPSS Collaboration and Deployment Services, organizations can efficiently deliver decision-making guidance to people and systems everywhere. There is no longer a need to email reports to groups of users or for someone to initiate a process manually when an analysis is completed. Instead, processes can be configured to let business users know that there are reports available for viewing.

It's also possible for analytic professionals to develop interfaces that enable business to run specific business-oriented analytics when they need to and then publish these results without performing any special coding. Through a standard web browser, business users can access and interact with these interfaces and display exactly the information they need, when they need it. Sales managers, for example, can customize how they view their data in a variety of ways. Marketing managers can use their reports to develop customized, cost-effective campaigns and achieve repeatable results—all without having an analyst dedicated to ad hoc reporting.

 [Learn more →](#)

IBM SPSS Analytic Server

Make predictive analytics easier for big data

IBM SPSS Analytic Server provides an open, integrated data-centric architecture that uses big data systems and is scalable to problems of almost any size. IBM SPSS Analytic Server enables IBM SPSS Modeler to use big data as a source for predictive modeling. Together, they can provide an integrated predictive analytics platform using data from Hadoop distributions and Spark applications. Users can discover insights in data stored in these big data frameworks as well as traditional relational database management systems (RDBMS) without the need to write complex code or scripts.

Predictive analytics includes running numerous iterations of the most relevant data to attain optimal results. When data sets grow to hundreds of millions of records, the time needed to perform such tasks becomes prohibitive. And the volume of data can be so large that it's impractical if not impossible to use traditional analytics platforms. As a result, accessing big data can be complex and requires specific skills and scripts.

The SPSS Modeler and SPSS Analytic Server solution is designed to support unstructured or semi-structured predictive analytics for big data. It enables the combination of structured and unstructured data to improve model accuracy. As a result, even business users can glean insights from big data quickly and efficiently without complex programming packages or scarce expertise.

Distributing analytics processing into Hadoop environments

Hadoop is a Java™-based programming framework that supports the processing of large data sets in a distributed computing environment. With the SPSS Modeler and SPSS Analytic Server combination, analytics processing is distributed into the Hadoop environment, eliminating the need to move data to the analytics and optimizing performance on large volumes of varied data. The solution supports multiple Hadoop platforms, including Apache Hadoop, Hortonworks HDP, Cloudera CDH and MapR.

Accessing data stored across the organization

Not all data used for modeling is found in Hadoop. In many cases, it resides in traditional RDBMSs or flat files. The combined IBM SPSS Modeler and IBM SPSS Analytic Server solution enables users to access data from a Hadoop distribution and combine it with external data from other sources. As a result, users can also discover insights through federation with a traditional RDBMS. In addition, the SPSS Modeler interface can be used to add data sources to a Hadoop distribution. This and the accessibility to data wherever it is stored can help business users get a complete picture and analyze all the data available to them that is relevant to the problem they are trying to solve.

Powering faster machine learning and real-time processing

IBM SPSS Modeler integrates with Apache Spark, an open source engine built specifically for data science, helping to simplify algorithm development and accelerate analytics results. Using Spark, data scientists can better extract value from big data, conduct deeper analyses and deliver results faster, all while reducing the time and effort required for coding. Spark accelerates analytics on Hadoop, delivering the necessary speed and agility to data scientists and developers working with big data at scale.

This capability is particularly important for real-time stream processing and machine learning, which requires iterative computation, a task that is normally prohibitively time consuming with massive data. Spark provides in-memory computing, which speeds the analytical process and enables users to take advantage of big data machine learning algorithms available both natively within SPSS Modeler and through the Spark machine learning library, MLlib.

 [Learn more →](#)

Boosting forecasting accuracy with analytics

Automotive aftermarket company anticipates consumer demand at the store level

A large auto glass specialist in the United States operating both bricks-and-mortar outlets and mobile stores, this company helps more than five million customers every year to repair or replace their vehicle glass. To get customers back on the road more quickly, the company aimed to allocate resources to where they are needed most. To do so, the company needed to forecast demand accurately at each store.

The company replaced spreadsheet-based enterprise planning with a collaborative approach based on IBM SPSS Modeler, IBM Cognos 8 Business Intelligence and Cognos TM1. Greater local involvement boosts forecast accuracy, driving smarter decision-making, while powerful automation cuts budget cycles by 40 percent, saving 450 person-hours a year. Sales forecast accuracy is an impressive 99.8 percent.

“We have only scratched the surface of what we can do with SPSS and TM1, and we have so much more to look forward to. Our journey with IBM started about ten years ago, and today the future is looking brighter than ever,” says the director of financial planning and analysis.

Why choose the IBM SPSS Modeler

IBM SPSS Modeler offers an easy-to-use, flexible and scalable platform that makes predictive analytics accessible to users with all skill levels and outfits projects of all sizes and complexity to help you and your organization to find new opportunities, improve efficiency and minimize risk. With solutions for individuals or the enterprise, you can take advantage of:



Power. With support for the complete data science cycle, from understanding to deployment, the portfolio offers a vast library of machine learning algorithms and capabilities such as text analytics and geospatial analysis.



Ease of use. An intuitive drag-and-drop interface guides users through the analytics processes. Extend your capabilities with open source technologies such as R, Python, Spark and Hadoop—with or without coding.



Scalability. Designed to grow with your business needs, options range from desktop analysis production system integration for a large-scale enterprise.



Flexibility. Choose the licensing option that best fits your needs—desktop, monthly subscription or enterprise perpetual license—or deploy as part of IBM Watson Studio.

For more information

To learn more about the IBM SPSS Modeler portfolio, contact your IBM representative or IBM Business Partner, or visit: ibm.com/analytics/data-science/predictive-analytics/spsstatistical-software

Take advantage of our active and growing open-source community, where you can find resources to help you expand the use of IBM predictive analytics software. Resources include blogs, videos, tutorials and an extensive library of more than 6,000 predictive extensions to help you take advantage of popular programming languages such as R, Python and Java. Join the community at: developer.ibm.com/predictiveanalytics

© Copyright IBM Corporation 2018

IBM Corporation
New Orchard Road
Armonk, NY 10504

Produced in the United States of America
December 2018

IBM, the IBM logo, ibm.com, Cognos, DB2, InfoSphere, SPSS and Watson are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at ibm.com/legal/copytrade.shtml

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

The client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

The client is responsible for ensuring compliance with laws and regulations applicable to it. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the client is in compliance with any law or regulation.