IBM SPSS Statistics: What’s New

New and enhanced features to accelerate, optimize and simplify data analysis

Analytics plays an increasingly important role in helping your organization achieve its objectives. The IBM® SPSS® Statistics family delivers the core capabilities needed for end-to-end analytics. To ensure that the most advanced techniques are available to a broader group of analysts and business users, we have made enhancements to the features and capabilities of IBM SPSS Statistics Base and its many specialized modules.

IBM SPSS Statistics 23 continues to offer advanced analytics through new data analysis techniques, enhanced features and output, and improved accessibility. This release focuses on increasing the analytic capabilities of the software through:

- **Extending the value of big data** — Uncover hidden causal relationships among large numbers of time series using Temporal Causal Modeling (TCM).
- **Geospatial analytics** — Explore the relationship between data elements that can be tied to a location to gain deeper insights.
- **Embedding analytics into the enterprise** — Apply the next generation of web reports, R programmability enhancements and improved accessibility.

Our suite of statistical software comes in three editions: IBM SPSS Statistics Standard, IBM SPSS Statistics Professional and IBM SPSS Statistics Premium. These editions group essential features and functionality, and are a convenient way to ensure you have the capabilities you need to generate the insights your organization requires for effective decision-making.
Discover causal connections between time series
Uncover hidden causal relationships among large numbers of time series using Temporal Causal Modeling (TCM), specialized methods that can be used to discover key temporal relationships in time series data. Now you can feed a large number of time series into TCM and find out which series are causally related. This differs from traditional time series modeling, where you must explicitly specify the predictors for a given time series, by automatically determining the best predictors for each target series.

For example, stock price data are temporal in nature—a value of a set of variables depends on the values of another set of variables at several points in the past. TCM can be used to analyze the change in value and discover the underlying causes.

Geospatial analytics: Add a new dimension to your data analysis
In SPSS Statistics 23, you can explore the relationship between data elements tied to a geographic location to reveal much deeper insights about people and events.

Find trends over time and space using Spatio-Temporal Prediction (STP)
With this new release, you can fit linear models for measurements taken over time at locations in 2D and 3D space, enabling organizations to predict “hot” areas and how those areas change over time. Some of the business applications for this feature include building management and branch performance analysis.

Figure 1: This image shows the causal connections associated with the top (best-fitting) 10 models in a temporal causal model system. A business that monitors key performance indicators (KPIs) and also tracks data on controllable metrics referred to as levers wants to determine the causal connections between the levers and the KPIs, so it can understand which levers affect which KPIs. The company also wants to know if there are causal connections between the KPIs themselves.

Figure 2: In this map and graph, SPSS Statistics displays the point density (Kernel Density Estimation) of the selected regions over time.
Create association rules that incorporate geospatial attributes

The Generalized Spatial Association Rule (GSAR) helps discover associations between spatial and non-spatial attributes. With GSAR, you can use historical data such as location, type of event and the time the event happened to describe the occurrences of events. This can be valuable in applications such as crime pattern analysis and epidemic surveillance.

Programmability enhancements for R

Users will have the ability to develop and test R programs using a full-featured, integrated R development environment within SPSS Statistics. In addition, R functions that use SPSS Statistics functionality can be written with command syntax from within R, and results can be returned to R.

New and improved categorical principal component analysis (CATPCA)

Categorical PCA capabilities have been expanded to include:

- Non-parametric bootstrapping for more stable estimates
- Clustering of cases in addition to variables
- New rotation options for better convergence
- An easier way to use continuous variables

Figure 3: As seen here, GSAR can be used to predict where crime is most likely to occur based on demographic information for the region, as shown in this map.

Figure 4: This is a plot of categorical principal component loadings with confidence intervals.
**Next generation of web output**
SPSS Statistics web reports have been completely redesigned, with more interactivity and functionality and web server support.

**Faster performance with bulk export/insert**
SPSS Statistics 23 can help you bulk load data for faster performance. The application writes the data to a text data file, then the bulk loader script writes the text data back to the database, providing superior performance when handling large datasets.

**Other V23 enhancements**
SPSS Statistics includes these additional capabilities:

- Enables users of the new release of Stata 13 to import, read and write Stata 9-13 files within SPSS Statistics.
- Supports enterprise users who need to access the software with their employee identification badges and badge readers

**System requirements**
Requirements vary by platform. Find the system requirements for your operating system at: ibm.com/software/analytics/spss/products/statistics/requirements

*Figure 5: Bulk loading options are available in the Export to Database Wizard.*
Reasons to upgrade your SPSS Statistics software
If you’re using an earlier version of IBM SPSS Statistics, you’ll gain all of these time-saving features — and many more — when you upgrade to the newest version.

Added in IBM SPSS Statistics 22
• Interactive output on smart devices
• Presentation ready output
• More powerful monte carlo simulation with support for simulating strings, automatic linear modeling and heat-maps
• Integration with IBM Cognos TM1
• Improved performance and scalability using SPSS Statistics server with SQL pushback
• Python plug-in as part of main installation
• Search for, download and install available programming extensions from within extensions
• Use a simplified method to specify user-defined estimands in SPSS Amos.
• Benefit from improved logging support for Enterprise Standard in the Platform Standards.
• Enable other applications to read/write encrypted Statistics data files with i/o dll.
• Generate pivot table output for non-parametric procedures.

Added in IBM SPSS Statistics 21
• Monte Carlo simulation for building better models to uncertain inputs
• Ability to compare two data files or data sets to ensure the data values and records are compatible
• Password protection of data and output files (encryption) to prevent others from seeing confidential information
• Improved and faster file merging
• Enhanced pivot tables with new features such as easier navigation and sorting operations
• Easier model specification in IBM SPSS Statistics Amos for structural equation modeling (SEM)
• Programming in SPSS Statistics with a Java plug-in
• Ability to import IBM® Cognos® Business Intelligence data into SPSS Statistics for additional analysis
• Exporting of output to Microsoft Excel 2007/2010
• Better scalability and performance with load balancing with IBM SPSS Collaboration and Deployment Services
• Single sign-on between SPSS Statistics client and server
• For security restricted environments, the option to run SPSS Statistics server as a non-root user on UNIX/LINUX
• Compression of data files for increased storage space
• Version compatibility to support new clients with old servers and vice versa, such as client V21 and server V20 OR client V20 and server V21)
Added in IBM SPSS Statistics 20
- Pre-built map templates and support for ESRI files in SPSS Statistics Base
- Faster pivot table output
- GLMM procedure in IBM SPSS Advanced Statistics can be run with ordinal values
- Non-graphical, programmatic method for specifying models in SPSS Amos
- Run SPSS Statistics Server jobs offline by disconnecting the SPSS Statistics client.
- Compress temporary files created by the sort procedure within SPSS Statistics Server to save disk space when sorting large files.

Added in IBM Statistics 19
- Automatic Linear Models (ALM) gives non-specialist users the tools to build powerful linear models automatically and predict numerical outcomes
- Generalized Linear Mixed Models (GLMM), in SPSS Advanced Statistics, lets you create more accurate models for predicting non-linear outcomes based on hierarchical/nested data or data that includes repeated measures
- Several new capabilities in IBM SPSS Direct Marketing
- Faster performing tables in SPSS Statistics Base
- More than a dozen performance and ease-of-use enhancements to the Syntax Editor, available in all products in the SPSS Statistics family
- The Statistics portal provides internal or external users with interactive online access to analysis (requires SPSS Statistics Server and SPSS Collaboration and Deployment Services)
- Compiled transformations is a feature in SPSS Statistics Server that improves the performance of SPSS Statistics programs that execute a large number of data transformations
- Analysts using SPSS Statistics Base can score customer data, access pre-built models and interface directly with data in Salesforce.com
- Pivot columns and crosstabulations in SPSS Statistics Base and IBM SPSS Custom Tables
- Work with smaller and sparse datasets on Linux and Mac operating systems in IBM SPSS Exact Tests
- Run SPSS Statistics Base Server on IBM System z (requires SuSE Linux)

Added in IBM SPSS Statistics 18
- Prepare data in a single step using the new Automated Data Preparation feature
- New Nonparametric tests in SPSS Statistics Base
- Post computed categories in SPSS Custom Tables
- SPSS Direct Marketing module
- SPSS Bootstrapping module
- Rule checking on Secondary SPC Charts
- IBM SPSS Statistics Developer
- Ability to view significance tests in the main results table in SPSS Custom Tables
- Interactive Model Viewer on Two-Step Cluster Analysis and Automated Data Preparation procedures
- Improved display of large pivot tables
- Improved performance on procedures within SPSS Statistics Base Server for Frequencies, Descriptives, Crosstabs
- Support for 64-bit hardware on desktop for Windows and Mac
- Support for Snow Leopard on Mac OS X 10.6
Added in IBM SPSS Statistics 17

- Syntax Editor with features to make it easier to create, test and deploy syntax jobs
- Switch user interface language
- Mac OS X and Linux platforms can connect clients to SPSS Statistics Server
- Updated plug-ins for Python, .NET and R
- Support for graphic packages written in R
- Create user-defined interfaces for existing procedures and user-defined procedures with Custom Dialog Builder
- Call front-end Python scripts or scripting APIs explicitly from within back-end Python programs
- Support for Predictive Enterprise View, a common data interface that can be defined once and used by all IBM SPSS analytic tools
- Administrative enhancements in SPSS Statistics Server, including optimized multithreading, visualization support and a “file in use” message to reduce errors in data created by more than one person writing to an SPSS Statistics file at the same time
- Read access to SPSS Statistics data files as an ODBC/JDBC data source, allowing these files to be read using SQL Codebook procedure to automatically describe the dataset
- Spell-checking of long text strings
- IBM SPSS EZ RFM module
- Multiple imputation of missing data in IBM SPSS Missing Values module
- Regularization methods: Ridge regression, the Lasso, Elastic Net in IBM SPSS Categories
- Model selection methods: 632(+), bootstrap, cross validation (CV), in IBM SPSS Categories
- Nearest Neighbor analysis in SPSS Statistics Base
- Median transformations function in COMPUTE procedure
- Option to use aggressive versus conservative rounding in COMPUTE procedure
- Create new variables that contain the values of existing variables from preceding or subsequent cases
- Graphboard integration, enabling users of SPSS Statistics products to deploy new or customer graph templates created in the new IBM SPSS Visualization Designer stand-alone module
- Wrapping and shrinking of wide tables in Word and PowerPoint
- Smartreader feature to allow the viewing and pivoting of SPSS Statistics output

Added in IBM SPSS Statistics 16

- Mac and Linux versions of SPSS Statistics
- Several multithreaded procedures for improved performance and scalability
- In the Data Editor: ability to customize variable view, spell checking for value labels and variable labels, sort by variable name, type, format, etc.
- Unicode support
- Import/export Excel 2007 data
- Syntax to change string length and basic data type of existing variables
- Creation of value labels and missing values on strings of any length
- Ability to set a permanent default working directory
- IBM SPSS Neural Networks module
- Complex Samples Cox Regression added to SPSS Complex Samples
- Latent Class Analysis in SPSS Amos
- Partial Least Squares regression
- Support for R algorithms
- Find and Replace feature in the Output Viewer

For more information on the latest version of SPSS Statistics, go to: ibm.com/software/analytics/spss/products/statistics.
About IBM Business Analytics
IBM Business Analytics software delivers data-driven insights that help organizations work smarter and outperform their peers. This comprehensive portfolio includes solutions for business intelligence, predictive analytics and decision management, performance management and risk management.

Business Analytics solutions enable companies to identify and visualize trends and patterns in such areas as customer analytics that can have a profound effect on business performance. They can compare scenarios; anticipate potential threats and opportunities; better plan, budget and forecast resources; balance risks against expected returns and work to meet regulatory requirements. By making analytics widely available, organizations can align tactical and strategic decision making to achieve business goals. For more information, see ibm.com/business-analytics.

Request a call
To request a call or to ask a question, go to ibm.com/business-analytics/contactus. An IBM representative will respond to your inquiry within two business days.