



**StatSoft®**

*data analysis • data mining • quality control • web-based analytics*

---

**New Features  
In  
STATISTICA 8**

March 31, 2007

---

**U.S. Headquarters: StatSoft, Inc. • 2300 E. 14th St. • Tulsa, OK 74104 • USA • (918) 749-1119 • Fax: (918) 749-2217 • info@statsoft.com • www.statsoft.com**

Australia: StatSoft Pacific Pty Ltd.  
Brazil: StatSoft South America  
Bulgaria: StatSoft Bulgaria Ltd.  
Czech Rep.: StatSoft Czech Rep. s.r.o.  
China: StatSoft China

France: StatSoft France  
Germany: StatSoft GmbH  
Hungary: StatSoft Hungary Ltd.  
India: StatSoft India Pvt. Ltd.  
Israel: StatSoft Israel Ltd.

Italy: StatSoft Italia srl  
Japan: StatSoft Japan Inc.  
Korea: StatSoft Korea  
Netherlands: StatSoft Benelux BV  
Norway: StatSoft Norway AS

Poland: StatSoft Polska Sp. z o.o.  
Portugal: StatSoft Ibérica Lda  
Russia: StatSoft Russia  
Spain: StatSoft Ibérica Lda

S. Africa: StatSoft S. Africa (Pty) Ltd.  
Sweden: StatSoft Scandinavia AB  
Taiwan: StatSoft Taiwan  
UK: StatSoft Ltd.

<b>ENHANCEMENTS TO EXISTING FUNCTIONALITY.....</b>	<b>4</b>
<b>System Level Functionality.....</b>	<b>4</b>
Projects .....	4
Resume Analysis / Rerun Analysis.....	4
ToolTip improvements .....	4
Wrapping variable name text.....	4
Multiple Version support.....	4
64-bit support.....	4
Application Background Window .....	4
<b>Microsoft Office Integration.....</b>	<b>4</b>
Microsoft Excel .....	4
Microsoft Word .....	4
<b>New Statistics and Statistics Enhancements.....</b>	<b>5</b>
Summaries of analyses .....	5
Descriptives, robust statistics, non-factorial breakdowns, non-factorial tables .....	5
T-test, independent by groups.....	5
Correlation matrices .....	5
Box-Cox Transformation.....	5
Botched Designs .....	5
Optimal Split Plot.....	5
Nonparametrics.....	5
Multivariate Quality Control .....	5
Attribute Agreement Gage Attribute .....	6
Capability Analysis - Attribute Gage.....	6
Capability ratios for True Position.....	6
Capability Analysis - Gage Capability .....	6
Process Capability ISO/DIN (Time dependent distribution model) .....	6
Statistics of block data.....	6
<b>New Graphs and Graph Enhancements .....</b>	<b>6</b>
Bag Plot .....	6
Bi Plot.....	6
Wafer maps.....	6
Overlaid Contour plots .....	7
Overlaid probability plots.....	7
Scatterplot with Error Bars .....	7
Copy graph by dragging .....	7
Drag graphs into Word documents .....	7
Graph Data as an array .....	7
Locking Graphs .....	7
Merging Graphs.....	7
Resolution control for exported graphs .....	7
Rotation of objects by CTRL + PAGE UP/PAGE DOWN .....	7
Scrolling legends .....	7
Selecting items from legend .....	7

**U.S. Headquarters: StatSoft, Inc. • 2300 E. 14th St. • Tulsa, OK 74104 • USA • (918) 749-1119 • Fax: (918) 749-2217 • info@statsoft.com • www.statsoft.com**

Australia: StatSoft Pacific Pty Ltd.  
Brazil: StatSoft South America  
Bulgaria: StatSoft Bulgaria Ltd.  
Czech Rep.: StatSoft Czech Rep. s.r.o.  
China: StatSoft China

France: StatSoft France  
Germany: StatSoft GmbH  
Hungary: StatSoft Hungary Ltd.  
India: StatSoft India Pvt. Ltd.  
Israel: StatSoft Israel Ltd.

Italy: StatSoft Italia srl  
Japan: StatSoft Japan Inc.  
Korea: StatSoft Korea  
Netherlands: StatSoft Benelux BV  
Norway: StatSoft Norway AS

Poland: StatSoft Polska Sp. z o.o.  
Portugal: StatSoft Ibérica Lda  
Russia: StatSoft Russia  
Spain: StatSoft Ibérica Lda

S. Africa: StatSoft S. Africa (Pty) Ltd.  
Sweden: StatSoft Scandinavia AB  
Taiwan: StatSoft Taiwan  
UK: StatSoft Ltd.

Styles and Axis settings .....	8
<b>By Group .....</b>	<b>8</b>
A new flexible By Group facility in all processes .....	8
<b>Workbooks .....</b>	<b>8</b>
Workbook multi-item display .....	8
Save as HTML.....	8
Advanced Structure Storage .....	8
Locking Workbook User Interface .....	8
Scrolling, selecting/expanding when dragging .....	8
<b>Reports .....</b>	<b>8</b>
Additional report options provide more flexibility .....	8
<b>Data Filtering, Cleaning, and Handling Outliers.....</b>	<b>9</b>
Filter duplicate cases .....	9
Filter Sparse Data .....	9
Process Relatively Invariant Variables .....	9
Recode Outliers .....	9
Process Missing Data.....	9
k-Nearest Neighbor Interpolation .....	9
<b>Spreadsheet .....</b>	<b>9</b>
Fast Text Importer User Interface.....	9
Cell references in Spreadsheet formulas.....	9
New Spreadsheet formulas .....	9
Changes to CrossTab (unstack) object.....	10
Increase Sort keys to 500.....	10
Analysis Case Selection Conditions .....	10
Text Labels .....	10
Missing Data.....	10
Merge Improvements.....	10
JMP 6 files supported .....	11
<b>Enhancements to WebSTATISTICA Server and Integration of the Desktop and Server Versions .....</b>	<b>11</b>
Integration with <i>STATISTICA</i> Server .....	11
<i>WebSTATISTICA</i> Performance monitors.....	11
<b>Enhancements to STATISTICA Data Miner .....</b>	<b>11</b>
Data Miner Recipe (DMR) – a new user interface .....	11
<b>Development System/Integration Tools .....</b>	<b>11</b>
OLE DB Provider for <i>STATISTICA</i> Spreadsheets .....	11
Graph Data as an array .....	11
Locking Workbook User Interface .....	11
Additional supported formulas .....	12
Logging interface.....	12
<i>STATISTICA</i> Query completion events .....	12
<i>WebSTATISTICA</i> Performance monitors.....	12

**U.S. Headquarters: StatSoft, Inc. • 2300 E. 14th St. • Tulsa, OK 74104 • USA • (918) 749-1119 • Fax: (918) 749-2217 • info@statsoft.com • www.statsoft.com**

Australia: StatSoft Pacific Pty Ltd.  
Brazil: StatSoft South America  
Bulgaria: StatSoft Bulgaria Ltd.  
Czech Rep.: StatSoft Czech Rep. s.r.o.  
China: StatSoft China

France: StatSoft France  
Germany: StatSoft GmbH  
Hungary: StatSoft Hungary Ltd.  
India: StatSoft India Pvt. Ltd.  
Israel: StatSoft Israel Ltd.

Italy: StatSoft Italia srl  
Japan: StatSoft Japan Inc.  
Korea: StatSoft Korea  
Netherlands: StatSoft Benelux BV  
Norway: StatSoft Norway AS

Poland: StatSoft Polska Sp. z o.o.  
Portugal: StatSoft Ibérica Lda  
Russia: StatSoft Russia  
Spain: StatSoft Ibérica Lda

S. Africa: StatSoft S. Africa (Pty) Ltd.  
Sweden: StatSoft Scandinavia AB  
Taiwan: StatSoft Taiwan  
UK: StatSoft Ltd.

<b>NEW PRODUCTS.....</b>	<b>13</b>
<b>SANN - STATISTICA AUTOMATED NEURAL NETWORKS .....</b>	<b>13</b>

## Enhancements to Existing Functionality

### System Level Functionality

#### Projects

You can save a *STATISTICA* Project file that will contain the “frozen” state of the current work session, including the locations of all object windows on the screen. Upon opening the project file, the respective analysis can be resumed where it was last saved. Options are provided to save the recorded macro associated with the analysis.

#### Resume Analysis / Rerun Analysis

Output that is placed in a workbook will also contain a recorded analysis macro for that output. You can choose to rerun or resume the analysis and can rerun the analysis using a different data set, providing a convenient manner to pick up work from where a project was stopped previously or to perform the same analysis using different data sets.

#### ToolTip improvements

- ToolTips will display the full text of a spreadsheet cell.
- ToolTips will display the numeric value of a text label.

#### Wrapping variable name text

Variable names will break at the nearest character for wrapping if there is no word-break boundary

#### Multiple Version support

Multiple, currently licensed versions of *STATISTICA* (including multiple languages) can be installed on the same computer enabling you to switch back and forth between them as needed.

#### 64-bit support

*STATISTICA* can be installed on 64-bit Windows operating systems.

#### Application Background Window

You can select from several standard background styles or can use an image of your choice.

### Microsoft Office Integration

- Microsoft Excel documents can be opened within *STATISTICA* directly (rather than importing them into the *STATISTICA* file format).
- Microsoft Word documents can be used directly as the output report type.

- Graphs can be placed in *STATISTICA* documents or external documents such as Microsoft Word and PowerPoint by dragging,

## **New Statistics and Statistics Enhancements**

### **Summaries of analyses**

Summary, integrated displays with both graphics and statistics available for applicable analyses

### **Descriptives, robust statistics, non-factorial breakdowns, non-factorial tables**

New statistics including Winsorized means, Trimmed means, Grubb's Test for outliers, Coefficient of Variation, Confidence interval for sample standard deviation, and Percent of valid observations have been added. (See also other tests for outliers in the *Data Cleaning and Filtering* section below.)

### **T-test, independent by groups**

Confidence interval for estimates is now available.

### **Correlation matrices**

Means and standard deviations can now be added to square correlation matrices.

### **Box-Cox Transformation**

Options to perform Box-Cox transformation have been added, including statistics, histograms, and normal probability plots. The results can be written back to the input spreadsheet.

### **Botched Designs**

Support for experiments that include “botched” runs (where factors are not set to the specified levels) in 2-level designs ( $2^{(k-p)}$ , screening, etc.)

### **Optimal Split Plot**

Optimal split-plot designs can be generated and analyzed

### **Nonparametrics**

Limit for nonparametric methods increased to 1-million cases (note that this change has been introduced only to support unusual applications, as more powerful and sensitive parametric tests are preferable to be used with large samples)

### **Multivariate Quality Control**

The new functionality supported includes:

- Hotelling T-Square charts
- Multiple Stream (Group) Hotelling T-Square charts
- Multivariate Exponentially Moving Average (MEWMA) charts
- Multivariate Cumulative Sum (MCUSUM) charts
- Generalized Variance Charts

## Attribute Agreement Gage Attribute

Results now include Fleiss kappa, Kohen kappa, Kendall's coefficient of concordance, and Kendall's correlation coefficient. Graphical results include plots to denote the Assessment agreement by appraiser or by item, with confidence intervals. An assessment disagreement table (that shows how far an appraiser disagrees with the standard) when the response is binary and the standard is known is also available.

## Capability Analysis - Attribute Gage

The results of this analysis includes a summary table and graphs and tests as described in the AIAG Measurement Systems manual

## Capability ratios for True Position

This method evaluates coordinate pairs compared to the “true values” for a center point. Output includes capability ratios and scatterplots for each X/Y pair of variables, with circles and lines that indicate the Tolerance region as specified.

## Capability Analysis - Gage Capability

The variability of gages is compared against the specification limits of the parts.

## Process Capability ISO/DIN (Time dependent distribution model)

Process Capability ISO/DIN (Time dependent distribution models) enable you to compute process capability indices consistent and in compliance with DIN (Deutsche Industrie Norm) 55319 (see Deutsches Institut fuer Normung e.V., 2002) and ISO 21747 (see ISO, 2006). These methods are applicable for situations where consecutive samples of observations are taken from an ongoing production line, and the goal is to estimate the process capability with respect to one or more quality dimensions measured in those samples. Specifically, these standards summarize different distribution “models” for how 1) the observations within each sample can be distributed, 2) how the moments (locations, dispersions, etc.) of consecutive samples can be distributed (e.g., normal, non-normal) over “time,” and 3) how best to estimate the process capability based on the resultant distribution (given the distribution of measurements within sample, and across samples/time).

## Statistics of block data

Statistics of block data functions will return the results in a separate output spreadsheet.

## *New Graphs and Graph Enhancements*

### Bag Plot

A type of scatter plot using a bivariate generalization of Tukey's univariate box-and-whisker plot to identify distributions (and outliers) in two-dimensional space.

### Bi Plot

A graph used in multivariate process analysis (MSPC) to identify relations between underlying variables and extracted factors to aid in effective root cause analysis.

### Wafer maps

A graphical method for rendering defect data in two-dimensional planes common in the semiconductor industry (wafer plots).

## Overlaid Contour plots

A type of contour plot that identifies “joint-regions” where multiple range requirements for Z are simultaneously satisfied. A graph useful for “visually optimizing” multivariate systems.

## Overlaid probability plots

Probability plots (Normal, Quantile-Quantile, Probability-Probability) can display multiple variables in one graph.

## Scatterplot with Error Bars

A type of scatterplot where points with the same X values are combined and a box/whisker (error bars) are generated at the points

## Copy graph by dragging

Graphs can be copied by dragging.

## Drag graphs into Word documents

Graphs can be placed in *STATISTICA* documents or external documents such as Word and PowerPoint by dragging.

## Graph Data as an array

Graph data can be accessed in automation as a data array.

## Locking Graphs

Graph files can be locked to prevent users from modifying them.

## Merging Graphs

Graphs can be merged by dragging or copy/paste onto the destination graph.

## Resolution control for exported graphs

You can specifying the resolution (dots per inch) for TIFF, GIF, PNG, JPG, and BMP files when saving a graph via the *Save As* menu

## Rotation of objects by CTRL + PAGE UP/PAGE DOWN

Graph objects (such as arrows) can be rotated with finer control using Windows-standard CTRL + PAGE UP and CTRL + PAGE DOWN.

## Scrolling legends

A graph legend that is too long to be displayed in the graph window can be scrolled so it can all be seen. The scrolling toolbar buttons enable you to scroll the legend either line-by line, or move all the way to the top/bottom of the legend

## Selecting items from legend

Point markers and fit lines can be selected from the graph legend.

---

U.S. Headquarters: StatSoft, Inc. • 2300 E. 14th St. • Tulsa, OK 74104 • USA • (918) 749-1119 • Fax: (918) 749-2217 • info@statsoft.com • www.statsoft.com

Australia: StatSoft Pacific Pty Ltd.  
Brazil: StatSoft South America  
Bulgaria: StatSoft Bulgaria Ltd.  
Czech Rep.: StatSoft Czech Rep. s.r.o.  
China: StatSoft China

France: StatSoft France  
Germany: StatSoft GmbH  
Hungary: StatSoft Hungary Ltd.  
India: StatSoft India Pvt. Ltd.  
Israel: StatSoft Israel Ltd.

Italy: StatSoft Italia srl  
Japan: StatSoft Japan Inc.  
Korea: StatSoft Korea  
Netherlands: StatSoft Benelux BV  
Norway: StatSoft Norway AS

Poland: StatSoft Polska Sp. z o.o.  
Portugal: StatSoft Ibérica Lda  
Russia: StatSoft Russia  
Spain: StatSoft Ibérica Lda

S. Africa: StatSoft S. Africa (Pty) Ltd.  
Sweden: StatSoft Scandinavia AB  
Taiwan: StatSoft Taiwan  
UK: StatSoft Ltd.

## Styles and Axis settings

Global graphic options control what portion of a graph definition is used when applying a graph style.

## By Group

- The By Group user interface has been significantly improved in Version 8 by adding a By Groups tab to most analysis specification dialogs. Note that the script-based By Group user interface implemented in Version 7 is still supported for backward compatibility.
- Output for By Group analysis can be directed into a single folder with labels for the respective group names or a folder for each group.
- Labels that contain the group information can be included in the header or title of all output.

## Workbooks

### Workbook multi-item display

The multi-item display allows for editing of the objects displayed.

### Save as HTML

Workbooks can now be saved as HTML displaying the workbook hierarchy of objects (tree view) in a browser window.

### Advanced Structure Storage

Options allow for more efficient storage of workbooks and provide significantly increased capacity.

### Locking Workbook User Interface

An automation feature that protects a worksheet from inadvertent changes from the user manipulating the workbook tree control.

### Scrolling, selecting/expanding when dragging

When dragging workbook items within the workbook tree, the workbook will automatically expand folders when hovered over and will allow scrolling around the tree by dragging to the bottom or top of the tree.

## Reports

New report options enable you to:

- Save *STATISTICA Enterprise* Reports as HTML.
- Specify that *STATISTICA* Spreadsheets in a *STATISTICA* Report should be saved in HTML tables when the *STATISTICA* Report is saved as HTML.
- Insert/delete objects into/from a report via automation. The special object properties added for *STATISTICA Enterprise* reports are also available via automation.
- Print each individual spreadsheet in a report as an object or as a full spreadsheet on separate pages. This option can be set on a per-spreadsheet basis.
- Define the layout of each page (either landscape or portrait)

---

U.S. Headquarters: StatSoft, Inc. • 2300 E. 14th St. • Tulsa, OK 74104 • USA • (918) 749-1119 • Fax: (918) 749-2217 • info@statsoft.com • www.statsoft.com

Australia: StatSoft Pacific Pty Ltd.  
Brazil: StatSoft South America  
Bulgaria: StatSoft Bulgaria Ltd.  
Czech Rep.: StatSoft Czech Rep. s.r.o.  
China: StatSoft China

France: StatSoft France  
Germany: StatSoft GmbH  
Hungary: StatSoft Hungary Ltd.  
India: StatSoft India Pvt. Ltd.  
Israel: StatSoft Israel Ltd.

Italy: StatSoft Italia srl  
Japan: StatSoft Japan Inc.  
Korea: StatSoft Korea  
Netherlands: StatSoft Benelux BV  
Norway: StatSoft Norway AS

Poland: StatSoft Polska Sp. z o.o.  
Portugal: StatSoft Ibérica Lda  
Russia: StatSoft Russia  
Spain: StatSoft Ibérica Lda

S. Africa: StatSoft S. Africa (Pty) Ltd.  
Sweden: StatSoft Scandinavia AB  
Taiwan: StatSoft Taiwan  
UK: StatSoft Ltd.

## ***Data Filtering, Cleaning, and Handling Outliers***

### **Filter duplicate cases**

Also called “de-duping,” using selected variables; duplicates (exact matches on all selected variables) can be removed from a data set.

### **Filter Sparse Data**

Cases or variables that exceed user-specified percentage of missing data can be removed from a data set.

### **Process Relatively Invariant Variables**

Variables that have a standard deviation below a user-specified threshold level can be removed from a data set.

### **Recode Outliers**

Using outlier tests such as Categorical, Normal, Grubb's, Percentile, and Tukey, outliers can be recoded to a user-defined value including a specified value, missing data, the mean, a specific percentile, or a boundary value.

### **Process Missing Data**

Missing data can be recoded to a specific value, mean, or median. Additionally, the variables that have missing data removed can be flagged if the variable contains a threshold percentage of missing data, and if flagged, the entire variable can optionally be removed.

### **k-Nearest Neighbor Interpolation**

Missing data can be replaced with values that are estimated (imputed) from the data using k-nearest neighbors, substituting the missing data with the “average” value for cases that are similar to – or near – the case in question, with respect to all values that are observed (not missing).

## ***Spreadsheet***

### **Fast Text Importer User Interface**

Faster options for importing text are now supported. The interface enables you to preview the data, intelligently selects the delimiters and types, and allows each individual variable to be selected in the preview, and the specific type defined.

### **Cell references in Spreadsheet formulas**

The new formulas can reference individual cells in spreadsheets in the same way as values of individual cells can be referenced in typical spreadsheet applications such as Microsoft Excel.

### **New Spreadsheet formulas**

New formulas for row-based calculations include:

- VCUR - the current variable number
- VREF - variable reference at runtime
- LAG - variable offset from the current row minus the lag parameter
- DIF - difference of the current value minus the variable specified by the parameter
- CUSUM - cumulative sum of the specified variable

---

**U.S. Headquarters: StatSoft, Inc. • 2300 E. 14th St. • Tulsa, OK 74104 • USA • (918) 749-1119 • Fax: (918) 749-2217 • info@statsoft.com • www.statsoft.com**

Australia: StatSoft Pacific Pty Ltd.  
Brazil: StatSoft South America  
Bulgaria: StatSoft Bulgaria Ltd.  
Czech Rep.: StatSoft Czech Rep. s.r.o.  
China: StatSoft China

France: StatSoft France  
Germany: StatSoft GmbH  
Hungary: StatSoft Hungary Ltd.  
India: StatSoft India Pvt. Ltd.  
Israel: StatSoft Israel Ltd.

Italy: StatSoft Italia srl  
Japan: StatSoft Japan Inc.  
Korea: StatSoft Korea  
Netherlands: StatSoft Benelux BV  
Norway: StatSoft Norway AS

Poland: StatSoft Polska Sp. z o.o.  
Portugal: StatSoft Ibérica Lda  
Russia: StatSoft Russia  
Spain: StatSoft Ibérica Lda

S. Africa: StatSoft S. Africa (Pty) Ltd.  
Sweden: StatSoft Scandinavia AB  
Taiwan: StatSoft Taiwan  
UK: StatSoft Ltd.

- DATA - value specified by the specific variable reference and row index
- NCASES - current number of cases in the spreadsheet
- NVARs - current number of variables in the spreadsheet
- Contains - a text function that searches the first argument for occurrence of the second argument
- Munger - a text function used to find substrings or insert or delete substrings
- Trim - a text function that will return text with all the trailing blanks removed
- Word - a text function that extracts the nth word from a character string
- Item - a text function that extracts the nth word from a character string where consecutive delimiter characters are seen as two delimiters with an empty word in between
- Hex - returns the hex representation of the argument
- Repeat - a text function that creates and returns a string that is the first argument repeated the number of times specified by the second argument
- CaseState - a number of new functions related to case states has been added.
- Floor - largest integer less than or equal to its argument
- Ceiling - the smallest integer greater than or equal to its argument
- Round - rounds the first argument to the number of decimal places given by the second argument
- FACT - returns the factorial of the integer of the argument
- COMBIN - returns the number of combinations of n things taken k at a time
- PERMUT - returns the number of permutations of n things taken k at a time
- CHOOSE - This function returns the ith value of an array
- MATCH - If value = Vi, returns Ri, else returns Rdefault, if any, else returns missing
- IN - This function returns a non-zero value (TRUE) if the first argument is equal to any of the remaining arguments, i.e., TRUE if val is in the list of values following. All arguments can be constants, variable references, or expressions. This works with text or numbers.
- Date Functions - fourteen functions for working with dates, months, days of the week, and time

## Changes to CrossTab (unstack) object

Options added to treat missing data in code or case ID variables as valid values and to preserve the relative variable order in output.

## Increase Sort keys to 500

Sort can be initiated using up to 500 sort keys

## Analysis Case Selection Conditions

To prevent common errors, the analysis case selection conditions must now be confirmed before this option becomes active.

## Text Labels

Warning displayed when text values are entered in a spreadsheet and text labels are not enabled for that variable.

## Missing Data

The default missing data code has been changed to -999999998

## Merge Improvements

Options are now offered to preserve the order of the cases when files are merged, resulting in data in the same relative order as in the first source spreadsheet.

## JMP 6 files supported

## ***Enhancements to WebSTATISTICA Server and Integration of the Desktop and Server Versions***

### **Integration with STATISTICA Server**

A new user interface is now available in all desktop versions of *STATISTICA* supporting a seamless integration between with the Server (Web) *STATISTICA* version. For example, you can now off-load time consuming tasks to the server with a touch of a button, monitor the progress of server computations from the desktop, and easily move *STATISTICA* objects between the computers.

### **WebSTATISTICA Performance monitors**

Performance monitor counters can be accessed using standard system tools.

### ***Enhancements to STATISTICA Data Miner***

- Data Mining menu made more accessible
- Interactive trees available with deployment
- Variables are considered by the tree-building algorithm one-by-one
- Limit for nonparametric methods increased to 1 million cases
- Lift table for all types added to Rapid Deployment

### **Data Miner Recipe (DMR) – a new user interface**

A recipe-like step-by-step process to guide you through the data mining process:

- Connect to data
- Modify/prepare data
- Perform computations
- Review results
- Save/Deploy
- Project files can be created and saved at any step of the process and Data Miner Recipe Projects can be deployed to *STATISTICA Enterprise* for scoring.

### ***Development System/Integration Tool***

### **OLE DB Provider for STATISTICA Spreadsheets**

Standard SQL can be used to query *STATISTICA* Spreadsheets from OLE DB compliant applications.

### **Graph Data as an array**

Graph data can be accessed in automation as a data array.

### **Locking Workbook User Interface**

An automation feature that protects a worksheet from inadvertent changes from the user manipulating the workbook tree control.

## Additional supported formulas

New formulas for row-based calculations include:

- VCUR - the current variable number
- VREF - variable reference at runtime
- LAG - variable offset from the current row minus the lag parameter
- DIF - difference of the current value minus the variable specified by the parameter
- CUSUM - cumulative sum of the specified variable
- DATA - value specified by the specific variable reference and row index
- NCASES - current number of cases in the spreadsheet
- NVARs - current number of variables in the spreadsheet
- Contains - a text function that searches the first argument for occurrence of the second argument
- Mungger - a text function used to find substrings or insert or delete substrings
- Trim - a text function that will return text with all the trailing blanks removed
- Word - a text function that extracts the nth word from a character string
- Item - a text function that extracts the nth word from a character string where consecutive delimiter characters are seen as two delimiters with an empty word in between
- Hex - returns the hex representation of the argument
- Repeat - a text function that creates and returns a string that is the first argument repeated the number of times specified by the second argument
- CaseState - a number of new functions related to case states has been added.
- Floor - largest integer less than or equal to its argument
- Ceiling - the smallest integer greater than or equal to its argument
- Round - rounds the first argument to the number of decimal places given by the second argument
- FACT - returns the factorial of the integer of the argument
- COMBIN - returns the number of combinations of n things taken k at a time
- PERMUT - returns the number of permutations of n things taken k at a time
- CHOOSE - This function returns the ith value of an array
- MATCH - If value = Vi, returns Ri, else returns Rdefault, if any, else returns missing
- IN - This function returns a non-zero value (TRUE) if the first argument is equal to any of the remaining arguments, i.e., TRUE if val is in the list of values following. All arguments may be constants, variable references, or expressions. This works with text or numbers.
- Date Functions - fourteen functions for working with dates, months, days of the week, and time

## Logging interface

An external logging interface that can be accessed from external programs is provided. Additional command line parameters can be used to enable this interface to write out logging information to a report window or an external text file

## STATISTICA Query completion events

Spreadsheets now have new events that get called for attached queries. This is useful for background queries (those you specify with Query.Refresh(True)), though it will be called for background and foreground queries. One function is called when query completes; another is called if there is an error in the query. The events can be hooked up to in a standard fashion from either SVB (using the WithEvents declaration) or from external languages that support connection points.

## WebSTATISTICA Performance monitors

Performance monitor counters can be accessed using standard system tools.

---

U.S. Headquarters: StatSoft, Inc. • 2300 E. 14th St. • Tulsa, OK 74104 • USA • (918) 749-1119 • Fax: (918) 749-2217 • info@statsoft.com • www.statsoft.com

Australia: StatSoft Pacific Pty Ltd.  
Brazil: StatSoft South America  
Bulgaria: StatSoft Bulgaria Ltd.  
Czech Rep.: StatSoft Czech Rep. s.r.o.  
China: StatSoft China

France: StatSoft France  
Germany: StatSoft GmbH  
Hungary: StatSoft Hungary Ltd.  
India: StatSoft India Pvt. Ltd.  
Israel: StatSoft Israel Ltd.

Italy: StatSoft Italia srl  
Japan: StatSoft Japan Inc.  
Korea: StatSoft Korea  
Netherlands: StatSoft Benelux BV  
Norway: StatSoft Norway AS

Poland: StatSoft Polska Sp. z o.o.  
Portugal: StatSoft Ibérica Lda  
Russia: StatSoft Russia  
Spain: StatSoft Ibérica Lda

S. Africa: StatSoft S. Africa (Pty) Ltd.  
Sweden: StatSoft Scandinavia AB  
Taiwan: StatSoft Taiwan  
UK: StatSoft Ltd.

# New Products

## **SANN - STATISTICA AUTOMATED NEURAL NETWORKS**

*STATISTICA Automated Neural Networks* is a high-performance and easy-to-use application that replaces the *STATISTICA Neural Networks* product offered with *STATISTICA Version 7*. It includes the latest technologies and state-of-the-art algorithms for building and deploying neural network models. Neural network modules include:

- Regression (for implementing regression with non-sequential data).
- Classification (for implementing classification analysis with non-sequential data).
- Time Series Regression (for implementing regression with time series data).
- Time Series Classification (for implementing classification with time series data).
- Cluster Analysis (using Kohonen maps)

---

**U.S. Headquarters: StatSoft, Inc. • 2300 E. 14th St. • Tulsa, OK 74104 • USA • (918) 749-1119 • Fax: (918) 749-2217 • info@statsoft.com • www.statsoft.com**

Australia: StatSoft Pacific Pty Ltd.  
Brazil: StatSoft South America  
Bulgaria: StatSoft Bulgaria Ltd.  
Czech Rep.: StatSoft Czech Rep. s.r.o.  
China: StatSoft China

France: StatSoft France  
Germany: StatSoft GmbH  
Hungary: StatSoft Hungary Ltd.  
India: StatSoft India Pvt. Ltd.  
Israel: StatSoft Israel Ltd.

Italy: StatSoft Italia srl  
Japan: StatSoft Japan Inc.  
Korea: StatSoft Korea  
Netherlands: StatSoft Benelux BV  
Norway: StatSoft Norway AS

Poland: StatSoft Polska Sp. z o.o.  
Portugal: StatSoft Ibérica Lda  
Russia: StatSoft Russia  
Spain: StatSoft Ibérica Lda

S. Africa: StatSoft S. Africa (Pty) Ltd.  
Sweden: StatSoft Scandinavia AB  
Taiwan: StatSoft Taiwan  
UK: StatSoft Ltd.