

Statgraphics Centurion Version 17 Enhancements

Version 17 of Statgraphics Centurion contains many significant enhancements to the program. These enhancements include:

1. 32 new statistical procedures.
2. Additional capabilities in 20 existing statistical procedures.
3. New user interface features such as a session log.

This document describes each of the enhancements.

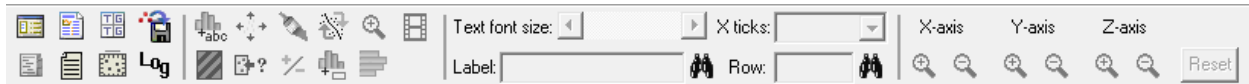
Analysis Procedures	4
Audit Trails and Electronic Signatures	7
Bivariate Density Statlet.....	8
Cause-and-Effect Diagram	11
Computer-Generated Designs	12
Correspondence Analysis.....	13
Crosstabulation	14
Curve-Fitting Statlet.....	15
Data Columns.....	17
Demographic Map	19
Demographic Map Brushing Statlet.....	21
Demographic Map Visualizer Statlet	23
Design of Experiments Wizard.....	24
Deviation Dashboard Statlet.....	26
Exponential Smoothing Statlet	28
Factor Analysis	29
Factorial Repeated Measures ANOVA	30
Frequency Histogram.....	31
G Chart	33
Graphics Options.....	35
Added text.....	35

Adding Objects	35
Axis tickmarks	36
Brushing scatterplots	37
Dashed and Dotted Lines	39
Overriding attributes	39
Recording videos.....	40
Text fonts	40
Zooming and panning	40
Hypothesis Tests (Describe).....	41
Hypothesis Tests (Compare).....	42
Item Reliability Analysis	43
Interactive Histogram Statlet.....	45
Kriging	46
Kriging Statlet.....	49
MIL-STD-105E Acceptance Sampling for Attributes	50
MIL-STD-414 Acceptance Sampling for Variables.....	52
MIL-STD-1916 Acceptance Sampling for Attributes	54
MIL-STD-1916 Acceptance Sampling for Variables.....	56
Multiple Sample Comparison.....	59
Multiple Variable Analysis.....	61
Multivariate Visualizer	63
Nonlinear Regression.....	69
Normal Probability Plot.....	70
One-Dimensional Visualizer	72
One Variable Analysis	74
Oneway ANOVA	75
Oneway Repeated Measures ANOVA.....	76
Open-High-Low-Close Candlestick Plot Statlet	77
P Chart.....	79
P' Chart.....	80
Power Transformations Statlet.....	81
Preferences	82

Principal Components Analysis.....	84
Probabilistic Fractal.....	85
Process Capability Analysis Statlet.....	86
Process Map.....	87
Reliability Demonstration Test Plan.....	89
Sample Size Determination Statlet	90
StatFolio Alerts.....	92
StatFolio Passwords	94
StatLog	95
Surface-Fitting Statlet	96
T Chart.....	98
Tabulation	100
Ternary Plot.....	101
Three-Dimensional Visualizer	102
Two-Dimensional Visualizer	103
U Chart	104
U' Chart	105
Miscellaneous	106
Data.....	106
Statgraphics Operators	106
Toolbar Shortcuts.....	106

Analysis Procedures

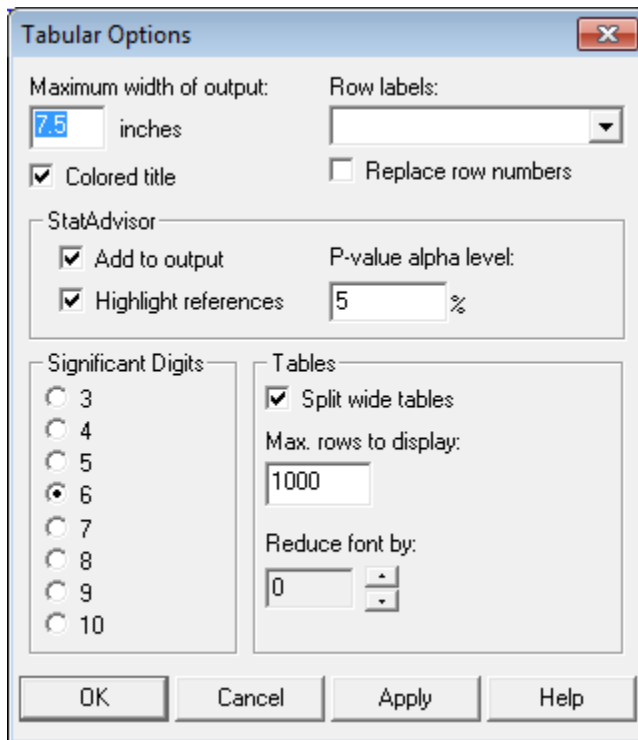
Significant enhancements have been made to the function of all statistical procedures. One important enhancement involves the analysis toolbar, which has been redesigned:



New buttons on the toolbar include:



Tabular options: active after a table has been maximized:



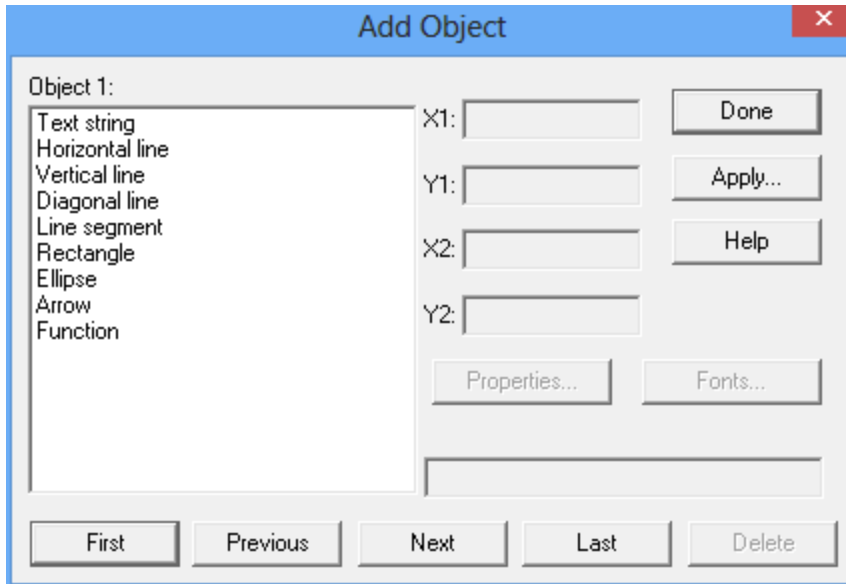
These options temporarily override the system-wide settings specified on the *Preferences* dialog box available under *Edit* on the main menu.




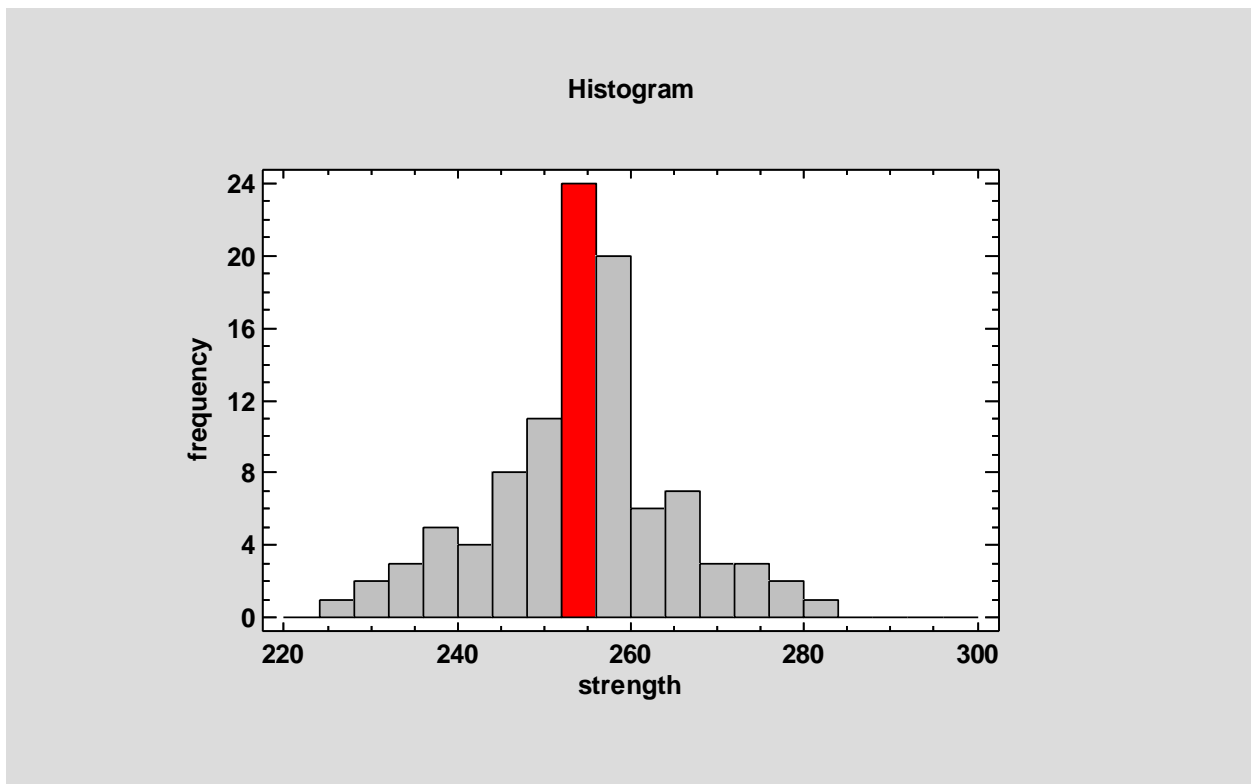
Save to logfile: copies all visible tables and graphs to the StatLog system logfile.



Add object: adds an additional object such a line, arrow or shape to a graph.



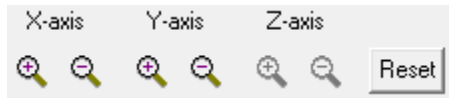
 **Override attributes:** overrides the default attributes of an object, such as the color of a single bar in a frequency histogram.



Change text font size: increases or decreases the size of all text in a table or graph.

X ticks:

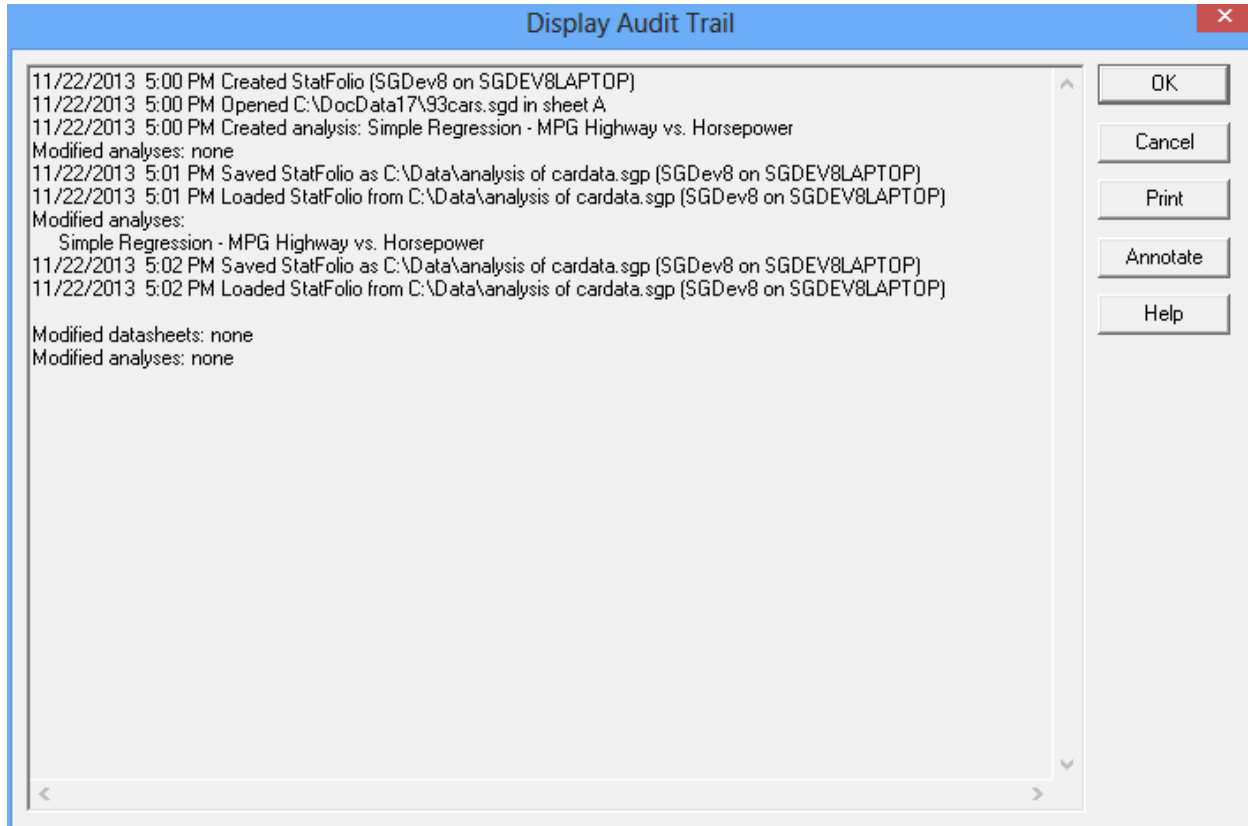
Rotate X-axis tickmarks: changes the orientation of the X-axis tickmarks.



Zoom: zooms a graph in or out along a particular axis.

Audit Trails and Electronic Signatures

Beginning with Version 16.2, Statgraphics Centurion saved as part of a StatFolio an audit trail that tracks important events in the life of a StatFolio. Version 17 adds an electronic signature to the audit trail, created from the username and computer name of the active Windows session.



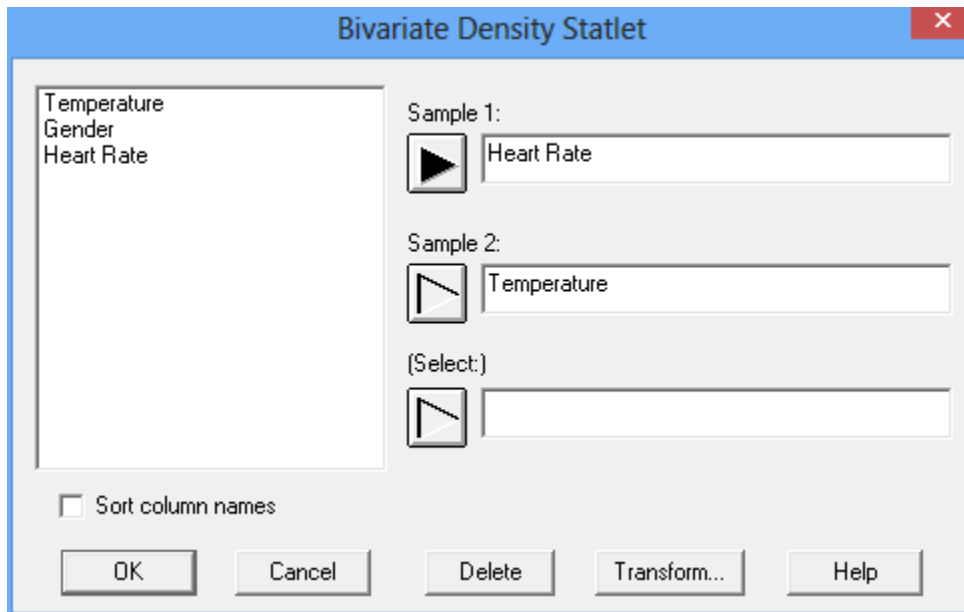
The signature is recorded each time any of the following events occurs:

1. A new StatFolio is created.
2. An existing StatFolio is loaded.
3. The current StatFolio is saved.

Version 17 also allows additional text to be added to the audit trail by pressing the *Annotate* button on the above window.

Bivariate Density Statlet

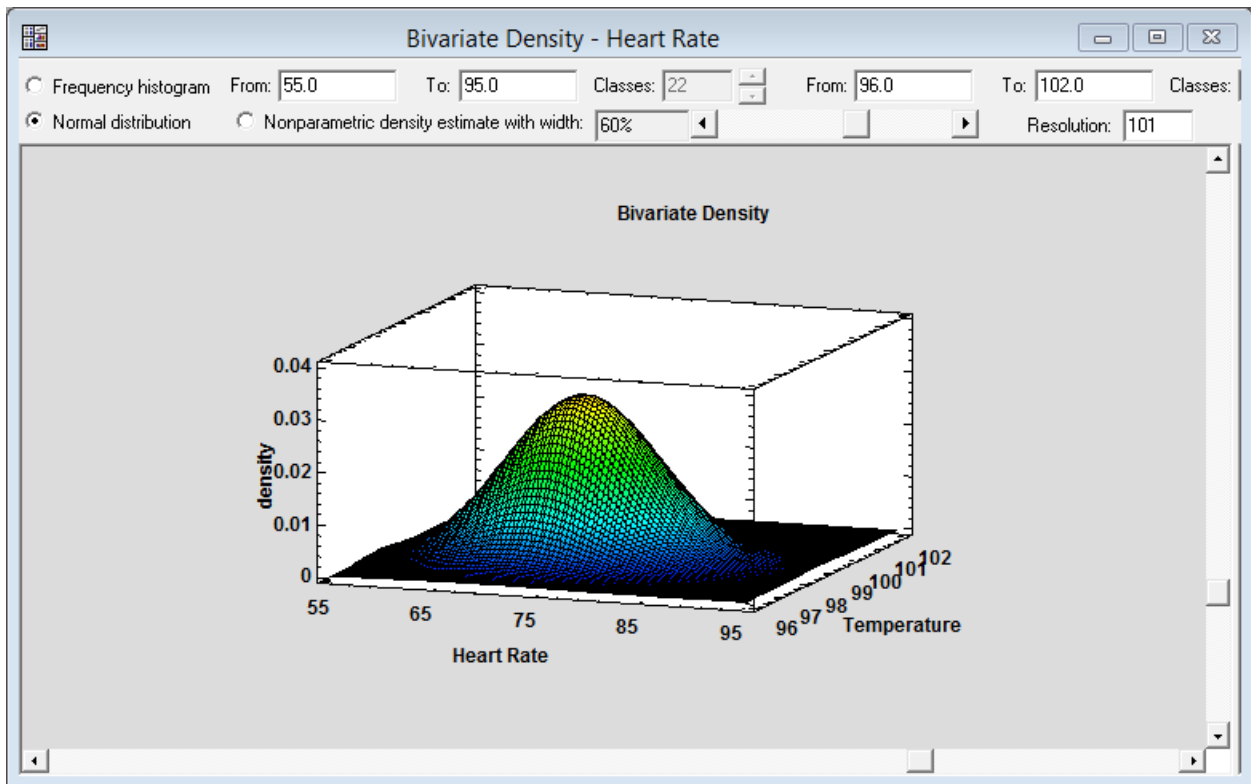
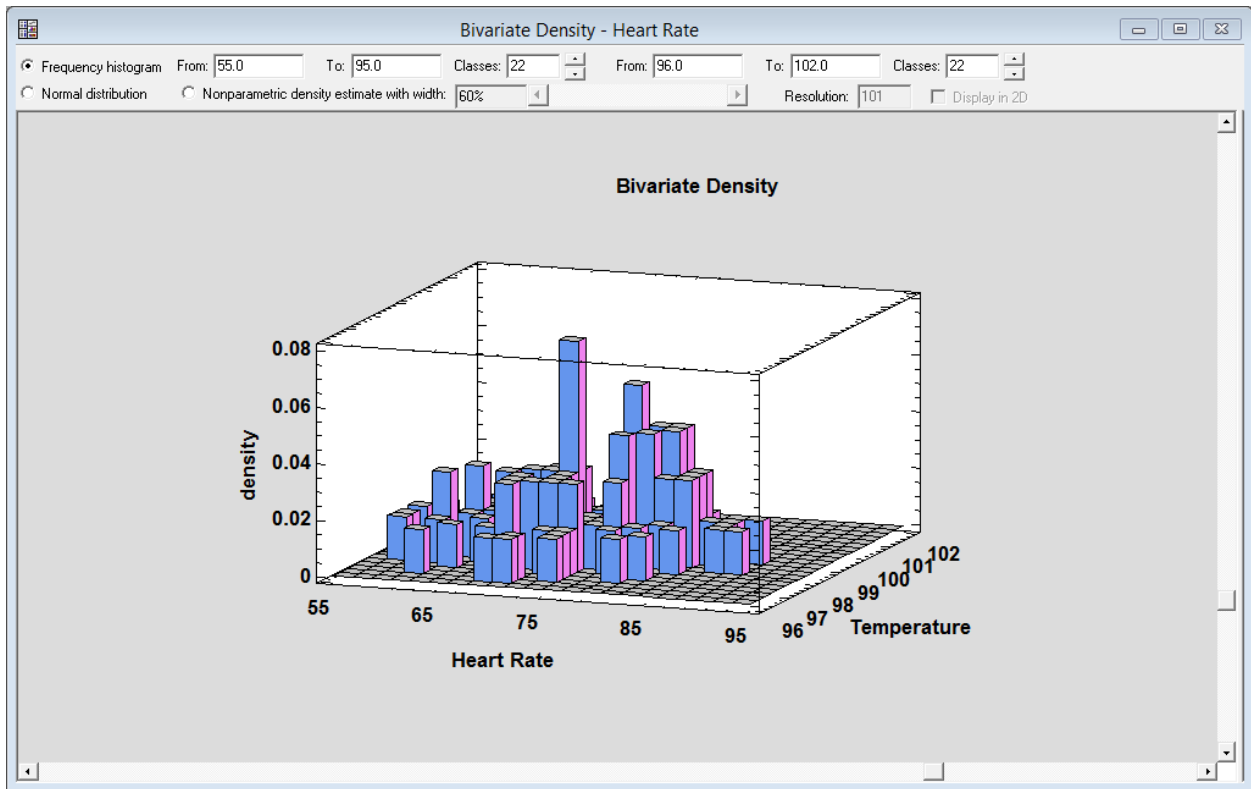
A new Statlet has been added to display the distribution of 2 random variables:

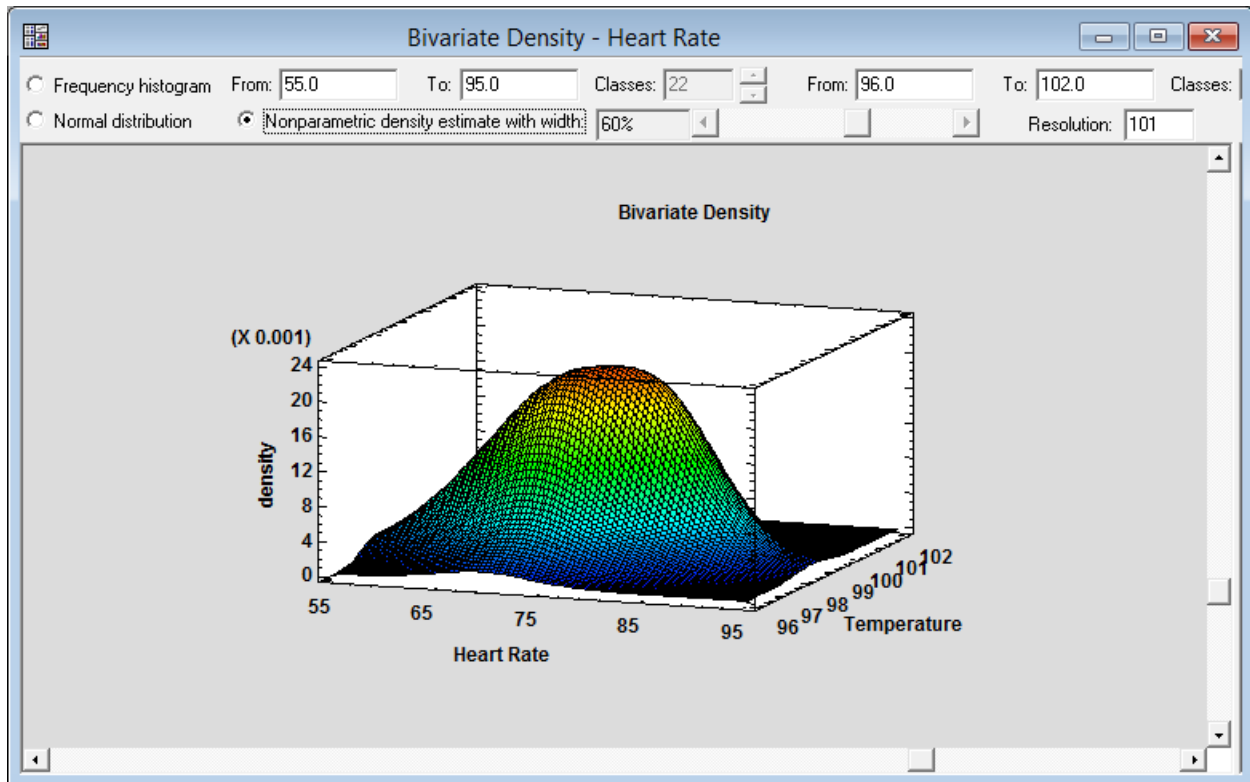


The Statlet creates 3 types of graphs:

1. A bivariate frequency histogram, which displays the observed counts in a rectangular grid array.
2. A bivariate normal distribution with the same means, standard deviations and correlation as the data.
3. A nonparametric density estimate which does not assume a normal distribution.

Users may interactively change the parameters that control the graphs.

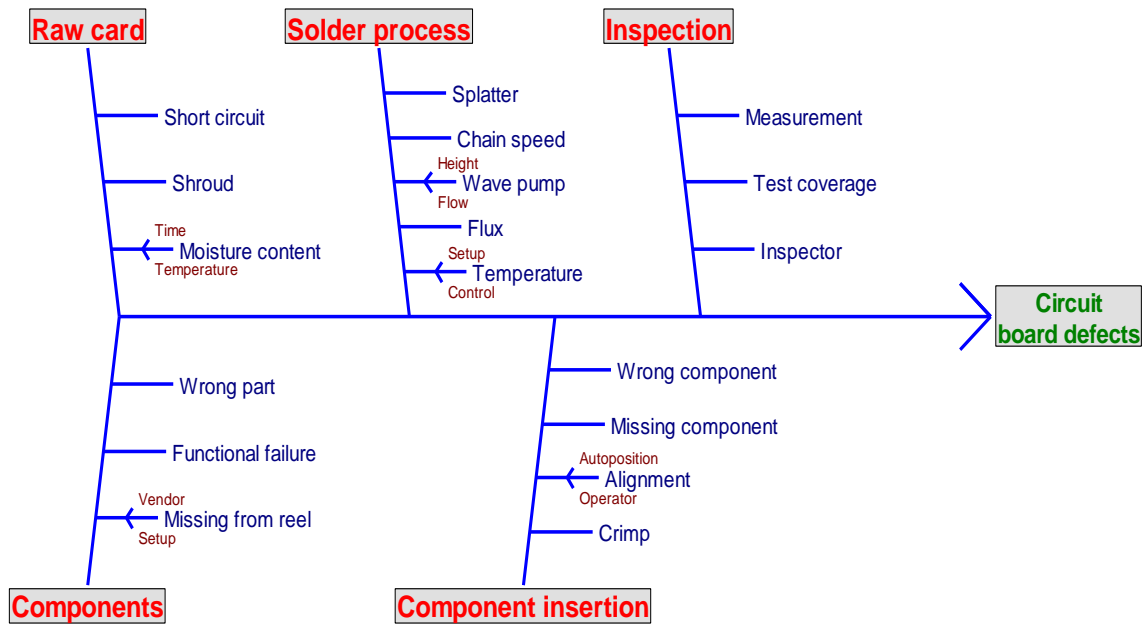




The cell definitions and frequencies for the bivariate histogram may also be saved to a datasheet in the format needed for input to the *Crosstabulation* procedure:

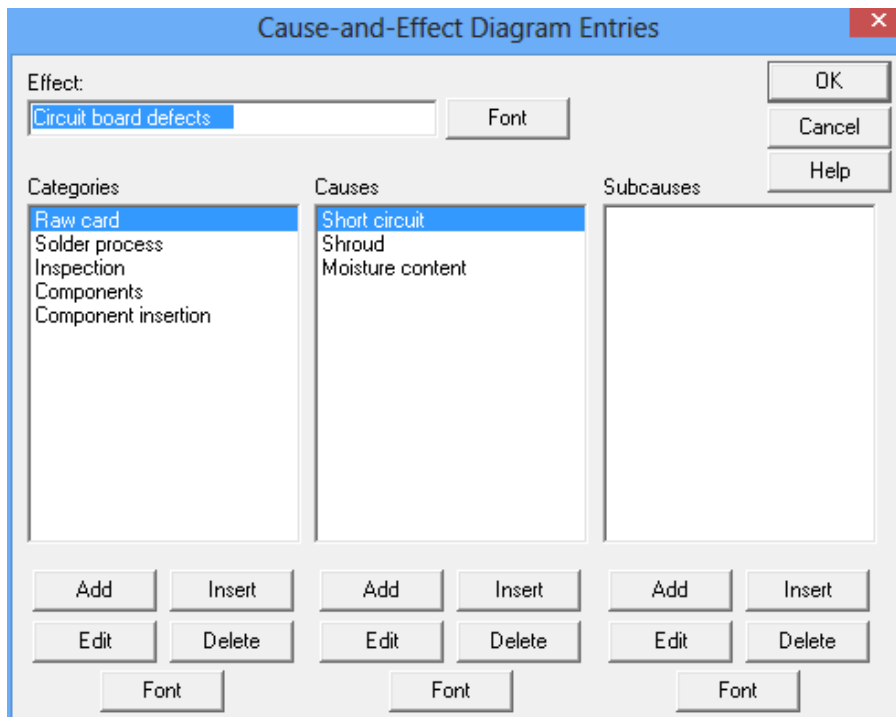
Cause-and-Effect Diagram

New features have been added to the Cause-and-Effect Diagram to improve the display:



These features include:

1. Better control over the colors and fonts for the labels at each level.
2. Boxes around the effect and primary categories.



Computer-Generated Designs

Computer Generated designs, which are a new type of experiment created by the DOE Wizard, are experimental designs which have optimal properties with respect to the estimation of specific statistical models. Given the definition of an experimental region, a model to be estimated, and the number of experimental runs that can be performed, the program searches for a set of runs that maximizes a selected design optimality criteria.

Computer generated designs are included as an option in the Experimental Design Wizard when constructing several types of designs, including screening designs, response surface designs, and mixture experiments. In addition to lower and upper limits for each experimental factor, constraints may be specified based on linear combinations of the factors. The optimality criteria offered include D-optimality, A-optimality, G-optimality, and I-optimality.

BLOCK	temperature degrees C	pressure bar	speed cpm	supplier
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				

Optimize

D-efficiency

G-efficiency

A-efficiency

I-efficiency

Number of coefficients: 18

Display

Original units

Coded units

Number of base runs: 24

Number of replicates: 0

Number of centerpoints: 0

Group runs in blocks of size: 2

Randomize run order

Average prediction variance:

OK

Cancel

Create

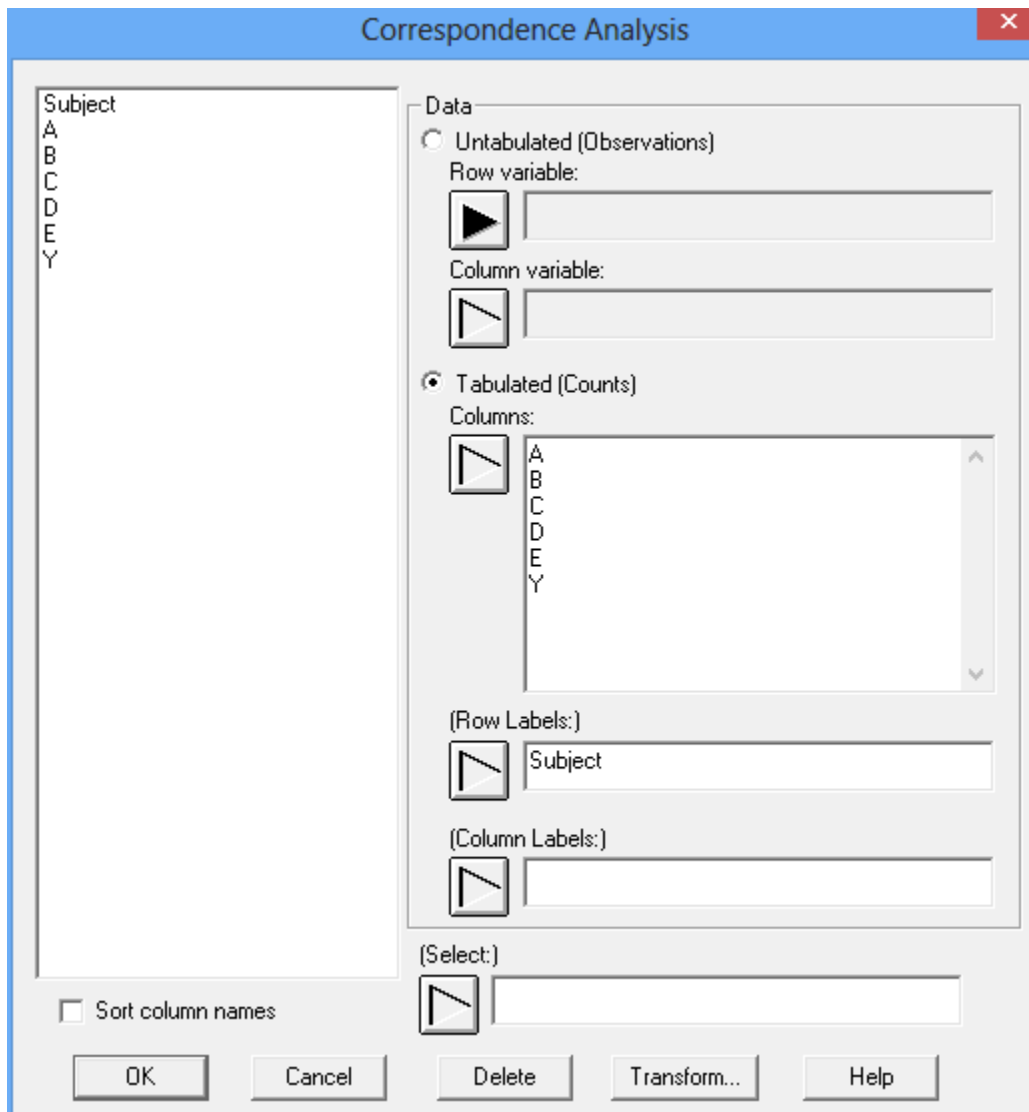
Advanced

Help

Unlike the D-optimal designs in Version 16, no set of candidate runs is required before generating the design.

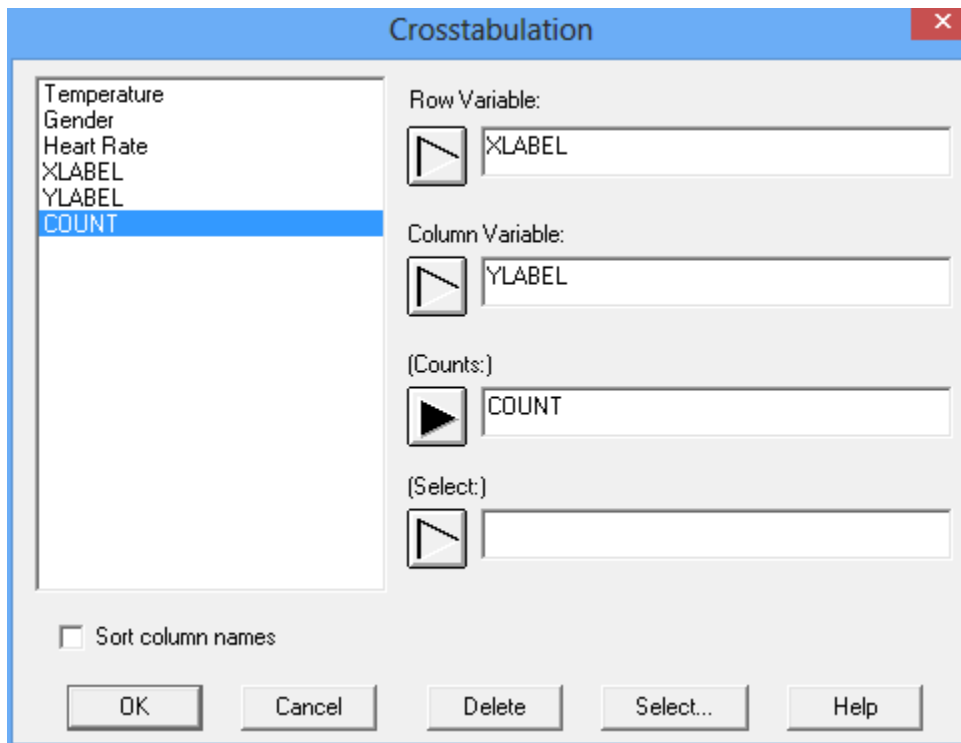
Correspondence Analysis

A new field has been added to the data input dialog box which allows *Column Labels* to be input from a column of the DataBook:



Crosstabulation

The data input dialog box for the *Crosstabulation* procedure has been modified to allow for specification of a *Counts* column:



If used, the *Counts* column specifies the frequency associated with each value in the *Row Variable* and *Column Variable* columns. If not used, each frequency is assumed to equal 1 as in Version 16.

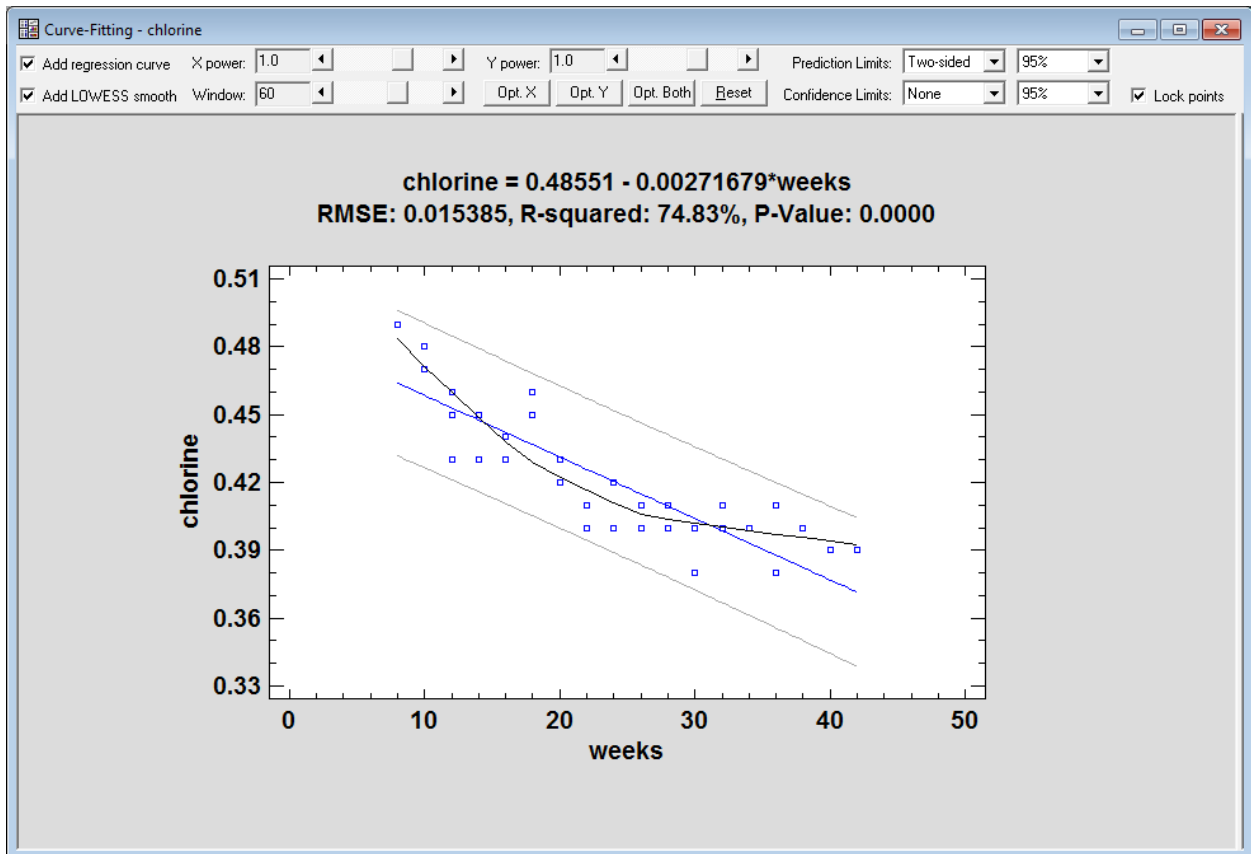
Curve-Fitting Statlet

This new procedure fits linear and nonlinear regression models involving a dependent variable Y and an independent variable X. The general model that is fit contains each variable raised to a power:

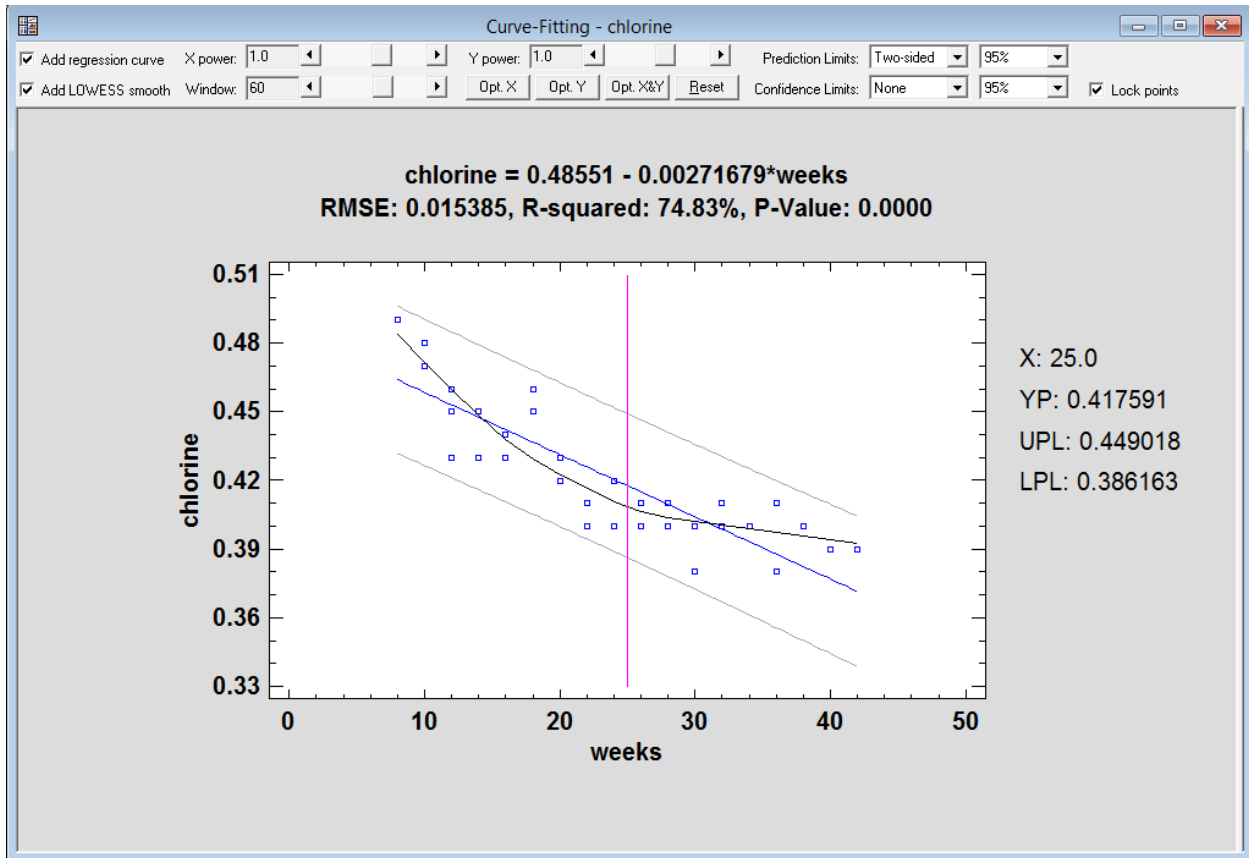
$$Y^p = a + b X^q$$

Using the Statlet, you may:

- (1) Visualize the effect of the powers by using the sliders on the toolbar to change them interactively.
- (2) Add a LOWESS smooth to the plot to compare it to the fitted model.
- (3) Interactively change the width of the smoothing window.
- (4) Use the optimize buttons to numerically optimize the powers.
- (5) Add one-sided or two-sided prediction and confidence limits.
- (6) Interactively investigate the effect of individual data points on the fitted model and the limits by dragging the points to new positions.

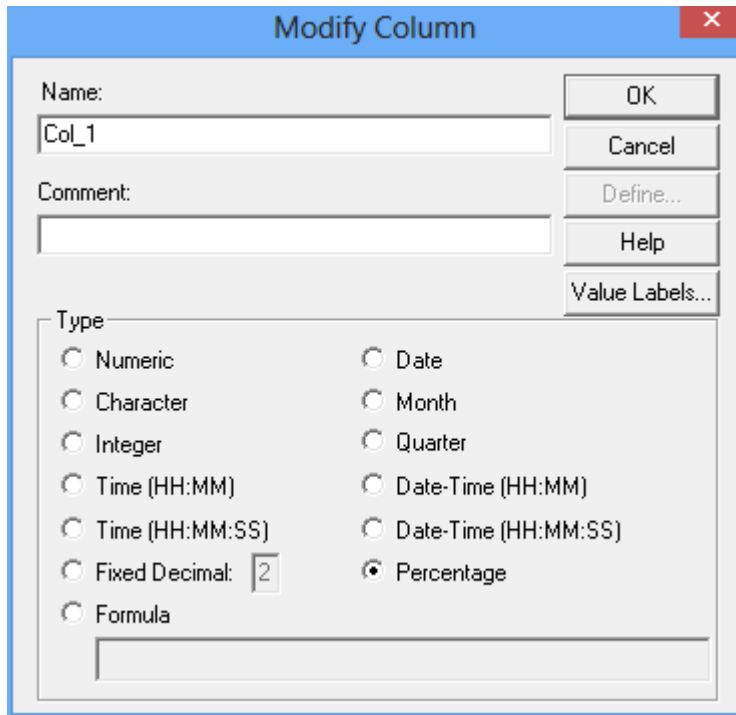


Using a vertical cursor, you can also make predictions at specific values of X:



Data Columns

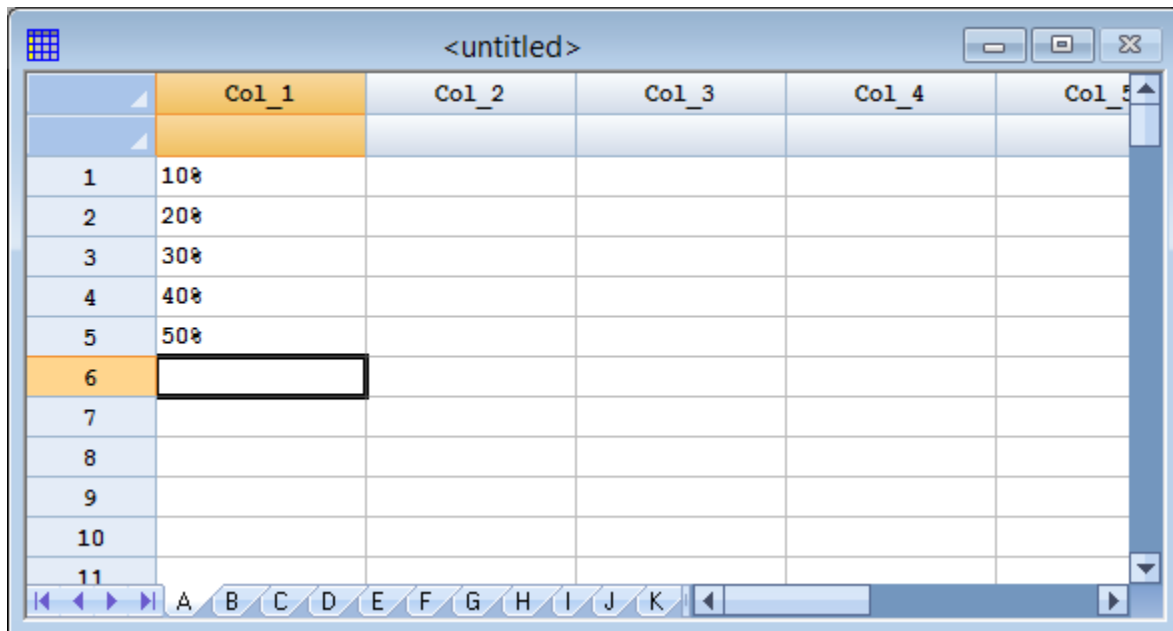
Two significant enhancements have been made to the columns of the DataBook in which data are placed. The first enhancement is the addition of a new *Percentage* data type:



The 'Modify Column' dialog box shows the following configuration:

- Name: Col_1
- Type: Percentage (selected)
- Fixed Decimal: 2

Data in percentage columns appear with a percentage sign:

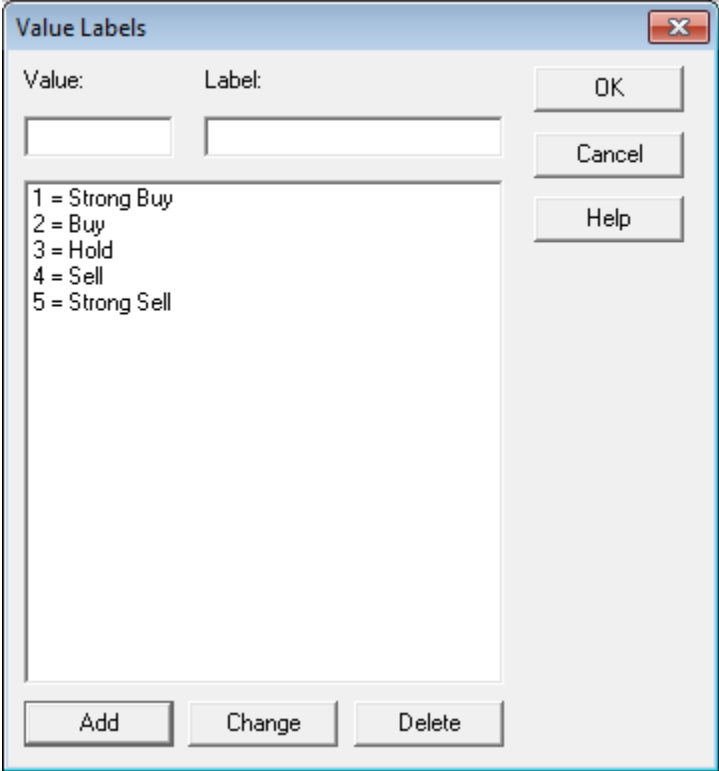


	Col_1	Col_2	Col_3	Col_4	Col_5
1	10%				
2	20%				
3	30%				
4	40%				
5	50%				
6					
7					
8					
9					
10					
11					

Internally, they are represented by their numeric values, such as 0.1 for 10%.

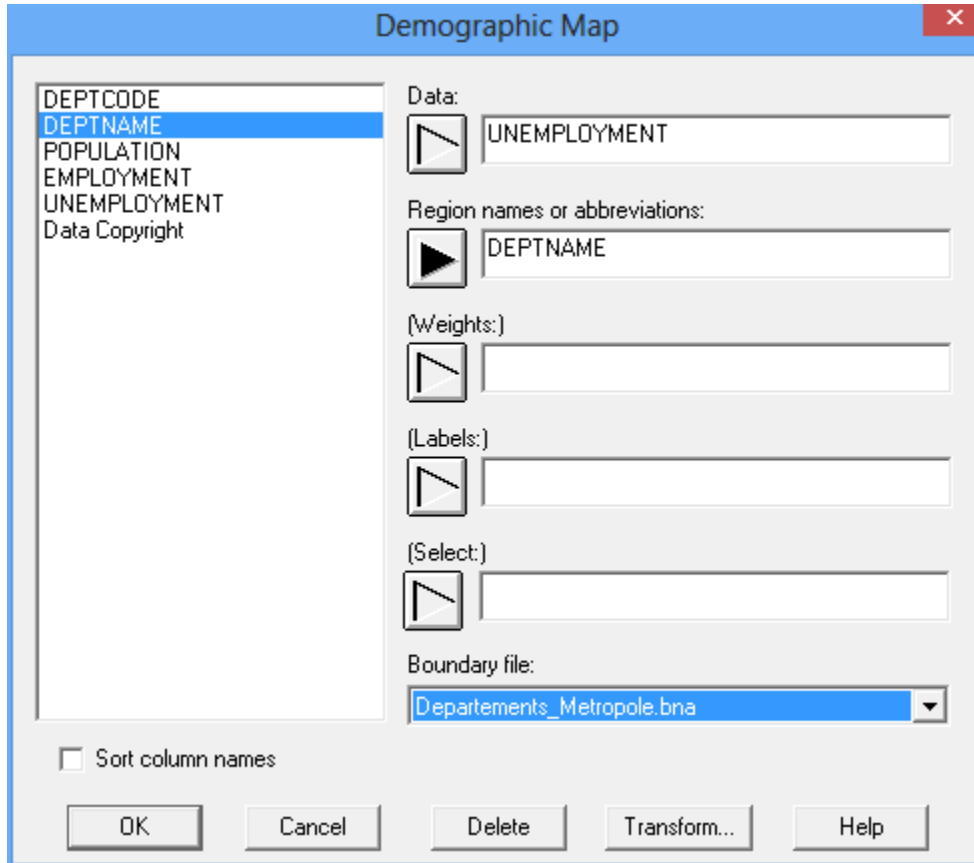
The second enhancement is the ability to specify value labels to be associated with numeric columns. Value labels are character strings that will replace specific values in the column when output for that column is displayed. For example, stock analysts sometimes classify individual securities as "Strong Buy", "Buy", "Hold", "Sell", or "Strong Sell". Rather than entering those strings in a character column, it is often easier to enter numbers such as 1, 2, 3, 4 or 5 in a numeric column. By creating value labels that correspond to each of those numbers, you can have the strings appear in all output in place of the corresponding numbers.

Value labels are created by pressing the *Value Labels* button on the *Modify Column* dialog box:



Demographic Map

The *Map by State* procedure has been renamed *Demographic Map* and moved to the *Plot* menu. It can now create maps for any regions defined by a BNA boundary file. Included with Statgraphics Centurion are sample boundary files, including one that defines department boundaries in France:



The *Analysis Options* dialog box is used to specify the type of data to be plotted and how to color the regions:

Demographic Map Options [X]

Data type:

Categorical

Discrete

Continuous

OK

Cancel

Help

Classes:

From:

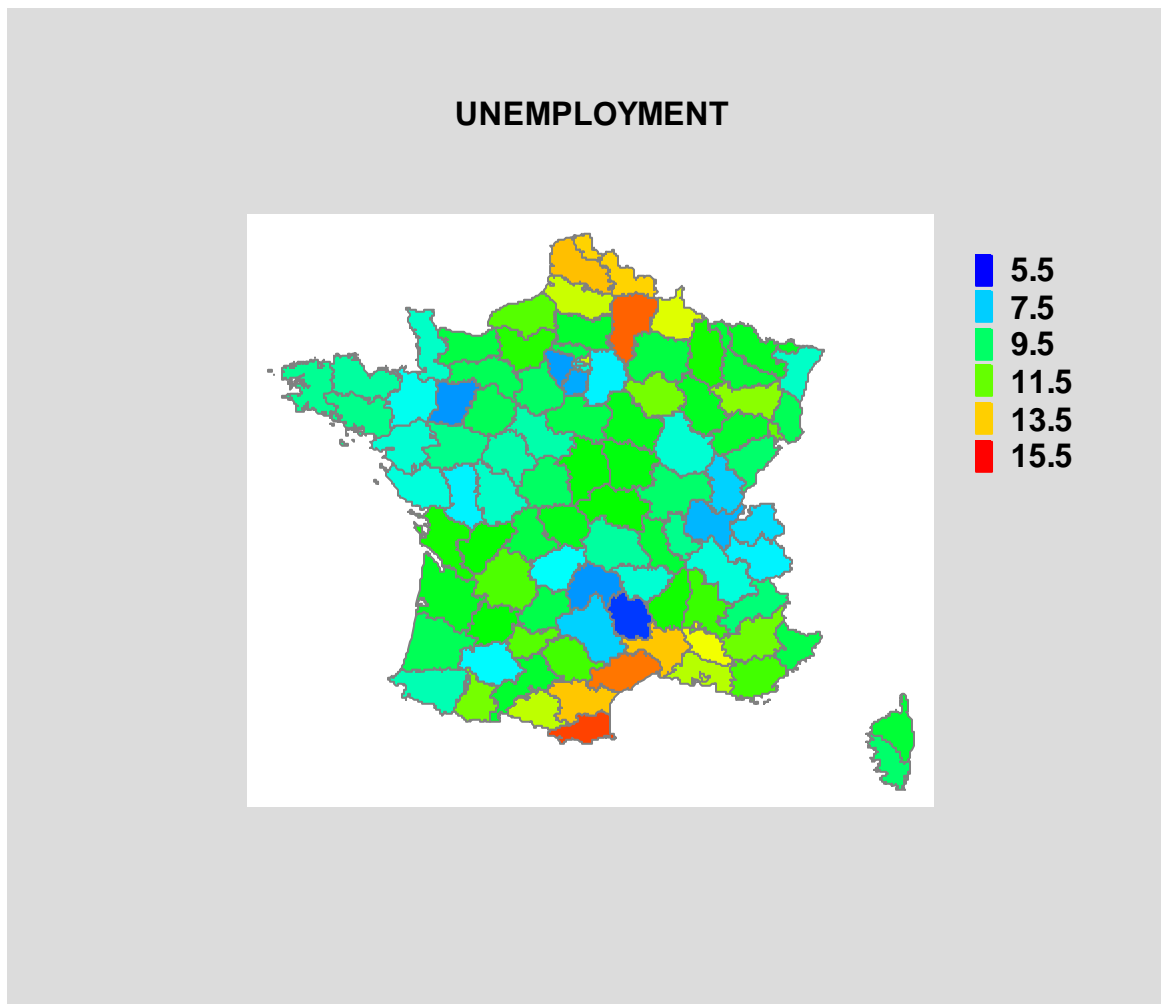
To:

By:

Use color gradient

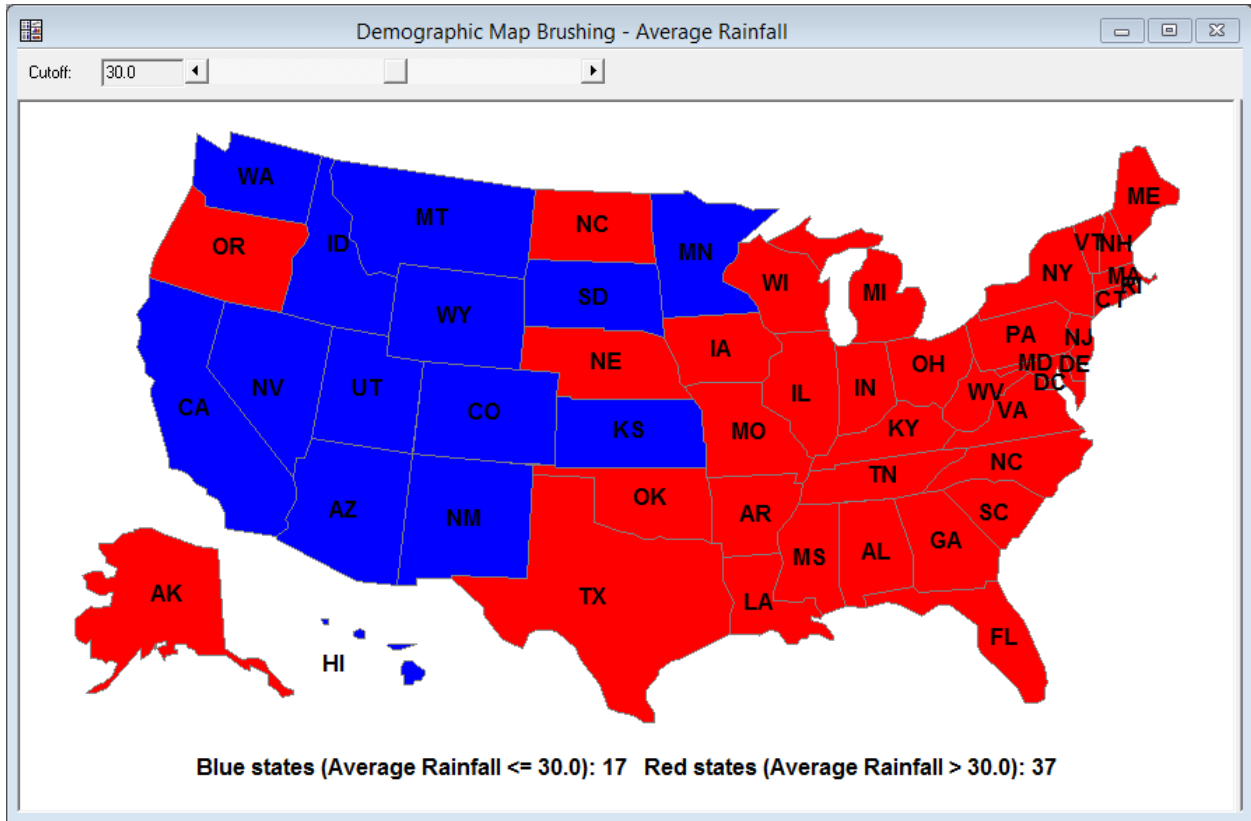
Hold scaling constant

Continuous data may be plotted using a color gradient:

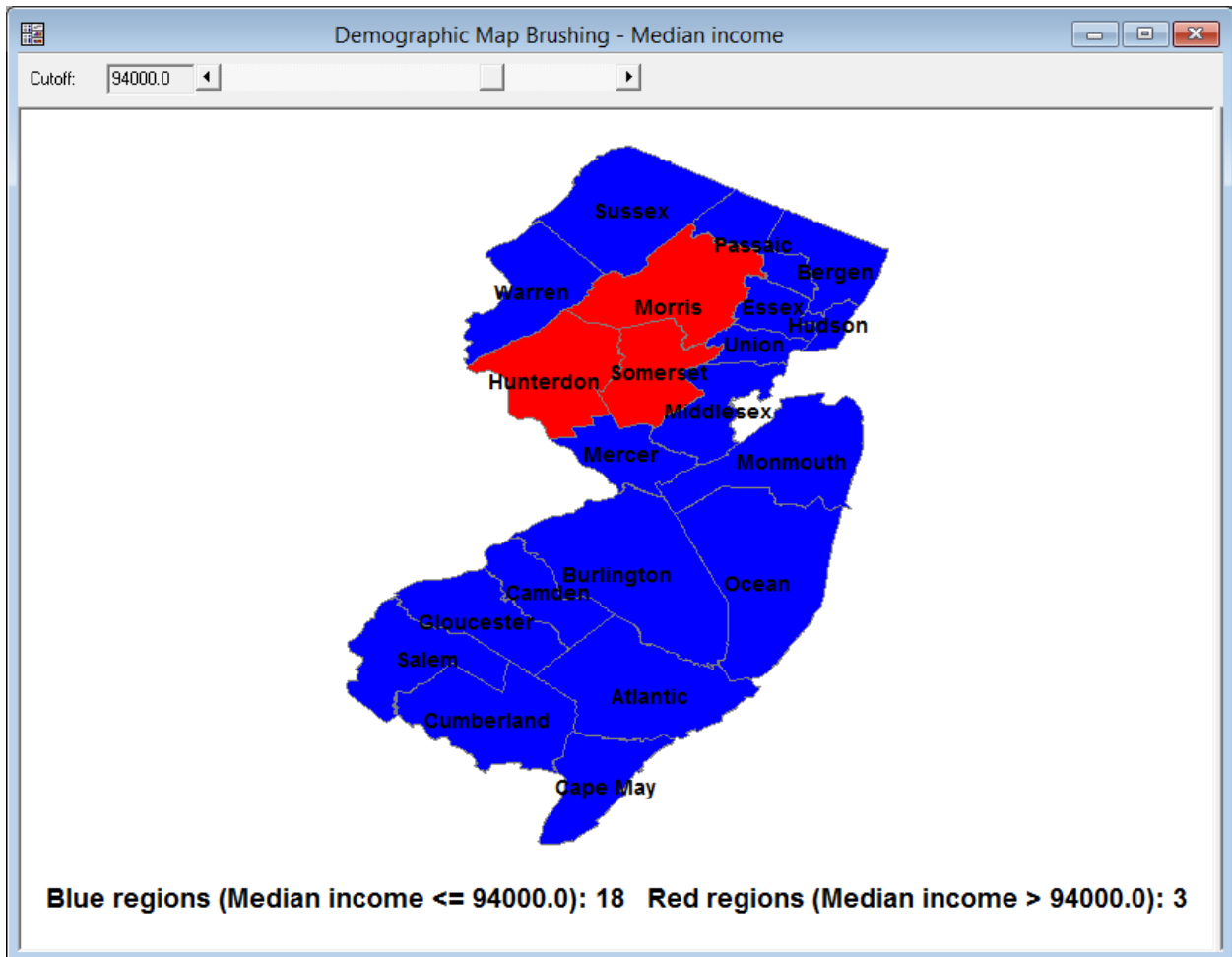


Demographic Map Brushing Statlet

This new Statlet plots a demographic map in which each region is colored either blue or red to illustrate the value of a selected variable. Using the Statlet controls, the analyst may change the cutoff that divides red from blue. Interactively changing the cutoff helps in visualizing the distribution of the specified variable throughout the map.

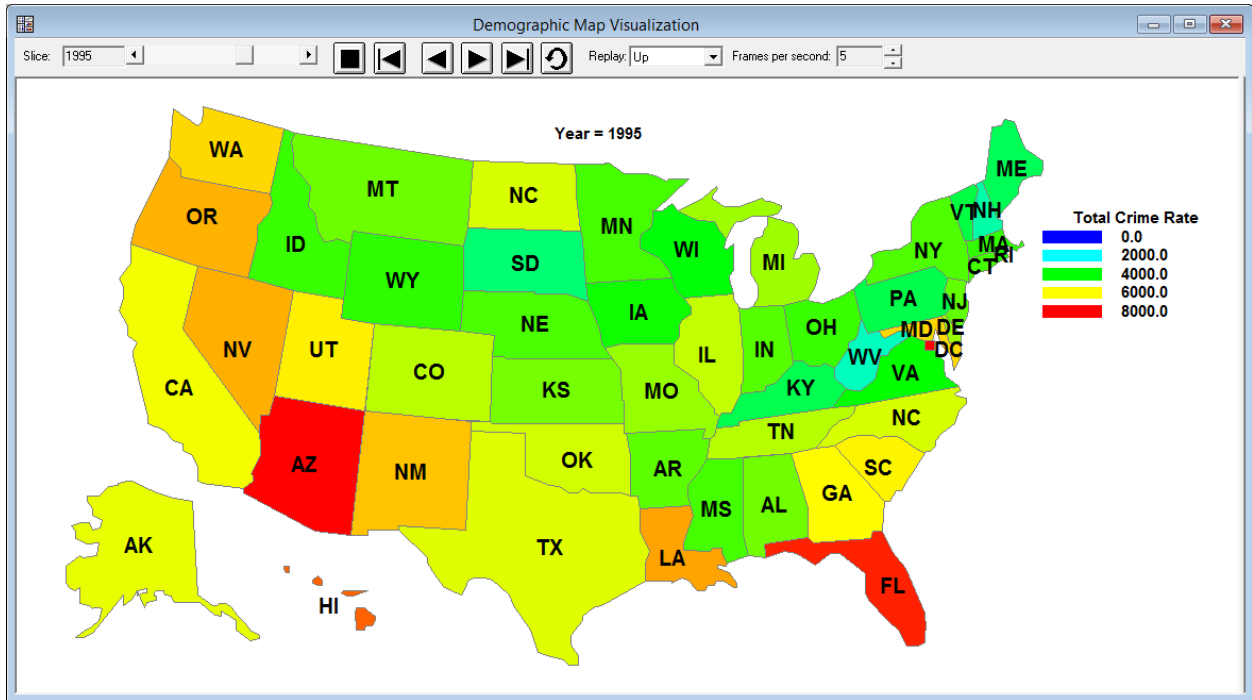


The procedure can map any regions defined by an Atlas Boundary File (file extension BNA). Such files are available for many states and countries:



Demographic Map Visualizer Statlet

This new Statlet is designed to illustrate changes in regional statistics over time. Given data for each region during p time periods, the program generates a dynamic display that illustrates how the data have changed in each region. Each region is drawn using a color that illustrates the level of the selected statistic.



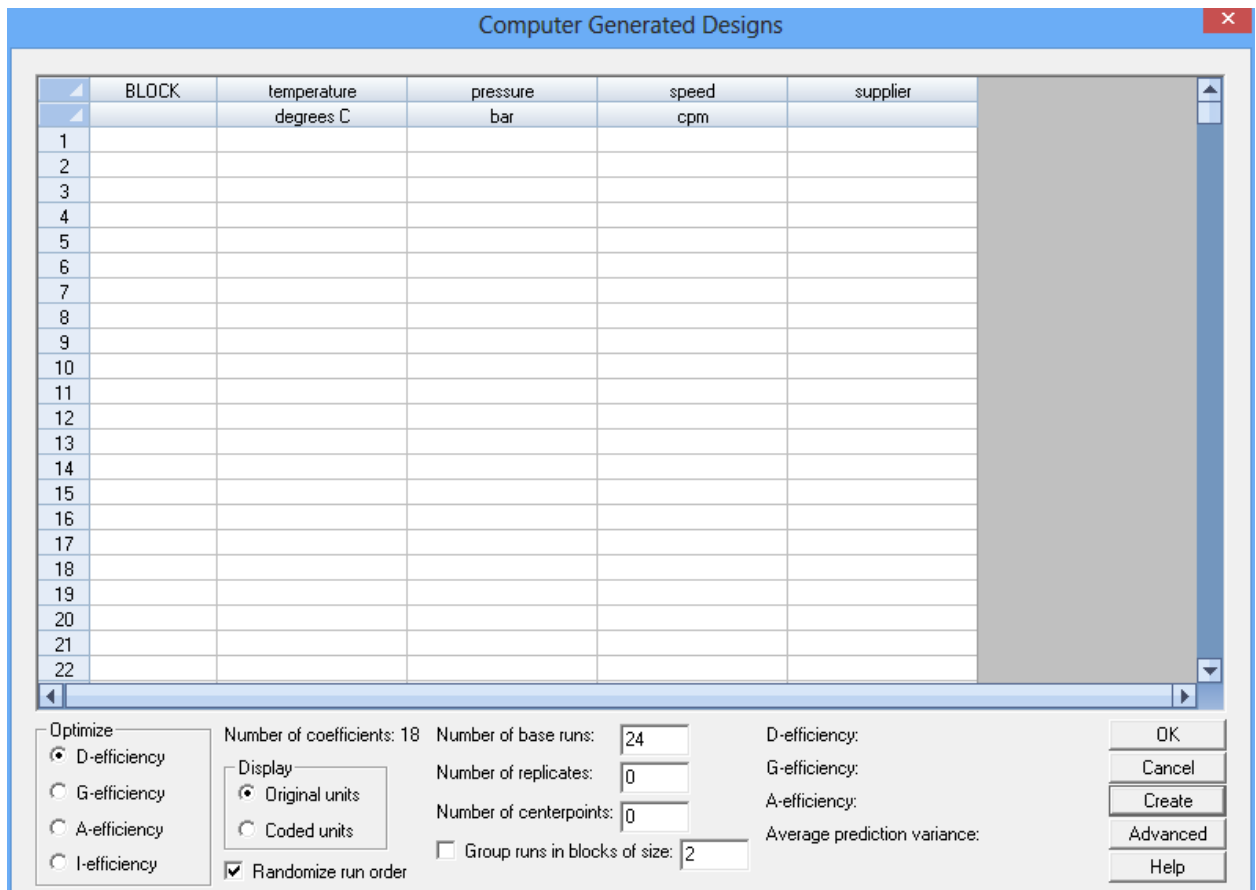
As time increases, the analyst can follow changes in each region. Various options are offered for smoothing the data and for dealing with missing values.

The procedure can map any regions defined by an Atlas Boundary File (file extension BNA).

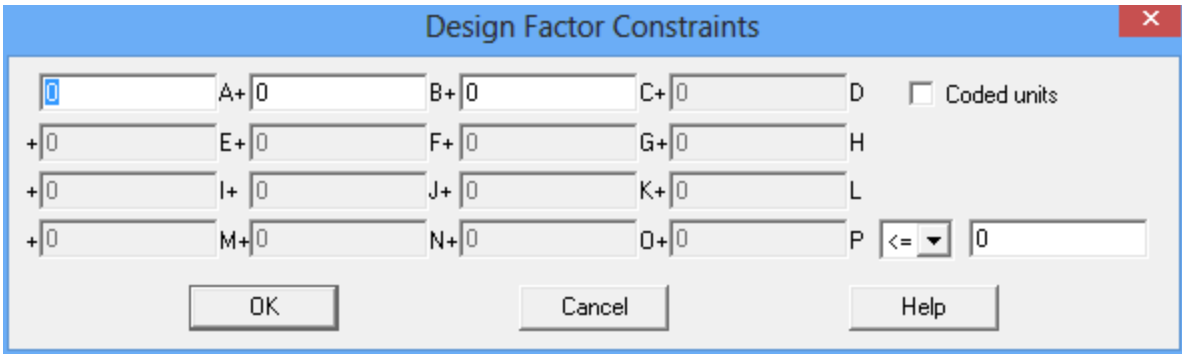
Design of Experiments Wizard

Several significant enhancements have been made to the Design of Experiments Wizard:

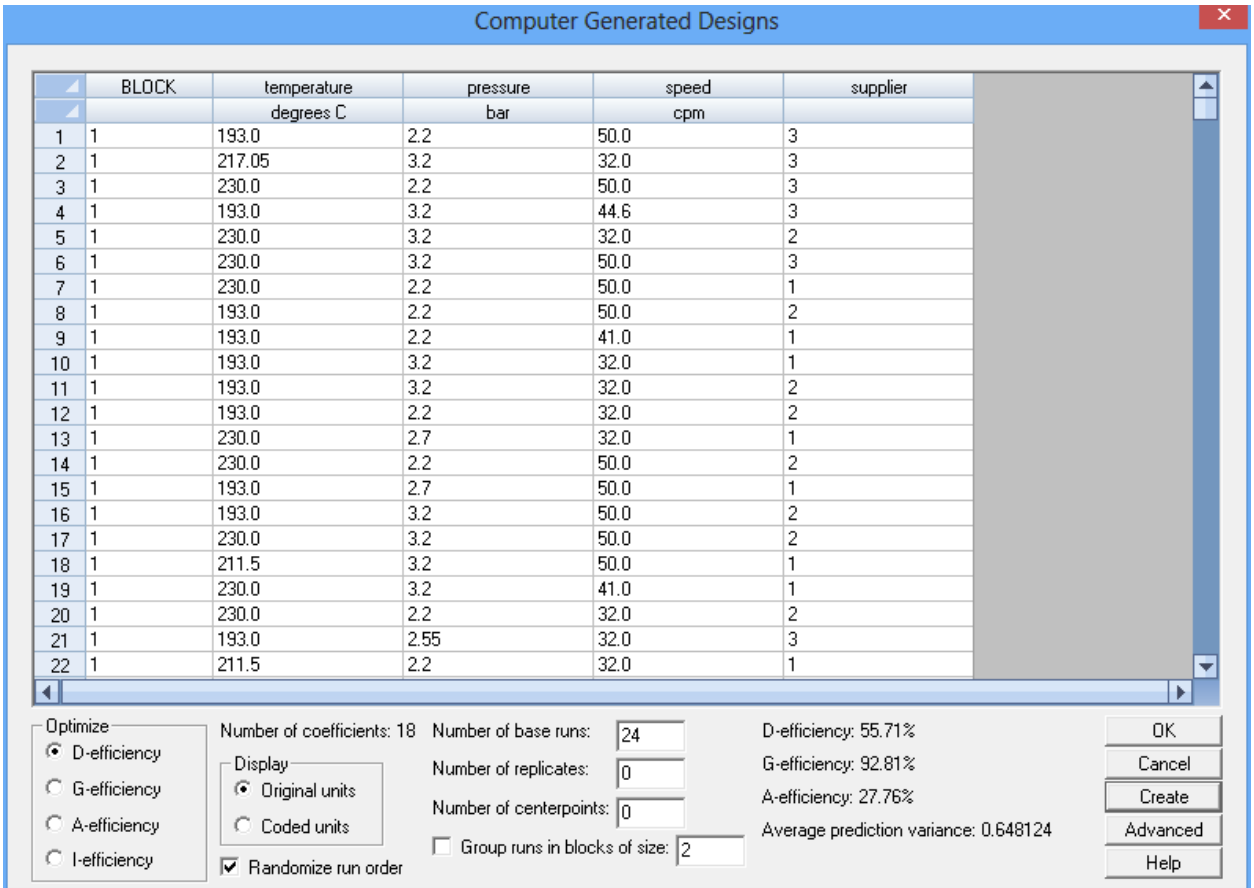
1. The maximum number of response variables has been increased from 16 to 32.
2. 3-level fractional factorial designs have been added to the list of available Response Surface Designs for experiments with 3 to 8 experimental factors.
3. A new class of designs called *Computer Generated Designs* has been added. These designs allow you to create experimental designs which have optimal properties with respect to the estimation of specific statistical models. Given the definition of an experimental region, a model to be estimated, and the number of experimental runs that can be performed, the program searches for a set of runs that maximizes a selected design optimality criteria. Unlike in Version 16, you do not have to specify a set of candidate runs.



4. For Computer Generated designs, the experimental region may be constrained based on linear combinations of the factors.

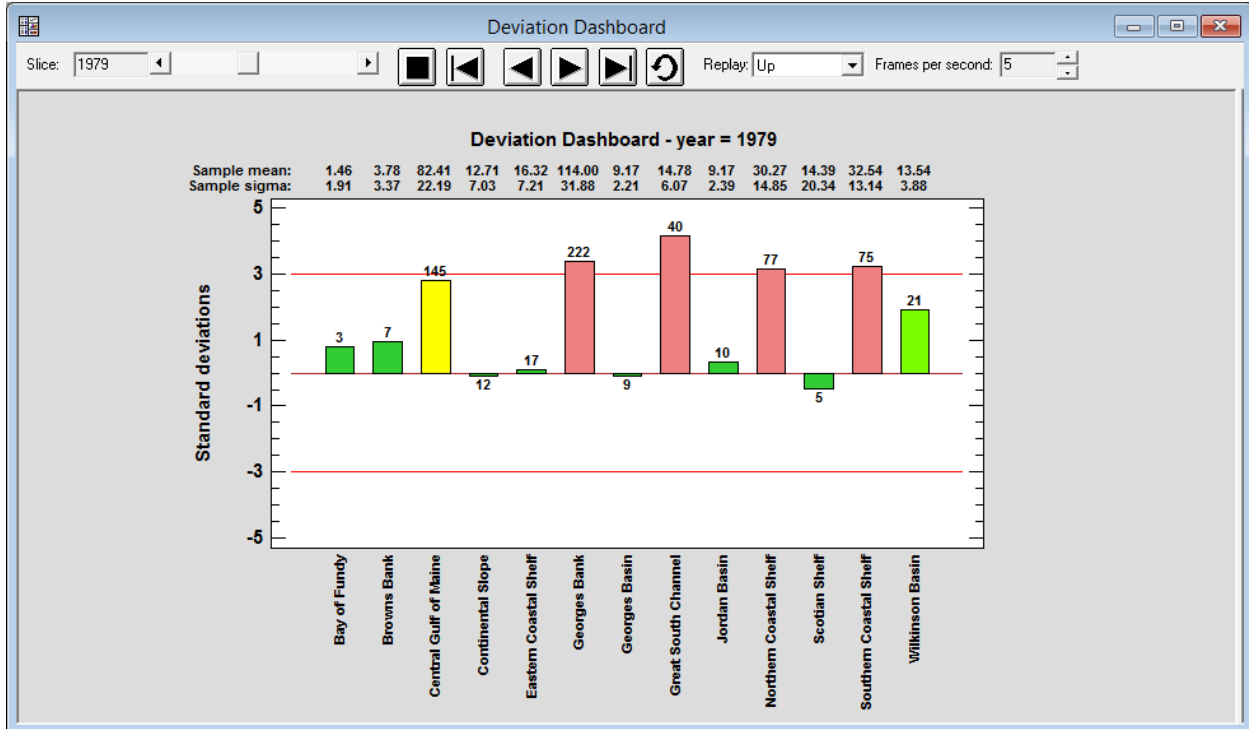


5. The optimality criteria offered now include D-optimality, A-optimality, G-optimality, and I-optimality.



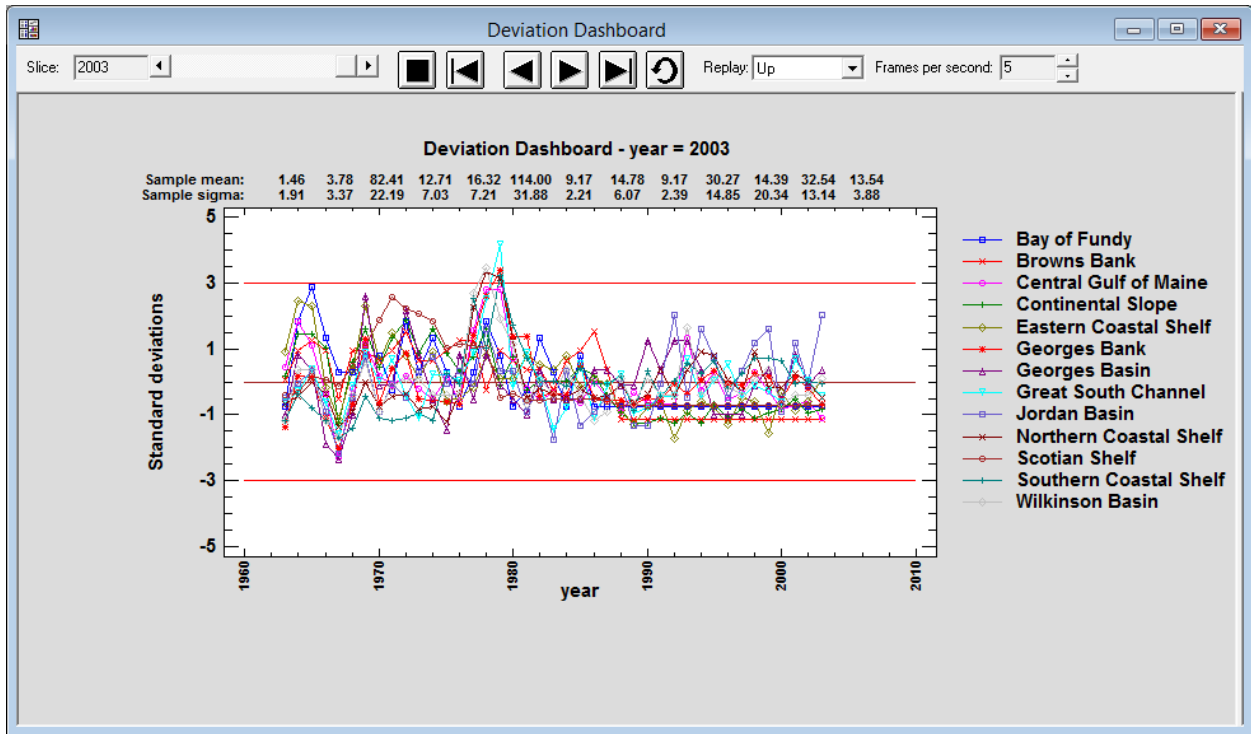
Deviation Dashboard Statlet

This new Statlet indicates the status of multiple variables by displaying how much they deviate from their respective means. The primary format shows each variable using a vertical bar, the color of which indicates how many standard deviations the variable is from its mean:



The toolbar controls may be used to dynamically change the time period indicated on the plot.

A secondary format show all variables displayed as a normalized control chart:

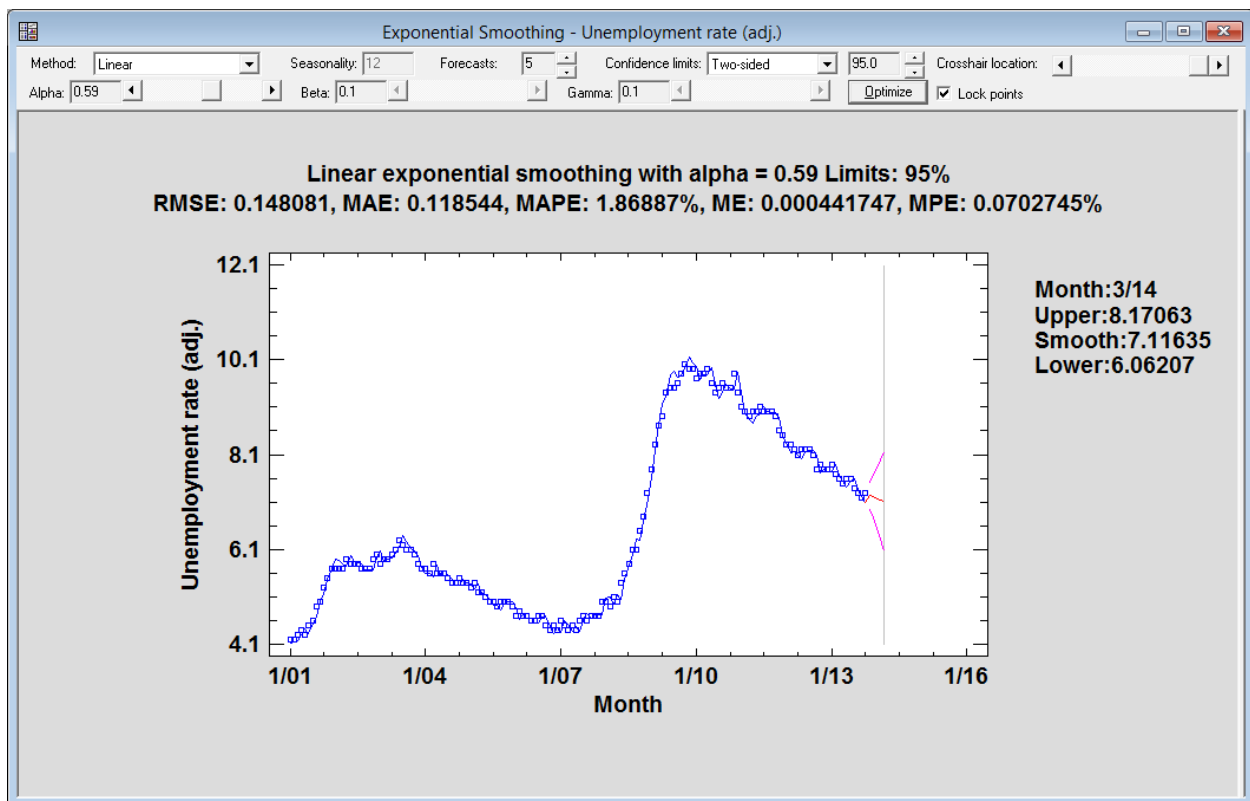


Exponential Smoothing Statlet

This new Statlet applies various types of exponential smoothers to a time series. It generates forecasts with associated forecast limits. Using the Statlet controls, the user may interactively change the values of the smoothing parameters to examine their effect on the forecasts.

The types of exponential smoothers included are:

1. Brown's simple exponential smoothing with smoothing parameter α .
2. Brown's linear exponential smoothing with smoothing parameter α .
3. Holt's linear exponential smoothing with smoothing parameters α and β .
4. Brown's quadratic exponential smoothing with smoothing parameter α .
5. Winters' seasonal exponential smoothing with smoothing parameters α , β , and γ .



Interesting features of this Statlet include:

1. Sliders that allow the user to interactively change the smoothing parameters. As the parameters change, so do the forecasts, forecast limits, and summary statistics.
2. An *Optimize* button that instructs the program to find optimal values of the smoothing parameters.
3. A slider that controls the location of the crosshair cursor. The forecast and forecast limits corresponding to the cursor location are displayed in the right margin of the graph.
4. Unless *Lock points* is checked, users may drag historical data to new positions along the vertical axis to observe the effect of individual data values on the forecasts.

Factor Analysis

New tests have been added to the *Factor Analysis* procedure to determine whether it is useful to perform a factor analysis on a set of variables. The factorability tests include the Kaiser-Meyer-Olsen (KMO) measure of sampling adequacy and Bartlett's test of sphericity:

Factorability Tests

Kaiser-Meyer-Olkin Measure of Sampling Adequacy
KMO = 0.920192

Bartlett's Test of Sphericity
Chi-Square = 1299.83
D.F. = 55
P-Value = 0.0

The StatAdvisor

The factorability tests provide indications of whether or not it is likely to be worthwhile attempting to extract factors from a set of variables. The KMO statistic provides an indication of how much common variance is present. For factorization to be worthwhile, KMO should normally be at least 0.6. Since KMO = 0.920192, factorization is likely to provide interesting information about any underlying factors.

Bartlett's test for sphericity tests the hypothesis that the correlation matrix amongst the variables is an identity matrix, indicating that they share no common variance. Since the P-value is < 0.05 , that hypothesis is rejected. Note: Bartlett's test is very sensitive and is usually ignored unless the number of samples per variable is no more than 5. In this case, the number of samples per variable equals 8.45455.

Factorial Repeated Measures ANOVA

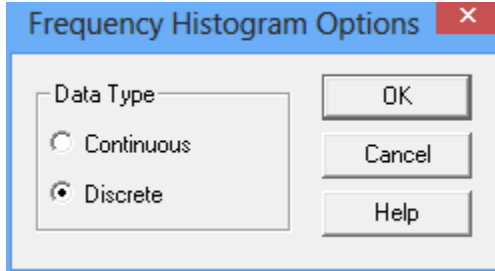
This new procedure analyzes experimental designs in which multiple measurements have been made on the same subjects, and in which one or more factors differentiate the subjects. In many cases, the measurements are taken sequentially in order to examine how the effect of the between-subject factors varies over time. The need for a special procedure stems from potential dependence between observations made on a given subject.

Included in the procedure is Mauchly's test for *sphericity*, which is a condition requiring that the variance between any two estimated treatment means be the same. Corrections to the standard F-test are included for cases when sphericity is not present. Alternative MANOVA tests are also included.

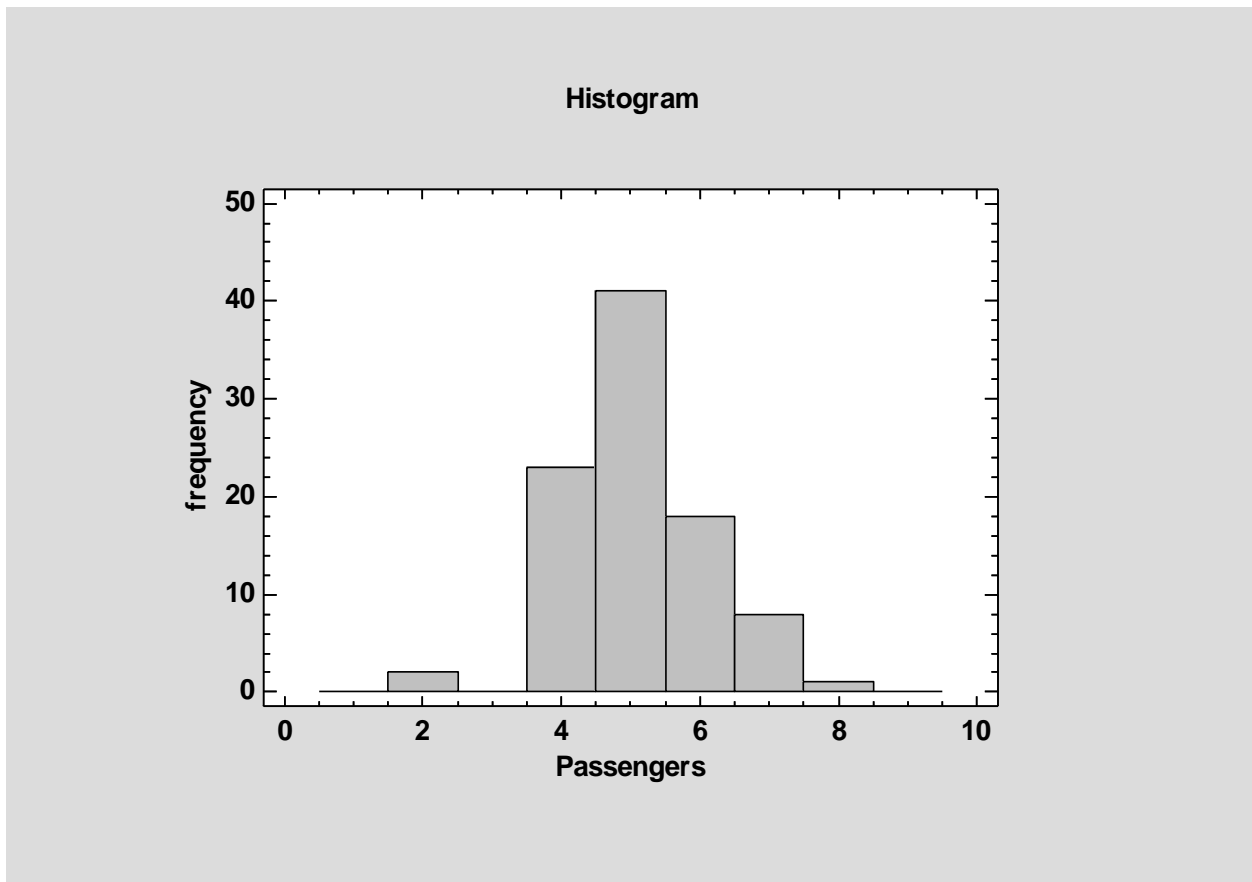
Sphericity Tests and Adjustments						
Mauchly's Sphericity Test						
<i>W</i>	<i>Chi-square</i>	<i>D.f.</i>	<i>P-value</i>			
0.631214	9.07439	5.0	0.106134			
Epsilon						
<i>Huynh-Feldt</i>	<i>Greenhouse-Geisser</i>	<i>Lower-bound</i>				
0.963045	0.776916	0.333333				
Tests of Within-Subject Effects						
<i>Source</i>	<i>Sphericity Correction</i>	<i>Sum of Squares</i>	<i>Df</i>	<i>Mean Square</i>	<i>F-Ratio</i>	<i>P-Value</i>
Time	None	289.615	3.0	96.5382	12.96	0.0000
	Huynh-Feldt	289.615	2.88913	100.243	12.96	0.0000
	Greenhouse-Geisser	289.615	2.33075	124.258	12.96	0.0000
	Lower-bound	289.615	1.0	289.615	12.96	0.0017
Time*Drug	None	527.417	6.0	87.9028	11.80	0.0000
	Huynh-Feldt	527.417	5.77827	91.2759	11.80	0.0000
	Greenhouse-Geisser	527.417	4.6615	113.143	11.80	0.0000
	Lower-bound	527.417	2.0	263.708	11.80	0.0004
Residual	None	469.219	63.0	7.44792		
	Huynh-Feldt	469.219	60.6718	7.73372		
	Greenhouse-Geisser	469.219	48.9457	9.58651		
	Lower-bound	469.219	21.0	22.3437		
Multivariate Tests						
<i>Effect</i>	<i>Test</i>	<i>Value</i>	<i>F</i>	<i>Hypothesis D.F.</i>	<i>Error D.F.</i>	<i>P-Value</i>
Time	Wilks' lambda	0.293279	15.2616	3.0	19.0	0.0000
	Pillai trace	0.706721	15.2616	3.0	19.0	0.0000
	Hotelling-Lawley trace	2.40973	15.2616	3.0	19.0	0.0000
	Roy's largest root	2.40973	15.2616	3.0	19.0	0.0000

Frequency Histogram

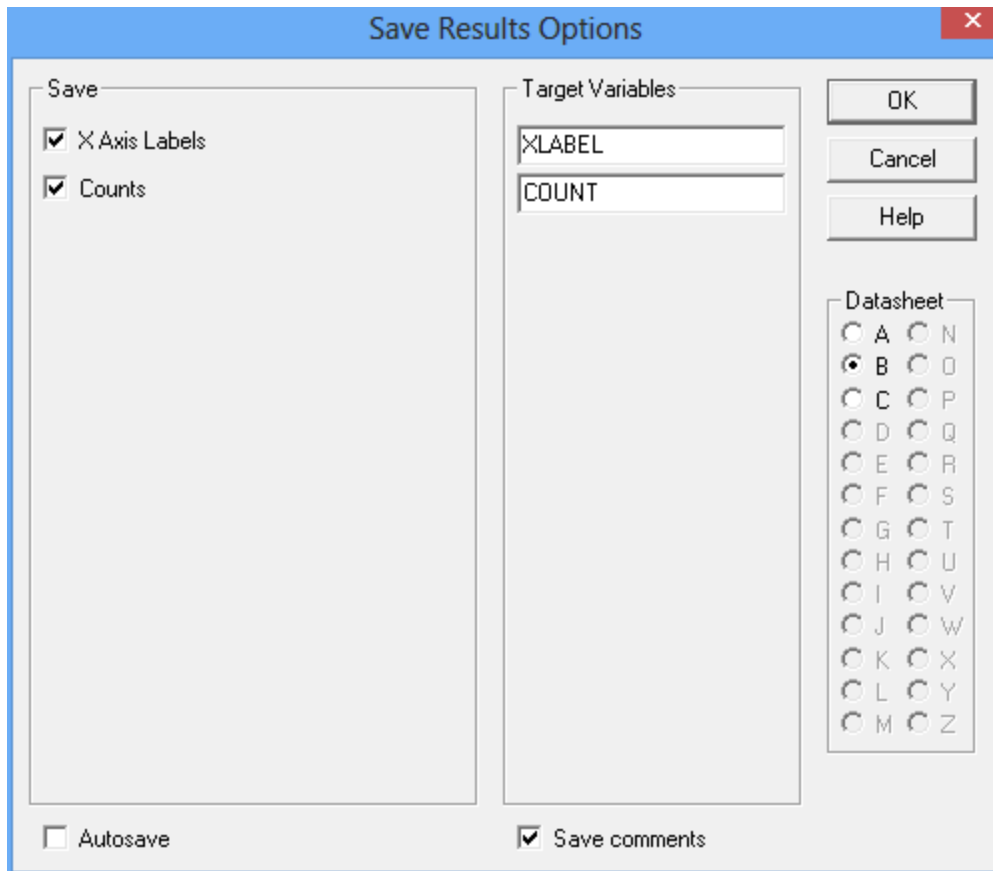
Two significant enhancements have been made to the *Frequency Histogram* procedure. The first enhancement is the addition of an *Analysis Options* dialog box on which the user specifies whether the data are continuous or discrete:



If the data are discrete, the histogram will automatically create a bar for each integer value within the range of the data:



The second enhancement is the addition of a *Save Results* dialog box which allows the user to save the calculated frequencies to a datasheet:



The resulting columns are in the format needed for input to the *Tabulation* procedure:

	XLABEL	COUNT	Col_3	Col_4	Col_5	Col_6
	X Axis	Counts				
1	0	0				
2	1	0				
3	2	2				
4	3	0				
5	4	23				
6	5	41				
7	6	18				
8	7	8				
9	8	1				
10	9	0				
11	10	0				
12						
13						
14						
15						

G Chart

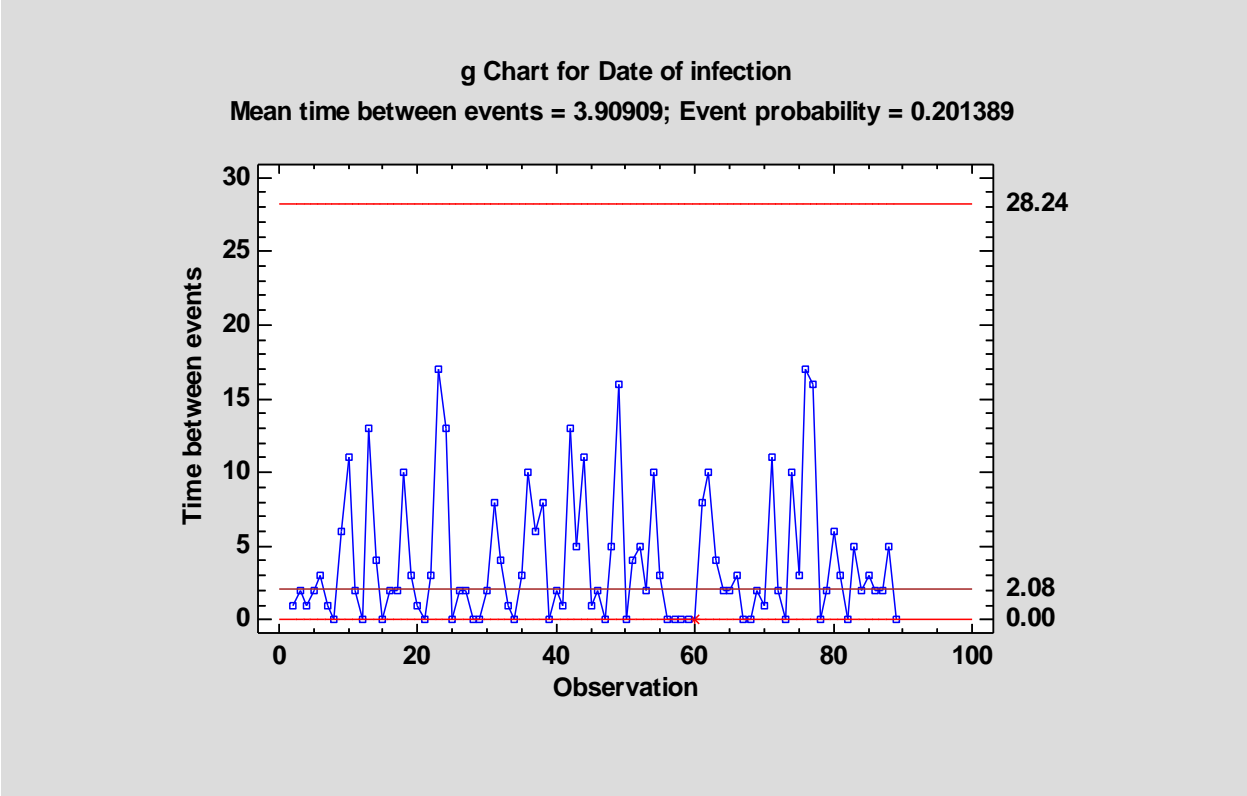
This new procedure creates a control chart for the length of time between the occurrence of rare events. It is based on the geometric distribution, which assumes that every time an opportunity for an event arises, there is a probability p that the event will occur. This type of chart is used frequently in the health sciences to monitor the occurrence of events such as post-surgical infections.

Input data may consist of either the times at which events occurred or the length of time between events:

The screenshot shows the 'g Chart' dialog box. The 'Data:' field is set to 'Date of infection'. The 'Type of data' section has 'Time of occurrence' selected. There are input fields for '(Labels:)', '(LSL:)', '(Nominal:)', and '(USL:)', each with a play button. A '(Select:)' field with a play button is also present. The 'Sort column names' checkbox is unchecked. The bottom buttons are 'OK', 'Cancel', 'Delete', 'Transform...', and 'Help'.

Specification limits may also be added to the chart.

The control chart is similar to others generated by Statgraphics:



The graph also displays the observed mean time between events and the estimated probability of an event at each point in time.

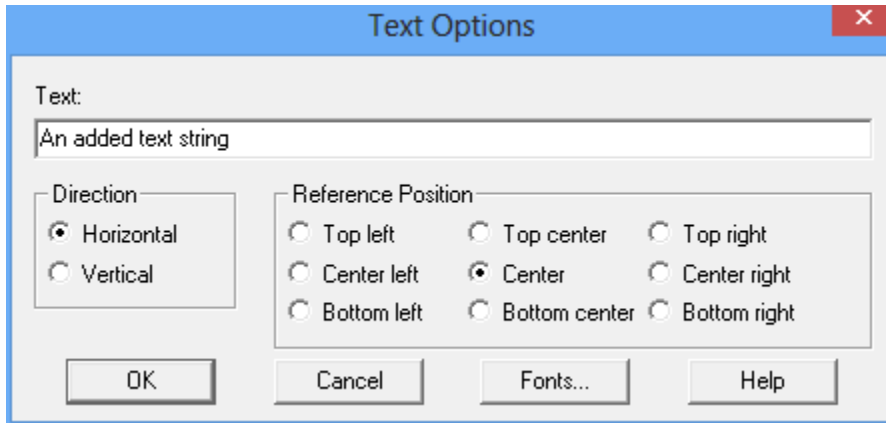
In addition to checking for standard runs rules violations, the program will check consecutive zeroes at a specified alpha level (the default value is 5 consecutive zeroes corresponding to $\alpha = 0.135\%$.)

Graphics Options

Several significant enhancements have been made to the graphics created by Statgraphics Centurion.


Added text

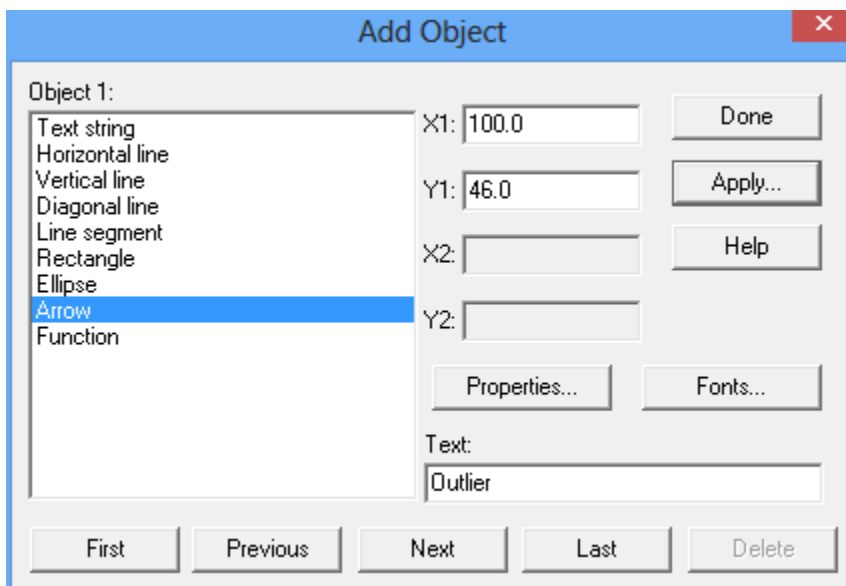
When new text strings are added to a graph, the reference position for the text may now be specified:



The reference position is the location on the text string that remains fixed when the size of the graph is changed. This makes it easier to position text relative to objects such as points.

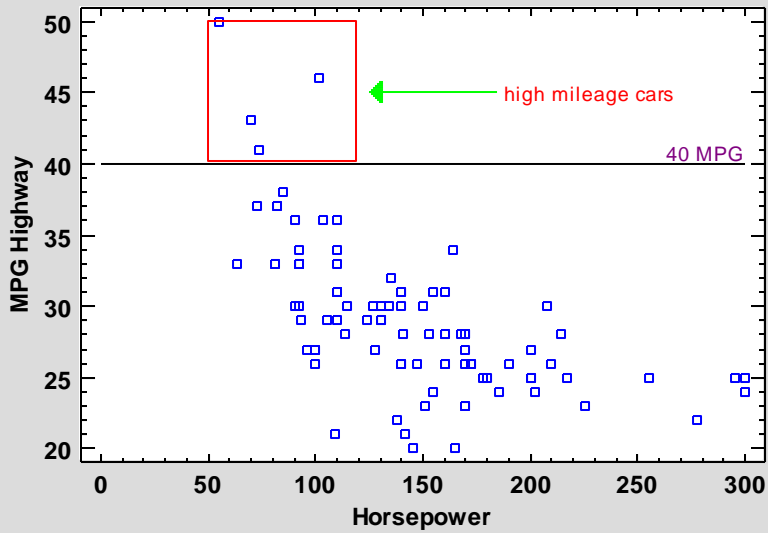
Adding Objects

A new button  has been added to the analysis toolbar that allows objects to be added to any graph. A dialog box is used to select the object to be added:



Objects that may be added include text strings (alone, within rectangles or ellipses, or next to an arrow), lines and line segments, arrows and functions. The plot below shows an X-Y Plot with added objects:

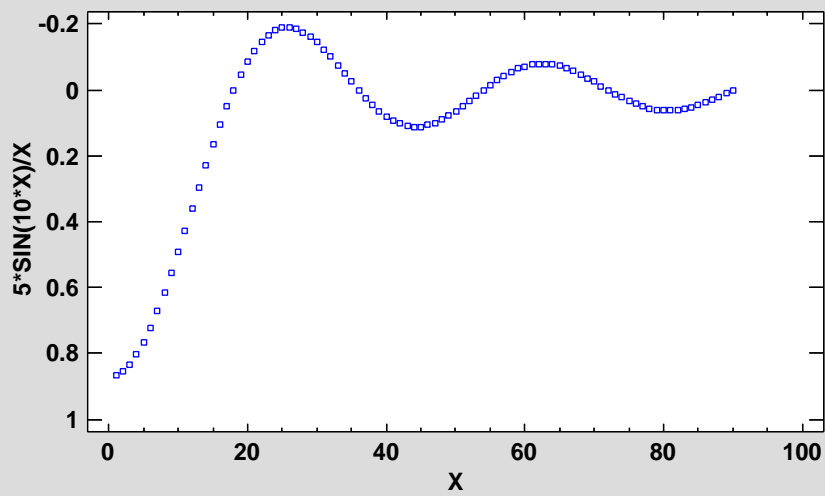
Plot of MPG Highway vs Horsepower

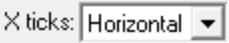


Axis tickmarks

Axis tickmarks may be reversed so that larger values are on the left (or top) and smaller values are on the right (or bottom):

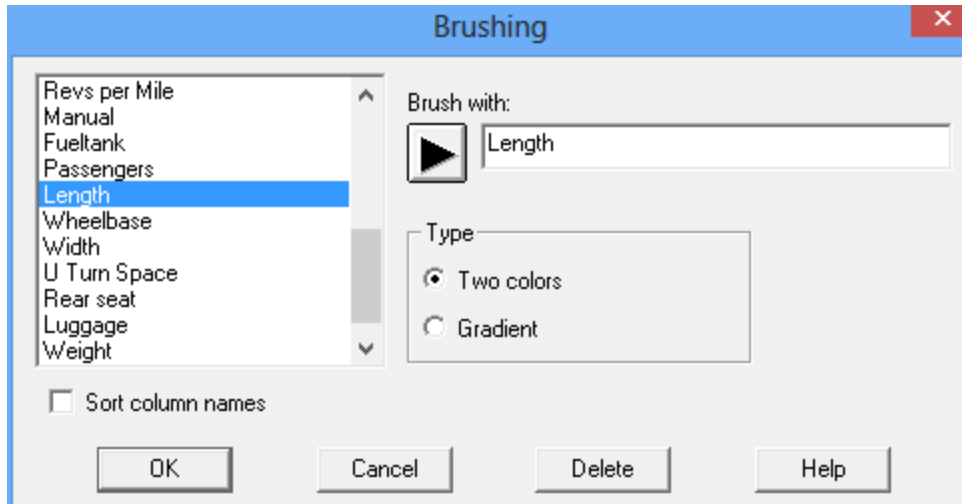
Plot of $5 \cdot \text{SIN}(10 \cdot X)/X$ vs X



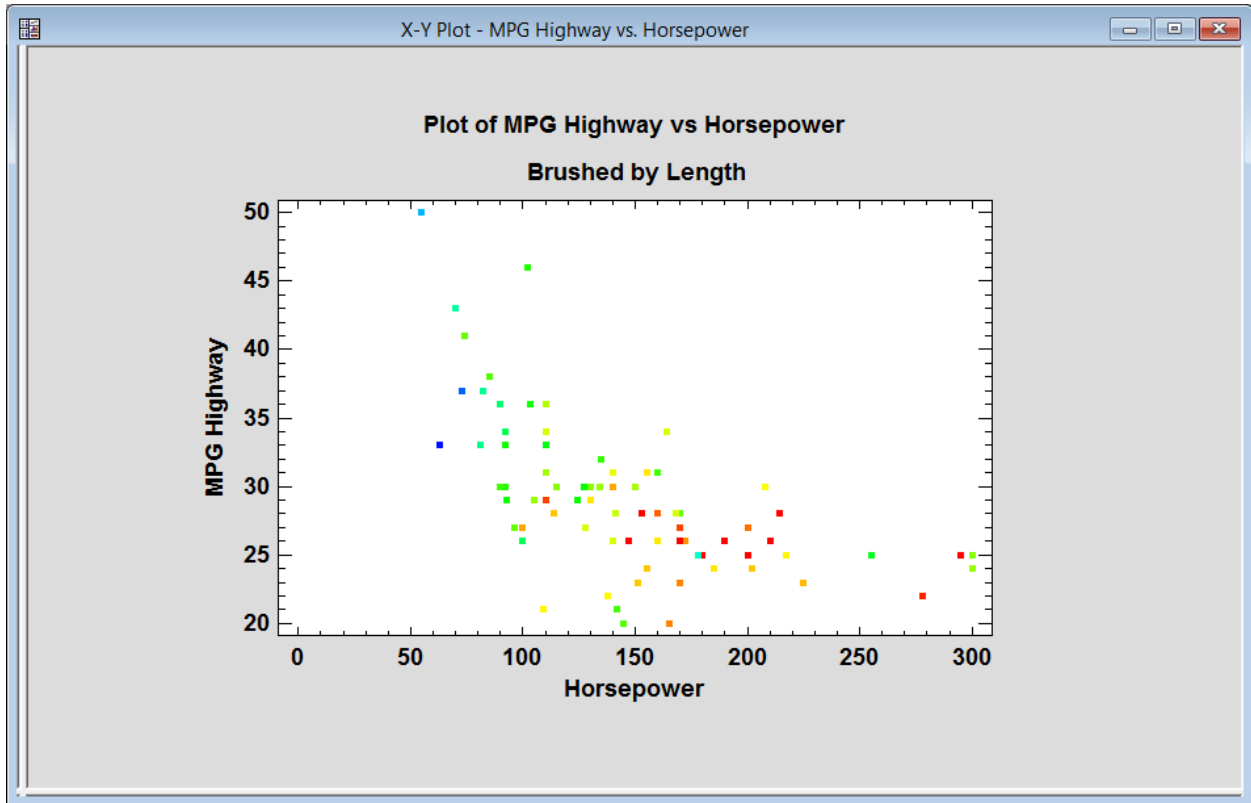
A new control on the analysis toolbar  also allows x-axis tickmarks to be rotated without having to access the *Graphics Options* dialog box.

Brushing scatterplots

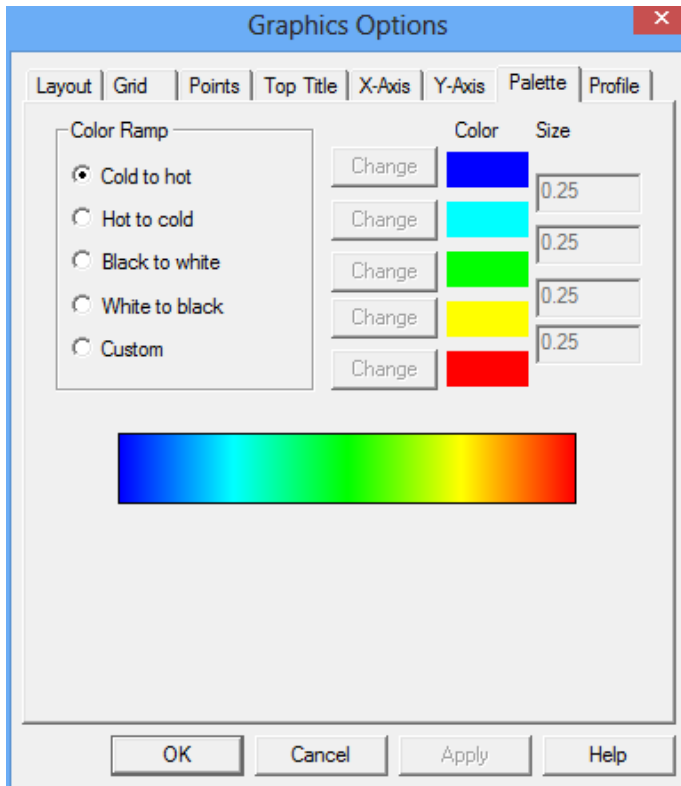
Scatterplots may now be brushing using either 2 colors (as in Version 16) or a color gradient:



When brushed using a gradient, each point is colored based on its location with respect to the current color palette (the default palette runs from blue at the minimum value to red at the maximum value):




The palette may be changed using *Graphics Options*:

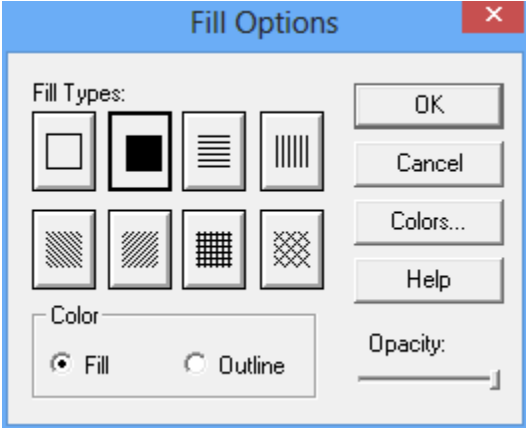


Dashed and Dotted Lines

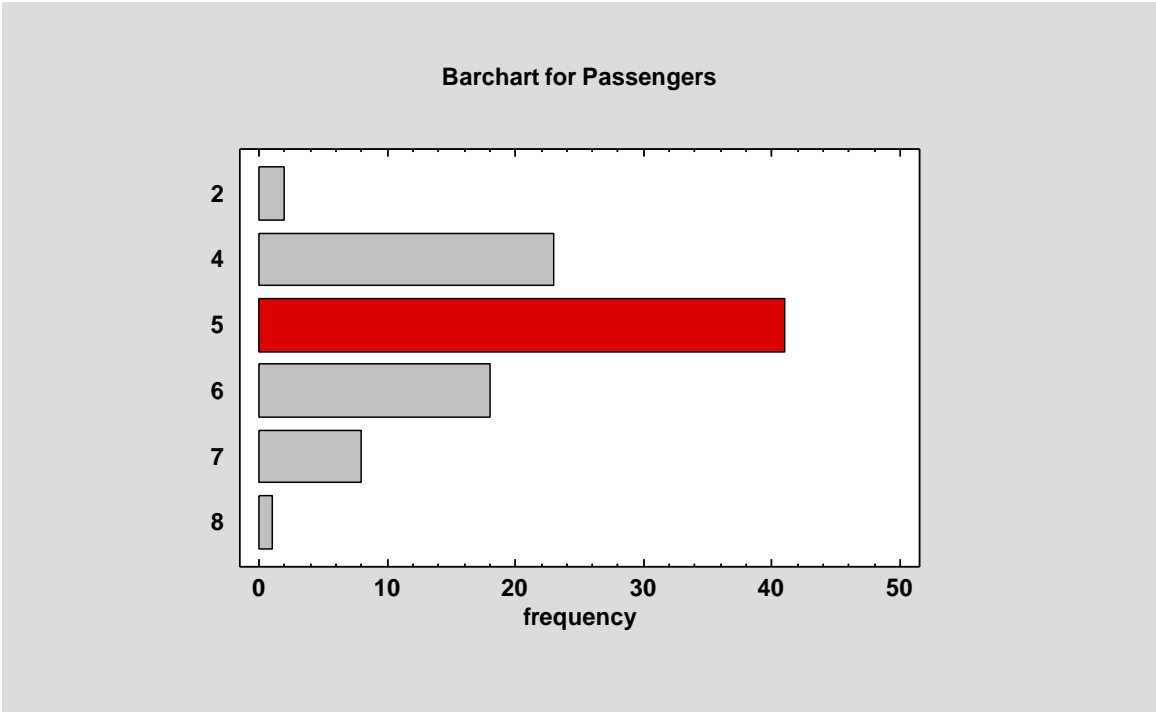
The width of dashed and dotted lines may now be increased. In Version 16, the width attribute applied only to solid lines.

Overriding attributes


A new button  has been added to the analysis toolbar to allow users to change the attributes of a single point, line, rectangle, or polygon. After clicking on an object, pushing the button displays a dialog box relevant for the object clicked:

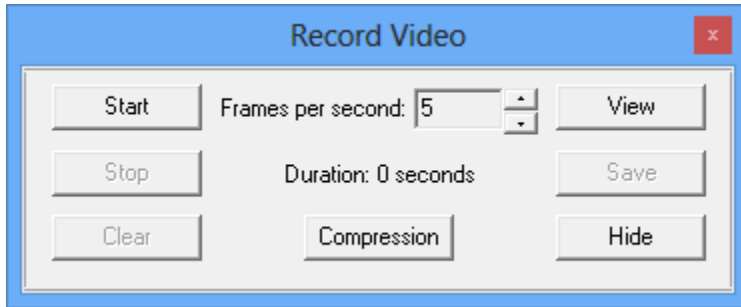


This can be useful for emphasizing selected objects:




Recording videos

A new button  has been added to the analysis toolbar to assist in recording videos. It displays the following dialog box:



The *Start* and *Stop* buttons are used to begin and end the recording. The *View* button plays back the last recorded video.

Text fonts

A new control has been added to the analysis toolbar to assist in increasing or decreasing the size of all text on a graph:  Changes occur immediately.

Zooming and panning

New buttons have been added to the analysis toolbar to facilitate the process of zooming and panning whenever a graph is displayed:



The + and – buttons expand or contract the current graph around its center. These buttons work in conjunction with the scrollbars along the margins of each graph, which are used for panning back and forth.

Another feature added to Version 17 is the automatic appearance of additional tickmarks as the graph is expanded.

Hypothesis Tests (Describe)

A new test has been added for the value of the Pearson correlation coefficient:

The screenshot shows a dialog box titled "Hypothesis Tests". Under the "Parameter" section, "Correlation Coefficient" is selected. The "Null Hypothesis" is set to 0.0. The "Sample Mean" is 0.0, "Sample Sigma" is 1.0, "Sample Proportion" is 0.5, and "Sample Rate" is 1.0. The "Sample Correlation" is 0.5 and the "Sample Size" is 100. Buttons for "OK", "Cancel", and "Help" are present.

Both two-sided and one-sided tests are provided.

The output also contains a confidence interval or confidence bound for the correlation:

Hypothesis Tests

Sample correlation coefficient = 0.85

Sample size = 100

95.0% confidence interval for correlation coefficient: [0.78457,0.896708]

Null Hypothesis: correlation = 0.9

Alternative: not equal

Computed Z statistic = -2.12801

P-Value = 0.033336

Reject the null hypothesis for alpha = 0.05.

Hypothesis Tests (Compare)

A new test has been added to compare the values of two Pearson correlation coefficients:

Hypothesis Tests (Compare)

Compare

- Normal Means
- Normal Sigmas
- Binomial Proportions
- Poisson Rates
- Correlation Coefficients

OK
Cancel
Help

Null Hypothesis for Difference of Correlation Coefficients:
0.0

Sample 1 Mean: 0.0 Sample 2 Mean: 0.0

Sample 1 Sigma: 1.0 Sample 2 Sigma: 1.0

Sample 1 Proportion: 0.5 Sample 2 Proportion: 0.5

Sample 1 Rate: 1.0 Sample 2 Rate: 1.0

Sample 1 Correlation: 0.56 Sample 2 Correlation: 0.62

Sample 1 Size: 100 Sample 2 Size: 100

Both two-sided and one-sided tests are provided.

Hypothesis Tests

Sample correlations = 0.56 and 0.62

Sample sizes = 100 and 100

Null Hypothesis: difference between correlations = 0.0

Alternative: not equal

Computed Z statistic = -0.641903

P-Value = 0.520934

Do not reject the null hypothesis for alpha = 0.05.

Item Reliability Analysis

This new procedure is designed to estimate the reliability or consistency of a set of variables. It is commonly used to assess how well a set of questions in a survey, each of which is designed to illicit information about the same characteristic, give consistent results.

The major output of the procedure is Cronbach's alpha. Alpha may be calculated directly from the input variables, or the variables may first be standardized so that they have equal variances. The effect on alpha when each variable is separately omitted is also estimated, so that unreliable questions can be identified.

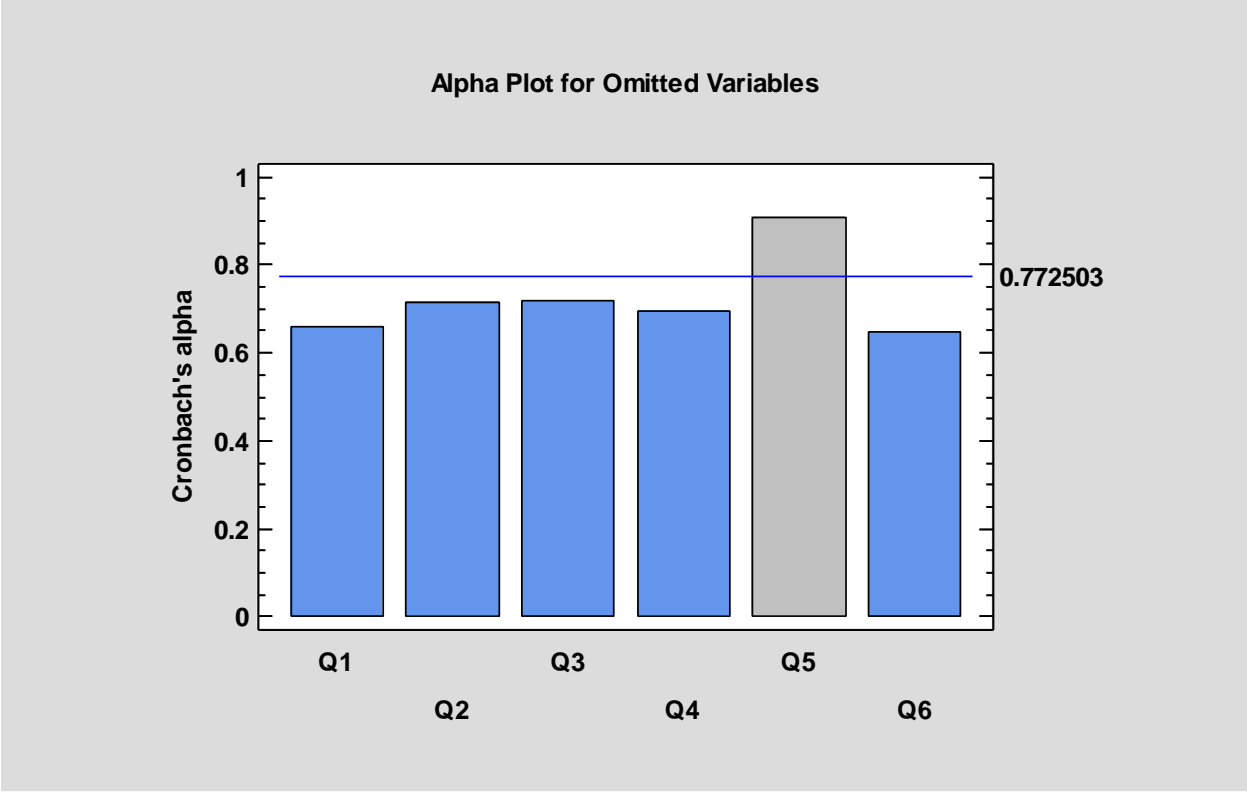
<i>Variable</i>	<i>Count</i>	<i>Sample Mean</i>	<i>Std. Deviation</i>
Q1	15	3.4	1.29835
Q2	15	3.33333	1.17514
Q3	15	3.53333	1.0601
Q4	15	4.06667	0.798809
Q5	15	3.06667	1.38701
Q6	15	3.73333	1.27988
Sum	15	21.1333	4.85308

Cronbach's alpha = 0.772503 (lower 95% confidence bound = 0.582388)

Omitted Item Statistics

<i>Omitted</i>	<i>Adj. Sum</i>	<i>Adj. Sum</i>	<i>Item-Total</i>	<i>Squared</i>	<i>Alpha if</i>
<i>Variable</i>	<i>Mean</i>	<i>Std. Deviation</i>	<i>Correlation</i>	<i>Multiple R</i>	<i>Omitted</i>
Q1	17.7333	3.76955	0.782267	0.742626	0.660188
Q2	17.8	4.03909	0.616996	0.728303	0.712931
Q3	17.6	4.13694	0.605884	0.595351	0.719254
Q4	17.0667	4.18273	0.810943	0.738456	0.696108
Q5	18.0667	4.84719	-0.138824	0.451907	0.905959
Q6	17.4	3.73784	0.83015	0.704872	0.645876

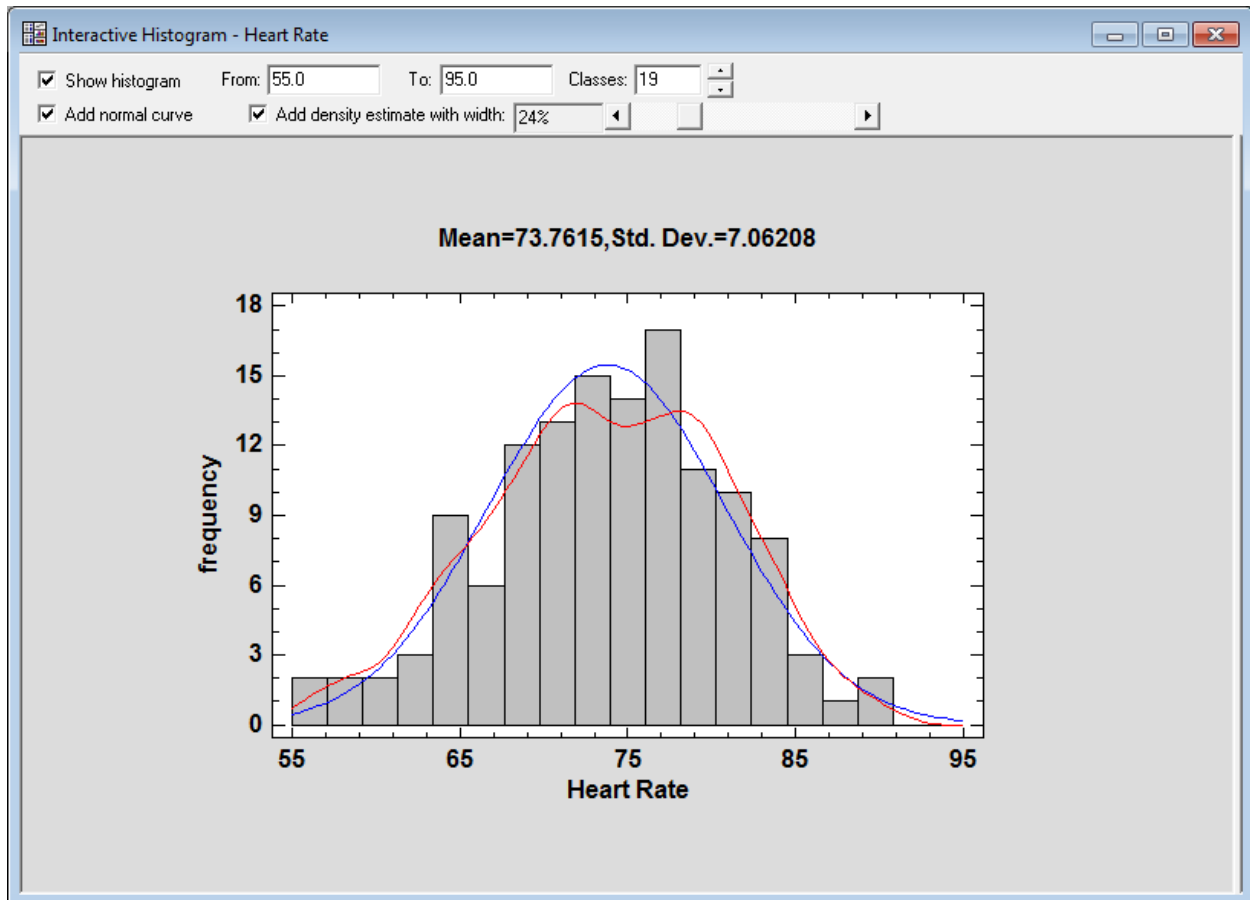
The *Alpha Plot* shows how alpha would change if individual questions were deleted:



Large increases in alpha, as for Q5 above, indicate that the corresponding question is not reliable.

Interactive Histogram Statlet

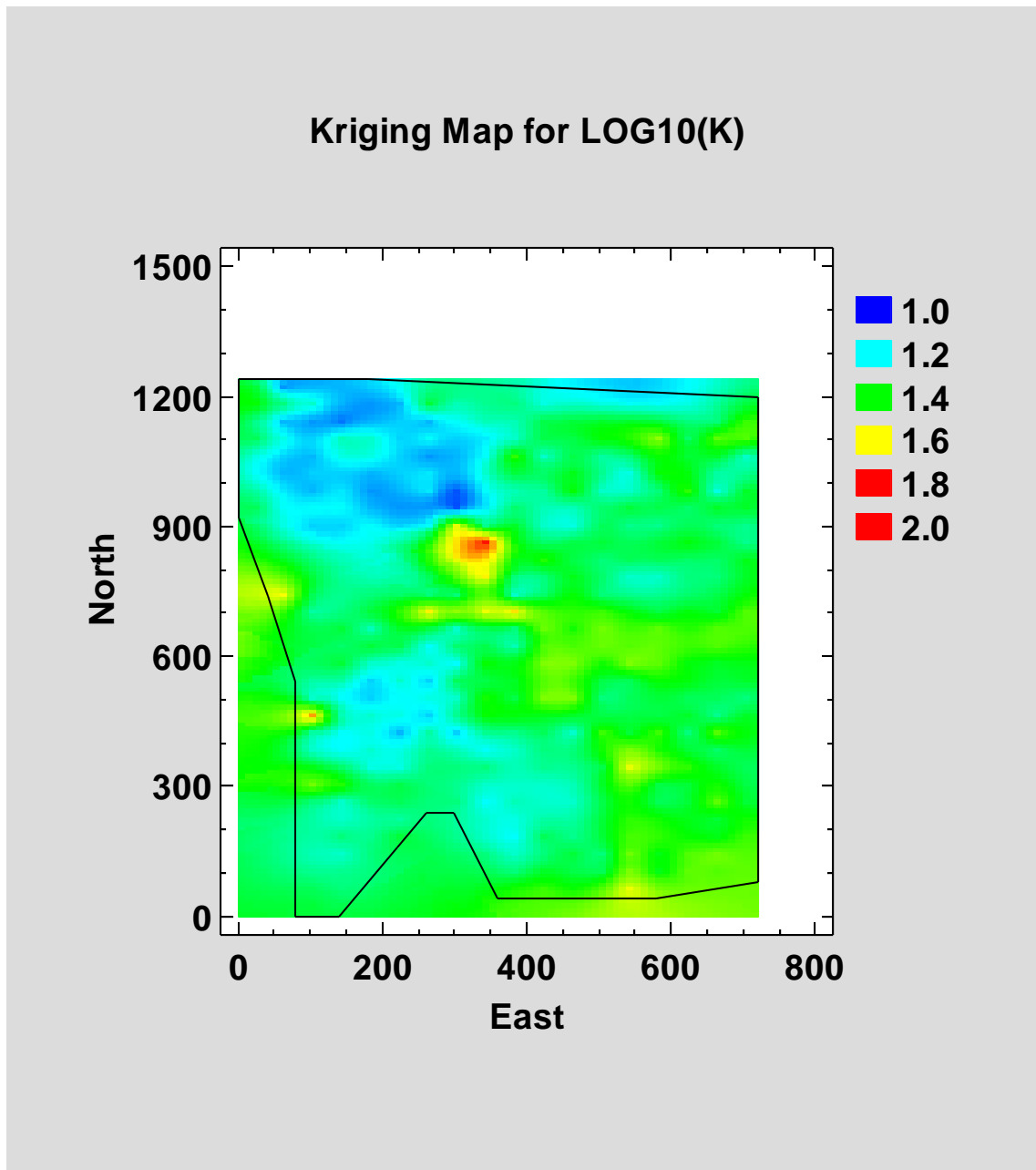
This new Statlet creates a frequency histogram for a column of numeric data. The controls on the toolbar make it easy to change the definition of the classes into which the data are grouped. The density function of a normal distribution with the same mean and standard deviation as the data may be superimposed on the histogram. In addition, a nonparametric density trace may be drawn.



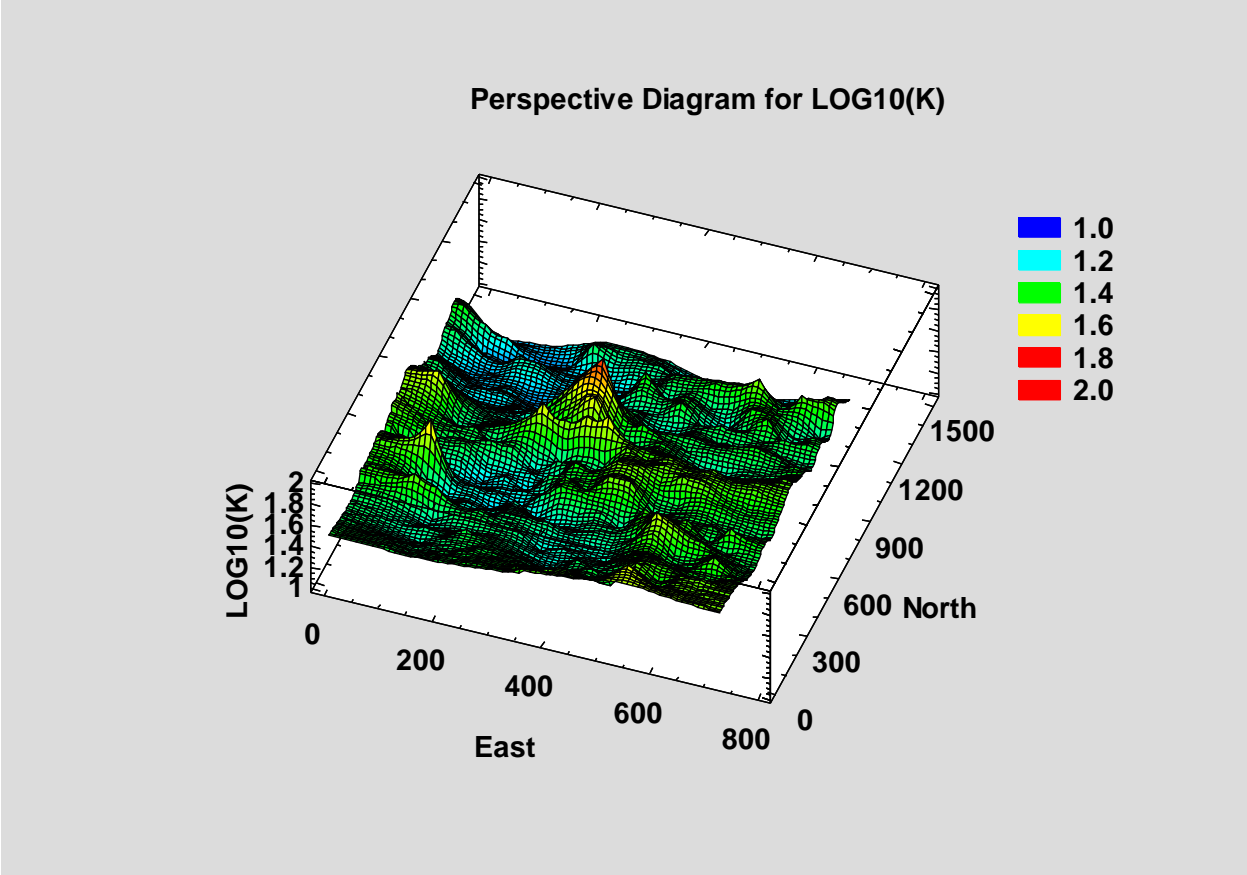
Kriging

A new procedure has been added for estimating the value of a random variable based on measurements made at locations distributed throughout a 2-dimensional region. Called Kriging, the procedure first creates a variogram to estimate the spatial dependence between measurements. Estimates are then made at unmeasured locations throughout the region.

The main output is a map similar to that shown below:

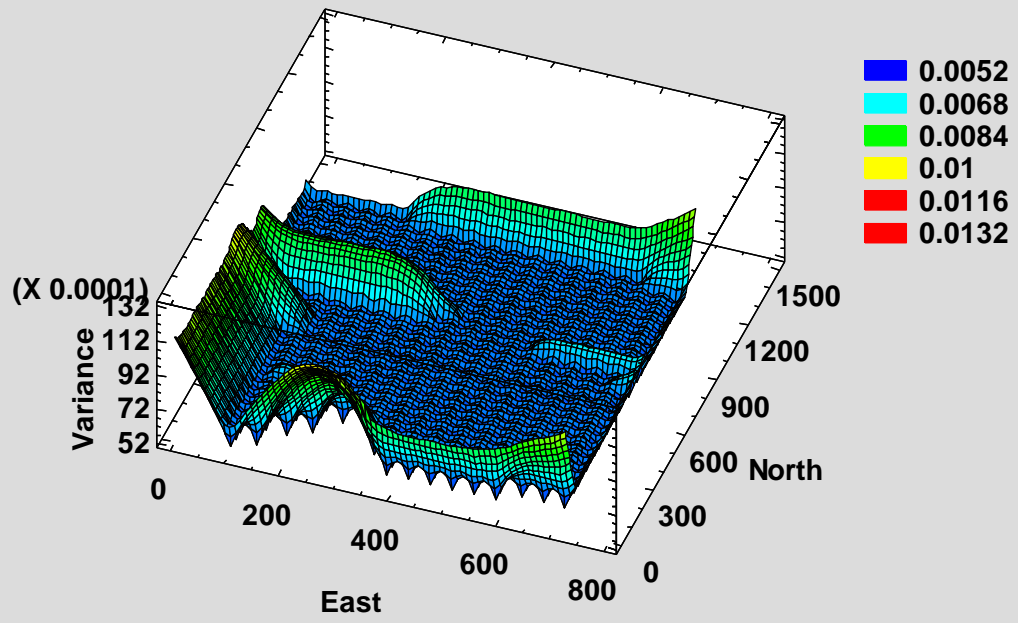


The estimates may also be viewed as a 3-dimensional perspective diagram:



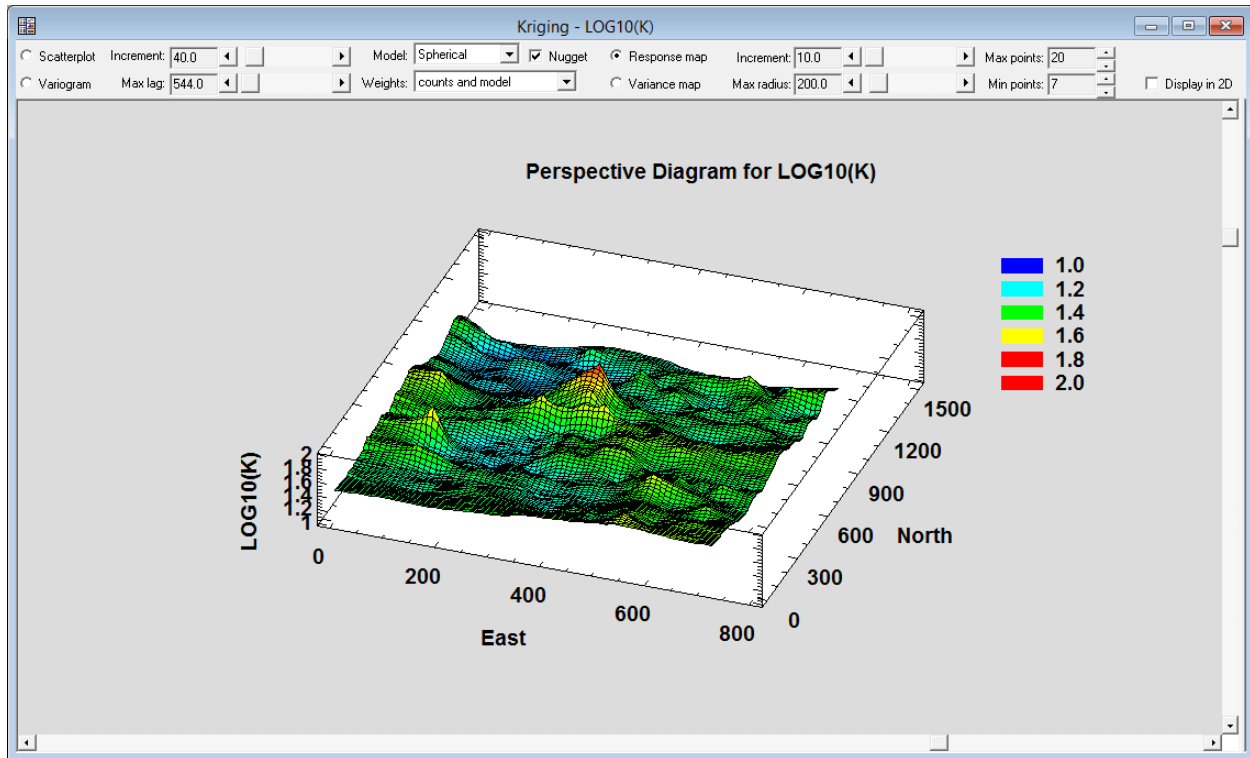
In addition to the estimates, the estimated variance is also obtained and plotted:

Perspective Diagram for Variance



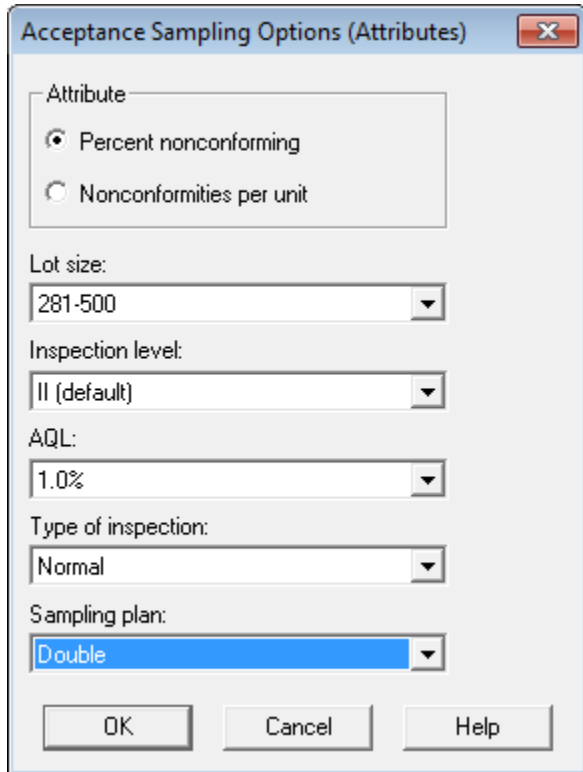
Kriging Statlet

A new Statlet has been added to visualize the effect of changing Kriging parameters:



MIL-STD-105E Acceptance Sampling for Attributes

This new procedure calculates required sample sizes for single, double and multiple sampling plans to be used with the ANSI Z1.4 (previously MIL-STD-105E) sampling methodology. The standard has been established for the acceptance or rejection of lots based on the evaluation of item attributes. Based on one or more samples taken from a batch or lot containing N units, the batch or lot is either accepted or rejected.



The image shows a dialog box titled "Acceptance Sampling Options (Attributes)". It contains several settings for attribute sampling:

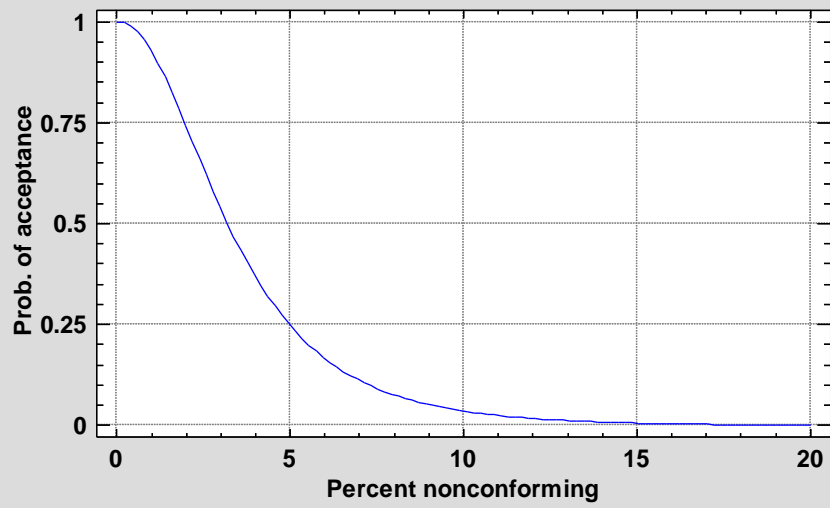
- Attribute:** A group box containing two radio buttons: "Percent nonconforming" (selected) and "Nonconformities per unit".
- Lot size:** A dropdown menu set to "281-500".
- Inspection level:** A dropdown menu set to "II (default)".
- AQL:** A dropdown menu set to "1.0%".
- Type of inspection:** A dropdown menu set to "Normal".
- Sampling plan:** A dropdown menu set to "Double".

At the bottom of the dialog are three buttons: "OK", "Cancel", and "Help".

The procedure also plots operating characteristic curves and average sample number curves to compare alternative sampling plans.

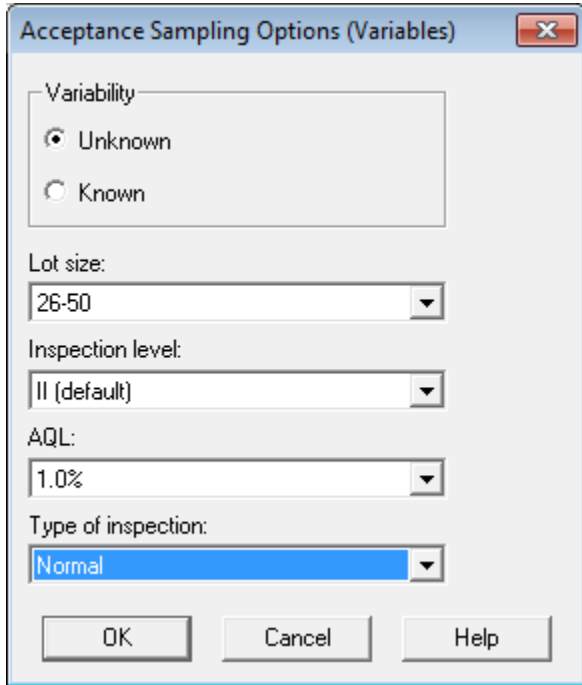
Operating Characteristic (OC) Curve

Lot size: 281-500; Inspection level: II; AQL: 1.0%; Type of inspection: Normal; Sampling plan: Double



MIL-STD-414 Acceptance Sampling for Variables

This new procedure calculates required sample sizes based on the ANSI Z1.9 (formerly MIL-STD-414) standard. ANSI Z1.9 is a standard sampling methodology that has been established for the acceptance or rejection of lots using measurements made on a sample of units taken from that lot. Based on a sample of n units taken from a batch or lot containing N items, the batch or lot is either accepted or rejected.

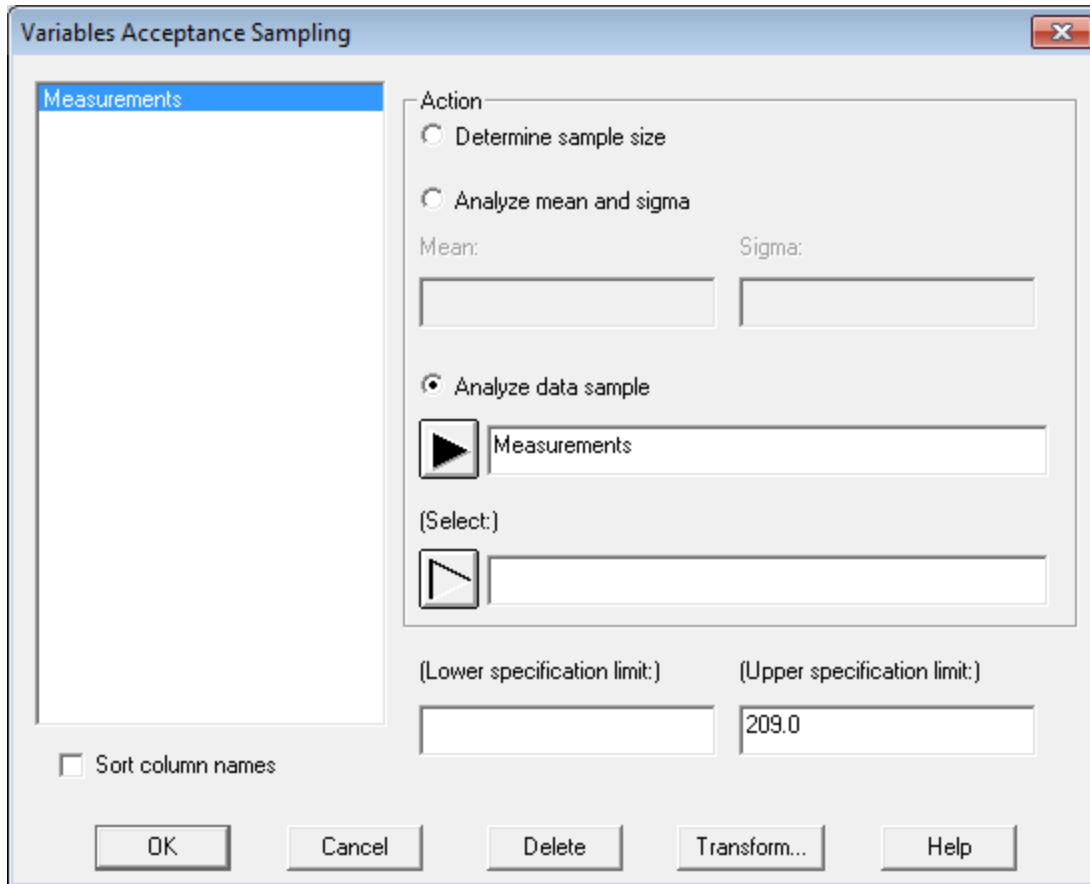


The image shows a dialog box titled "Acceptance Sampling Options (Variables)". It contains the following settings:

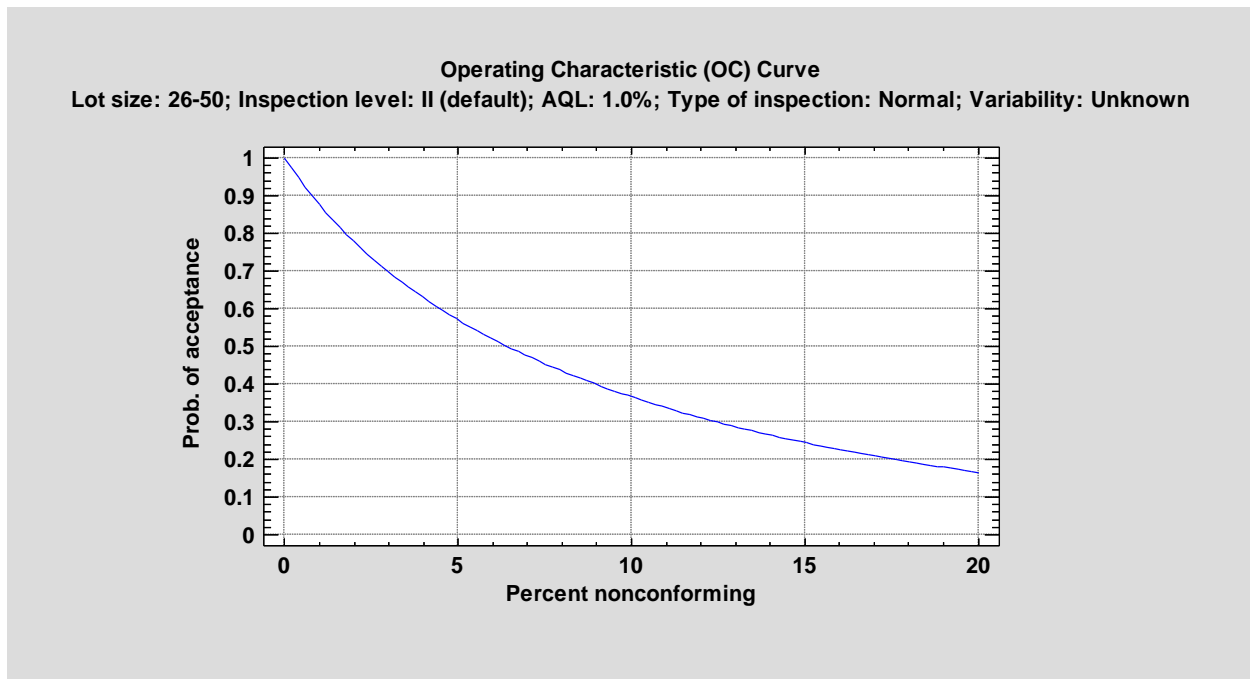
- Variability:** Radio buttons for "Unknown" (selected) and "Known".
- Lot size:** A dropdown menu set to "26-50".
- Inspection level:** A dropdown menu set to "II (default)".
- AQL:** A dropdown menu set to "1.0%".
- Type of inspection:** A dropdown menu set to "Normal".

At the bottom of the dialog box are three buttons: "OK", "Cancel", and "Help".

In addition to calculating the required sample sizes, the procedure will analyze a set of measurements and determine whether or not the lot should be accepted.

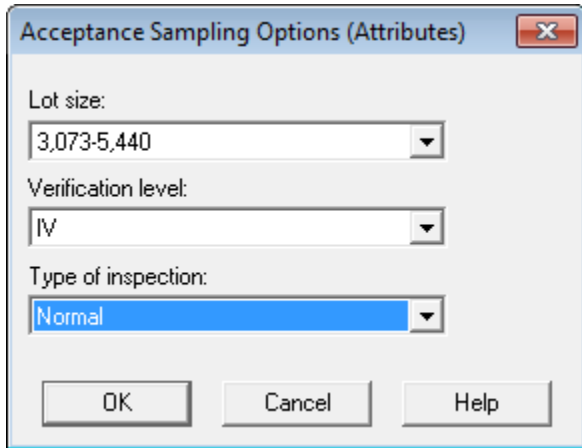


The procedure also plots operating characteristic curves to compare alternative sampling plans.



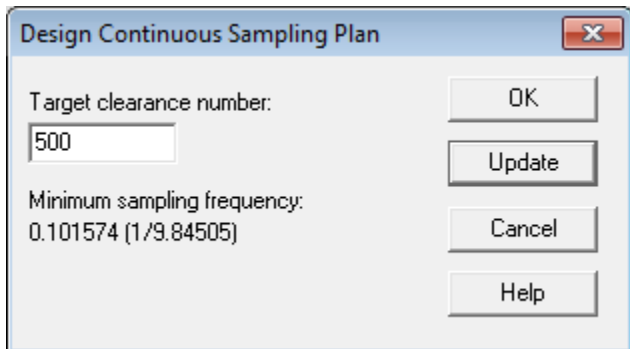
MIL-STD-1916 Acceptance Sampling for Attributes

This procedure calculates the required sample sizes needed to implement MIL-STD-1916. MIL-STD-1916 is a standard sampling methodology that has been established for the acceptance or rejection of lots based on the evaluation of item attributes. A sample of n units is taken from a batch or lot containing N units. If the sample contains no nonconforming units, the batch or lot is accepted. Otherwise, the batch or lot is rejected. Such a plan is often referred to as a “zero acceptance number sampling plan”. The MIL-STD-1916 standard specifies the appropriate sample sizes for different verification levels, which are usually specified by contract.



The dialog box titled "Acceptance Sampling Options (Attributes)" contains three dropdown menus. The first is labeled "Lot size:" and has the value "3,073-5,440". The second is labeled "Verification level:" and has the value "IV". The third is labeled "Type of inspection:" and has the value "Normal". At the bottom of the dialog are three buttons: "OK", "Cancel", and "Help".

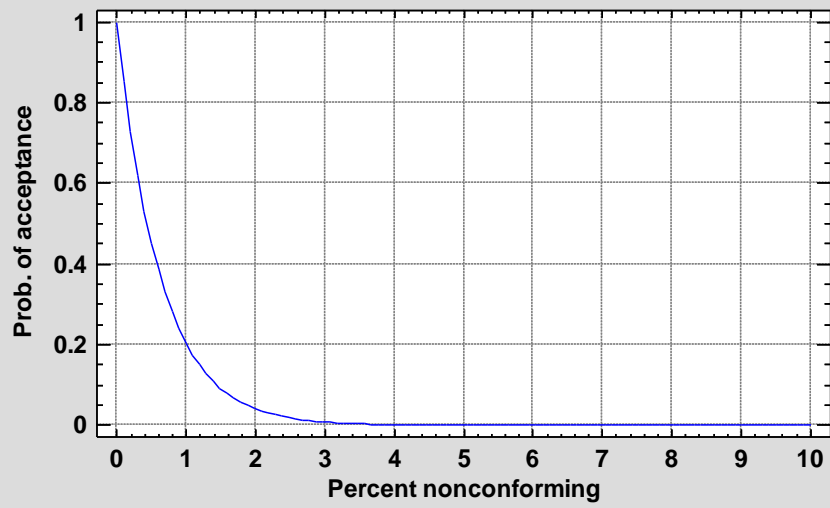
It also determines the required sampling frequency for implementing continuous sampling plans.



The dialog box titled "Design Continuous Sampling Plan" contains a text input field for "Target clearance number:" with the value "500". Below it, the text "Minimum sampling frequency:" is followed by the value "0.101574 (1/9.84505)". On the right side of the dialog are four buttons: "OK", "Update", "Cancel", and "Help".

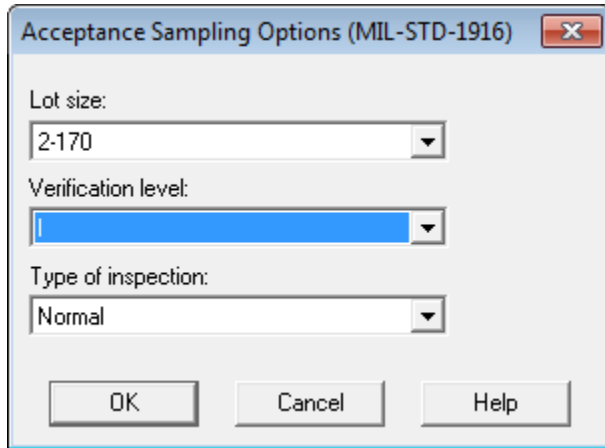
The procedure also plots operating characteristic curves to compare alternative sampling plans.

Operating Characteristic (OC) Curve
Lot size: 3,073-5,440; Verification Level: IV; Type of inspection: Normal



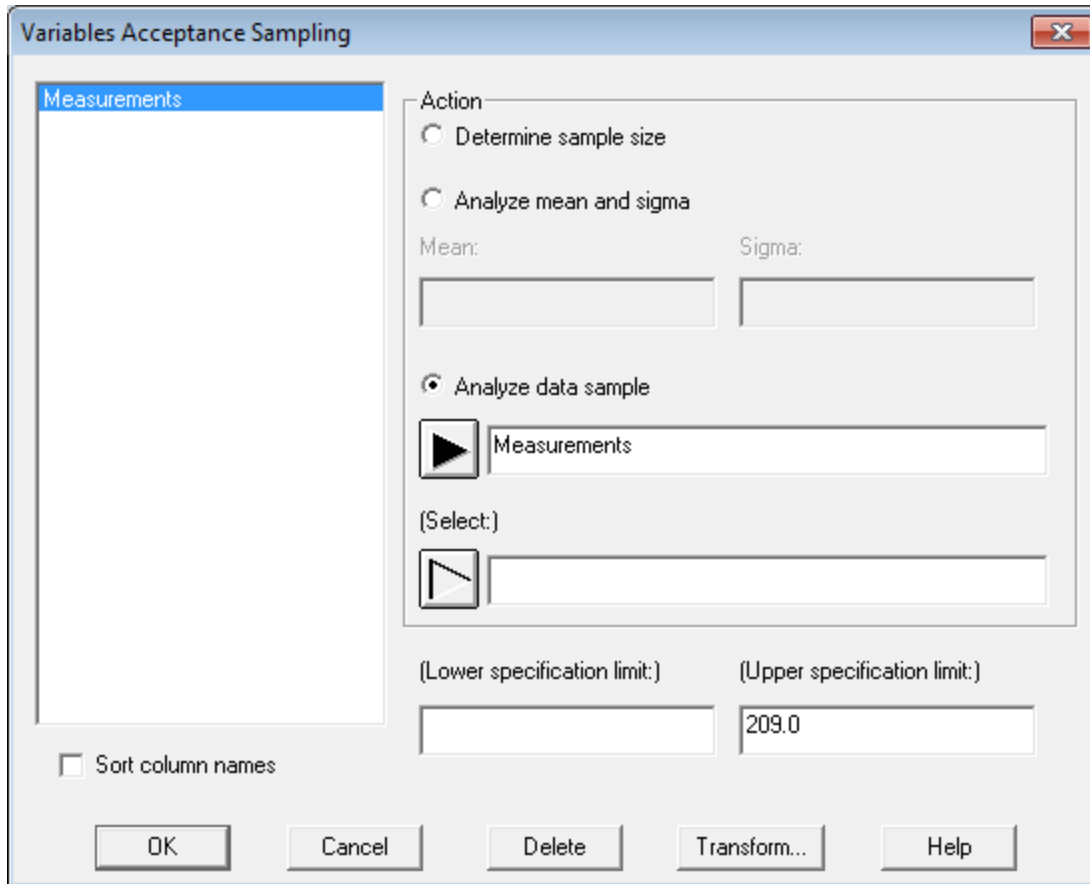
MIL-STD-1916 Acceptance Sampling for Variables

This procedure calculates the required sample sizes needed to implement MIL-STD-1916. MIL-STD-1916 is a standard sampling methodology that has been established for the acceptance or rejection of lots based on the evaluation of item measurements. A sample of n units is taken from a batch or lot containing N units. If the sample contains no nonconforming units and the sample mean and sample standard deviation yield acceptable quality indices, the batch or lot is accepted. Otherwise, the batch or lot is rejected. To calculate the quality indices, either one or two specification limits for the measurements must be entered.



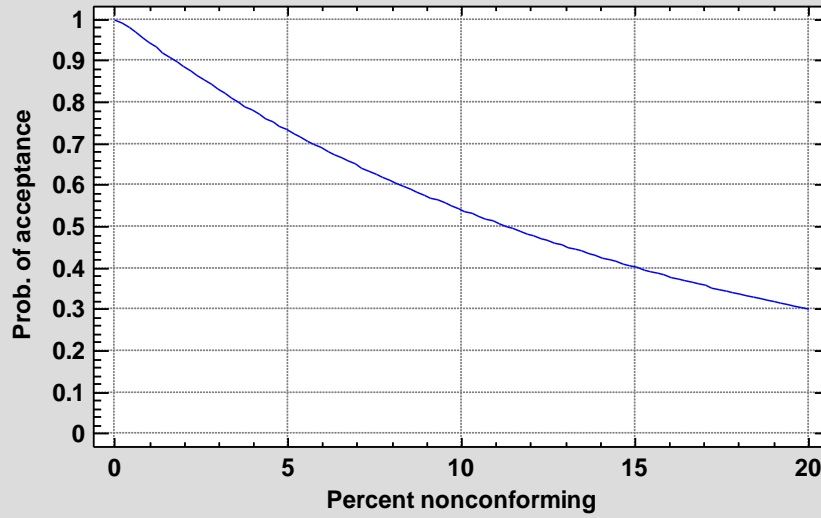
The image shows a dialog box titled "Acceptance Sampling Options (MIL-STD-1916)". It contains three dropdown menus: "Lot size" with the value "2-170", "Verification level" with the value "I", and "Type of inspection" with the value "Normal". At the bottom of the dialog box are three buttons: "OK", "Cancel", and "Help".

In addition to calculating the required sample sizes, the procedure will analyze a set of measurements and determine whether or not the lot should be accepted.



The procedure also plots operating characteristic curves to compare alternative sampling plans.

Operating Characteristic (OC) Curve
Lot size: 2-170; Verification Level: I; Type of inspection: Normal



Multiple Sample Comparison

A new method has been added for simultaneously comparing several means. Called the *Games-Howell* procedure, it is similar to the Tukey procedure, except that it does not assume equal variances within each group. This is a good choice if the *Variance Check* indicates significant differences amongst the group variances.

Multiple Range Tests			
Method: 95.0 percent Games-Howell			
	Count	Mean	Homogeneous Groups
Task 6	11	28.8182	X
Task 5	12	29.5	XX
Task 2	12	31.0833	XX
Task 1	13	31.9231	XX
Task 3	10	35.8	XX
Task 4	10	38.0	X
Contrast	Sig.	Difference	+/- Limits
Task 1 - Task 2		0.839744	7.0518
Task 1 - Task 3		-3.87692	6.80671
Task 1 - Task 4		-6.07692	6.19608
Task 1 - Task 5		2.42308	7.0067
Task 1 - Task 6		3.1049	8.16719
Task 2 - Task 3		-4.71667	7.42918
Task 2 - Task 4	*	-6.91667	6.91298
Task 2 - Task 5		1.58333	7.59217
Task 2 - Task 6		2.26515	8.60071
Task 3 - Task 4		-2.2	6.65731
Task 3 - Task 5		6.3	7.37977
Task 3 - Task 6		6.98182	8.43302
Task 4 - Task 5	*	8.5	6.86642
Task 4 - Task 6	*	9.18182	8.065
Task 5 - Task 6		0.681818	8.55813

* denotes a statistically significant difference.

The output of the Kruskal-Wallis test has also been expanded to include Bonferroni intervals for the difference between each pair of medians:

Kruskal-Wallis Test

	<i>Sample Size</i>	<i>Average Rank</i>
Task 1	13	33.3846
Task 2	12	30.5833
Task 3	10	46.4
Task 4	10	50.35
Task 5	12	26.7083
Task 6	11	23.3636

Test statistic = 15.9995 P-Value = 0.00684551

95.0 percent Bonferroni intervals

<i>Contrast</i>	<i>Sig.</i>	<i>Difference</i>	<i>+/- Limits</i>
Task 1 - Task 2		2.80128	23.2346
Task 1 - Task 3		-13.0154	24.4129
Task 1 - Task 4		-16.9654	24.4129
Task 1 - Task 5		6.67628	23.2346
Task 1 - Task 6		10.021	23.7774
Task 2 - Task 3		-15.8167	24.8512
Task 2 - Task 4		-19.7667	24.8512
Task 2 - Task 5		3.875	23.6947
Task 2 - Task 6		7.2197	24.2272
Task 3 - Task 4		-3.95	25.9562
Task 3 - Task 5		19.6917	24.8512
Task 3 - Task 6		23.0364	25.3595
Task 4 - Task 5		23.6417	24.8512
Task 4 - Task 6	*	26.9864	25.3595
Task 5 - Task 6		3.3447	24.2272

* denotes a statistically significant difference.

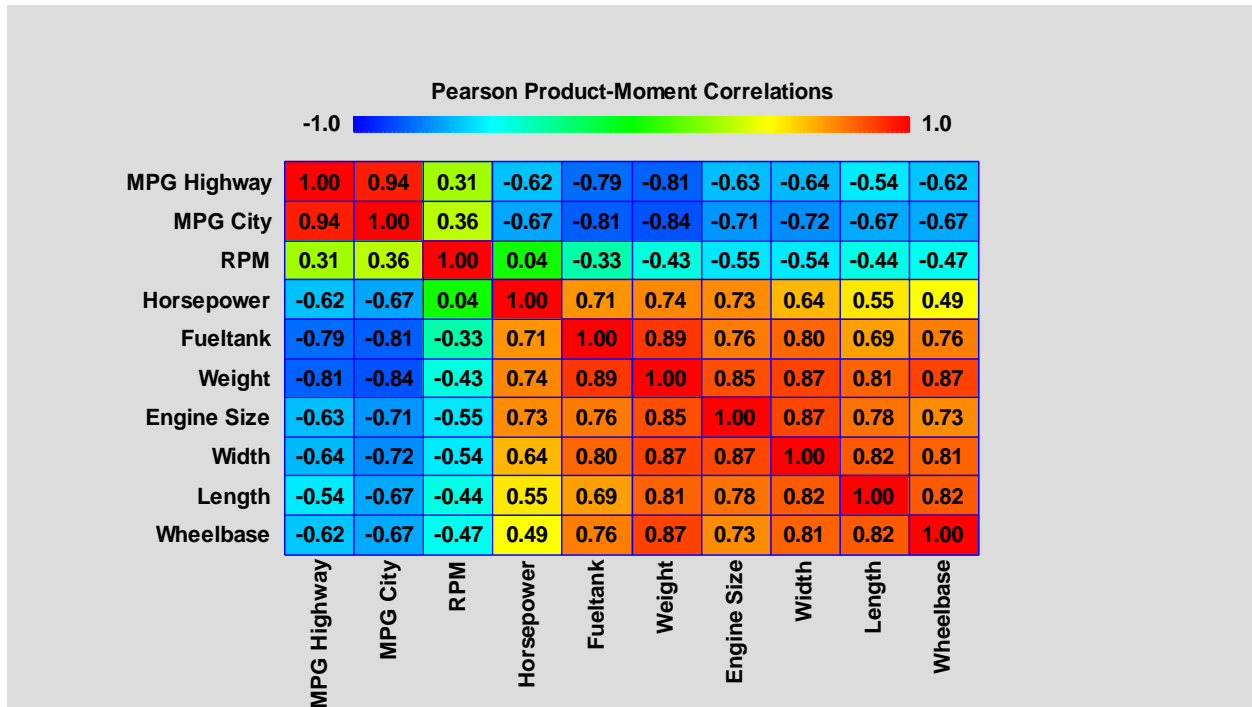
Bonferroni intervals permit the simultaneous comparison of many differences at a specified family confidence level.

Multiple Variable Analysis

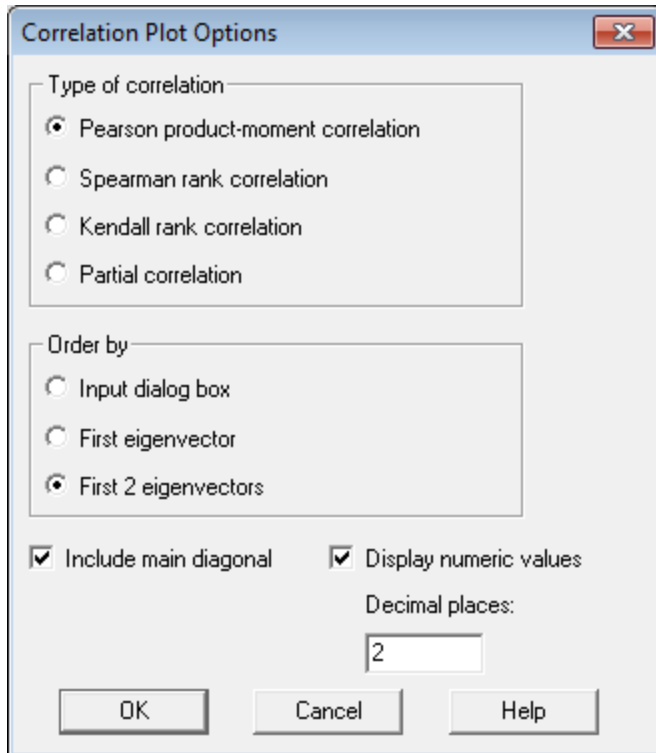
Two major enhancements have been made to the Multiple Variable Analysis procedure:

1. A *Correlation Plot* has been added to display correlations graphically.
2. The partial correlations now allow the user to specify which variables are partialled out.

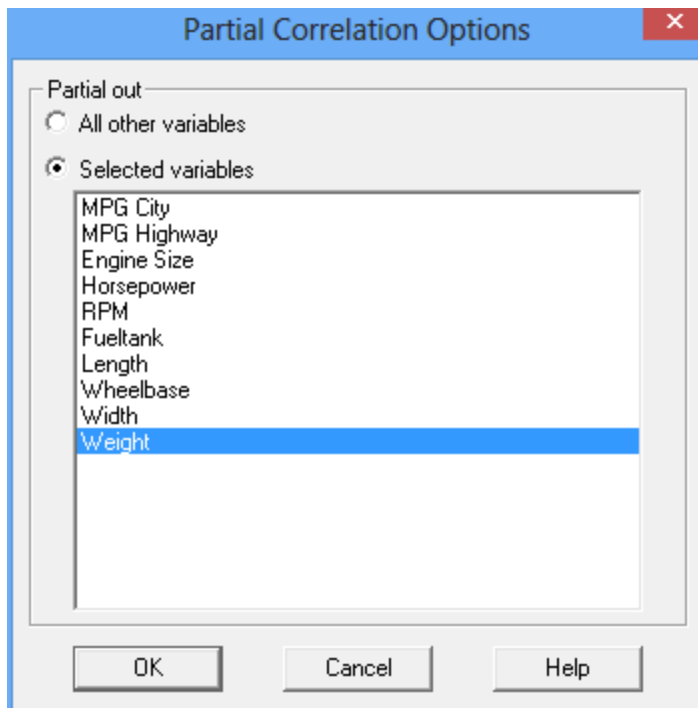
The Correlation Plot (also called a Corrgram) displays the estimated correlations or partial correlations in the form of a matrix with colored cells:



Various options are available to control the ordering of the variables:



Estimation of partial correlations has been enhanced by letting users specify the set of variables to be partialled out, rather than always partialing out all variables:



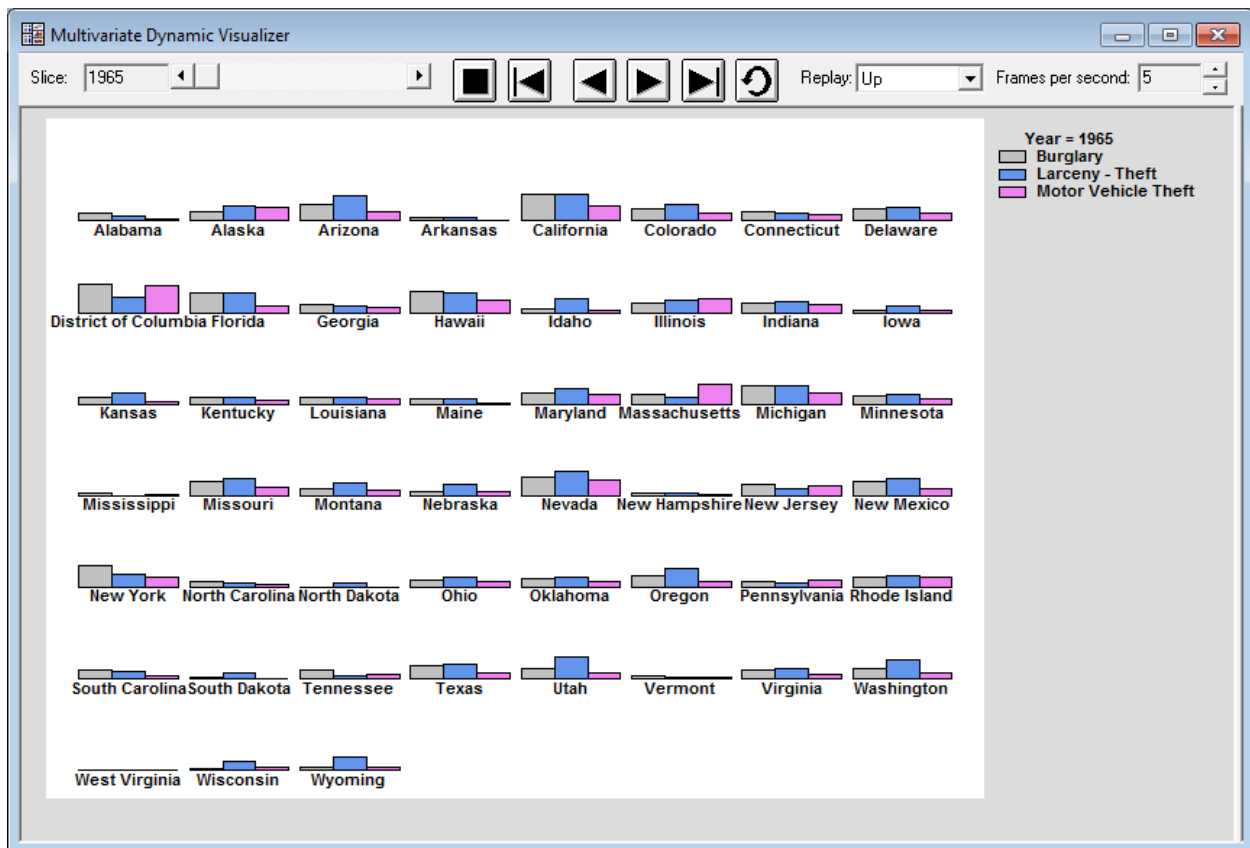
Multivariate Visualizer

This new Statlet is designed to plot multiple time series in a manner that helps users visualize the changes in multiple variables over time. Given n samples for each of m variables observed over p time periods, the program generates a dynamic display that illustrates how each of the variables has changed over time. Typical applications include plotting:

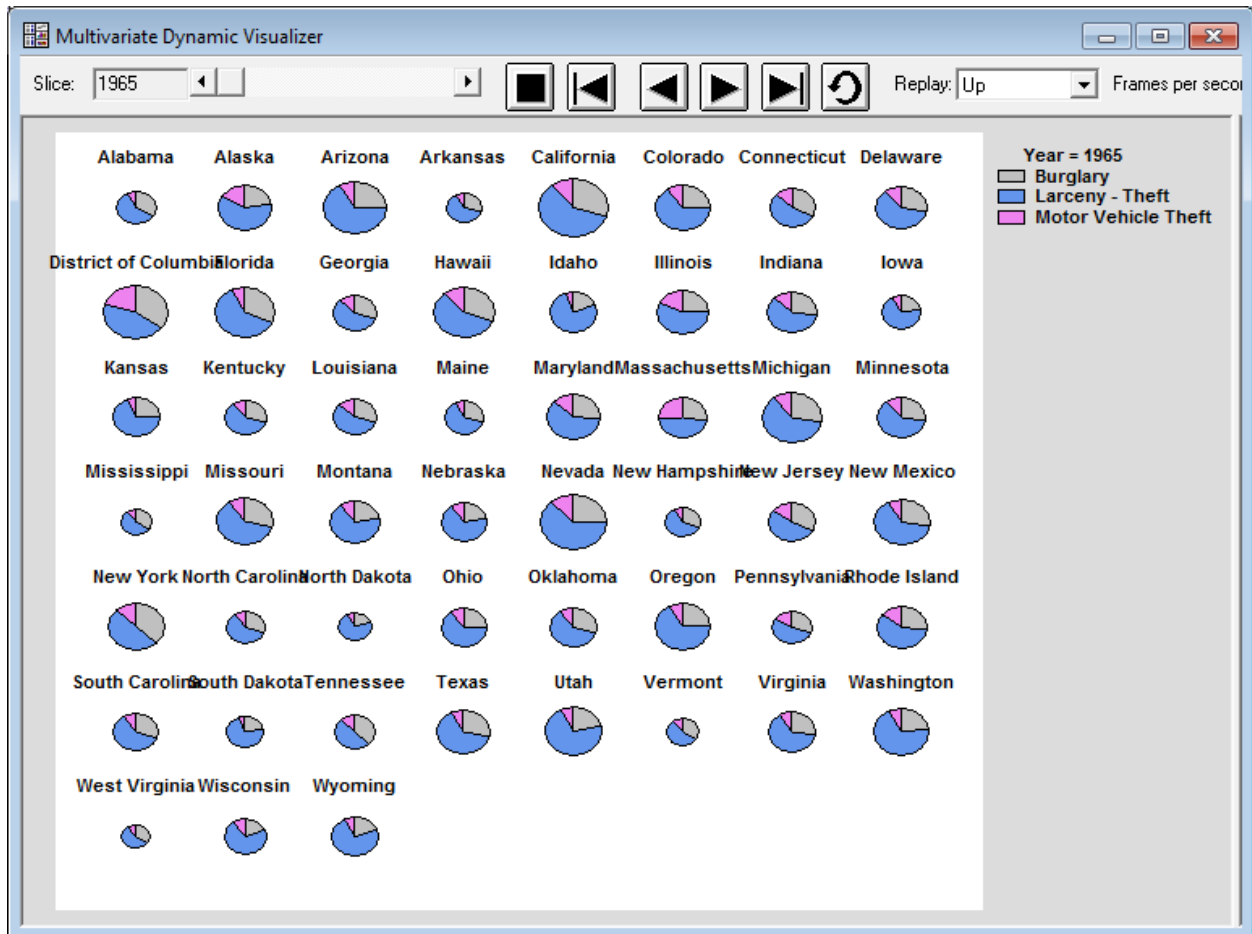
1. Yearly demographic variables for different countries.
2. Quarterly sales figures for multiple divisions within a company.
3. Monthly economic indices on a state-by-state basis.
4. Daily closing stock prices for multiple equities within a portfolio.

Several different types of plots may be created, including barcharts, piecharts, profile plots, strip plots, starplots, and Chernoff faces. As time evolves, the analyst can follow changes in all of the variables in all of the samples simultaneously. Various options are offered for smoothing the data and for dealing with missing values.

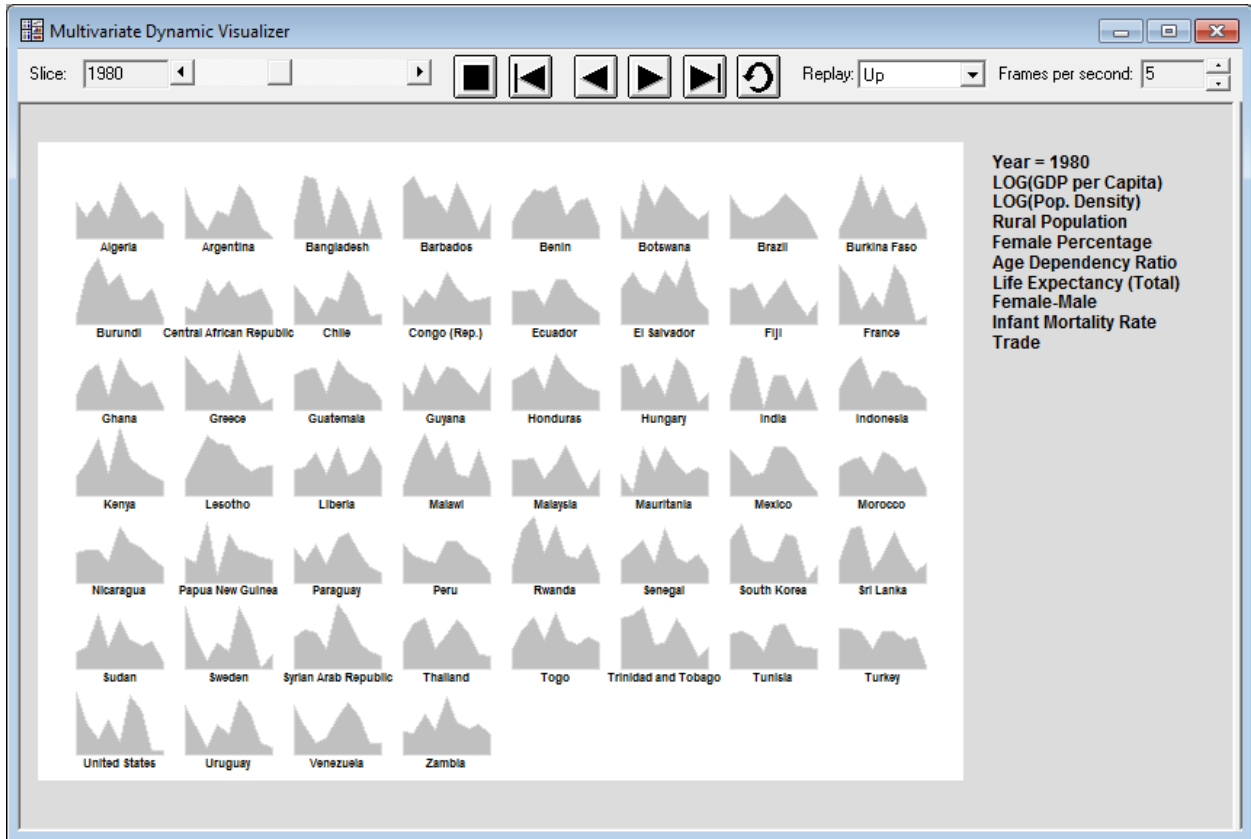
Barcharts:



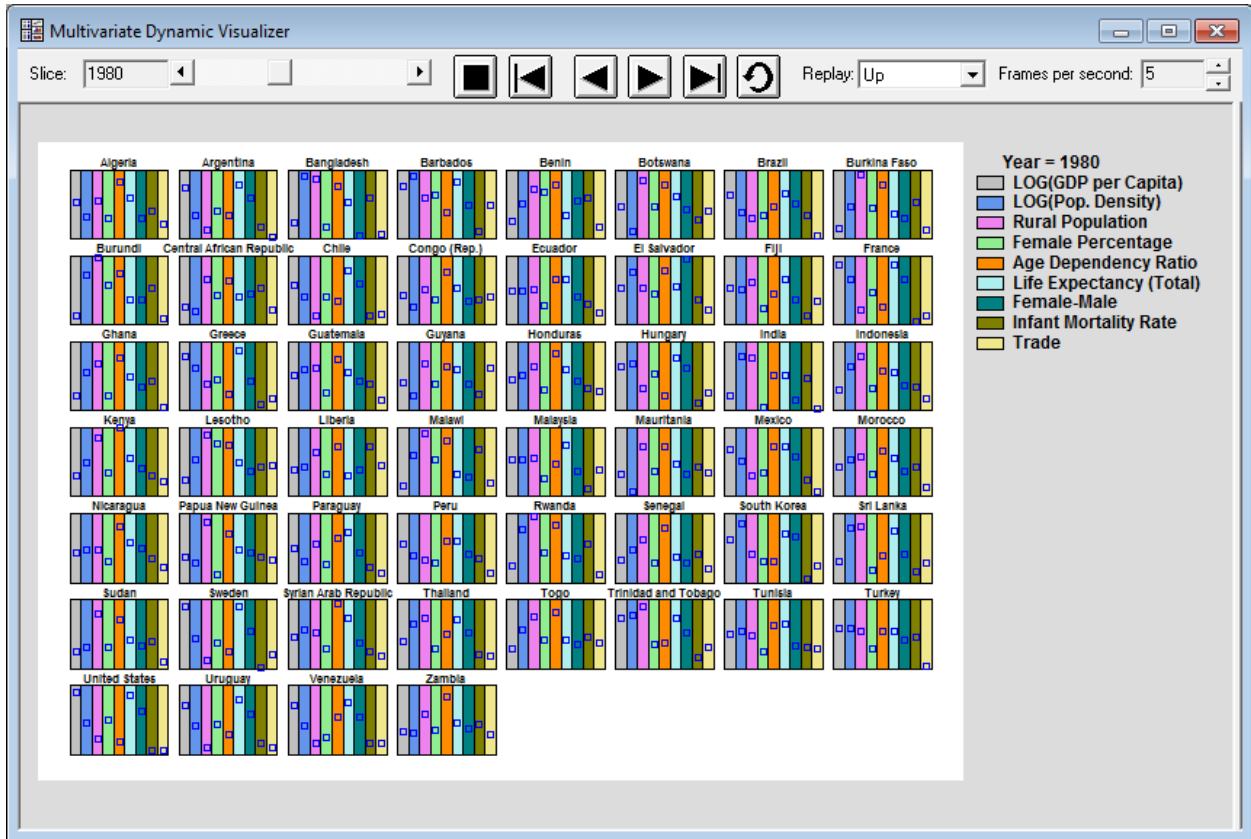
Piecharts:



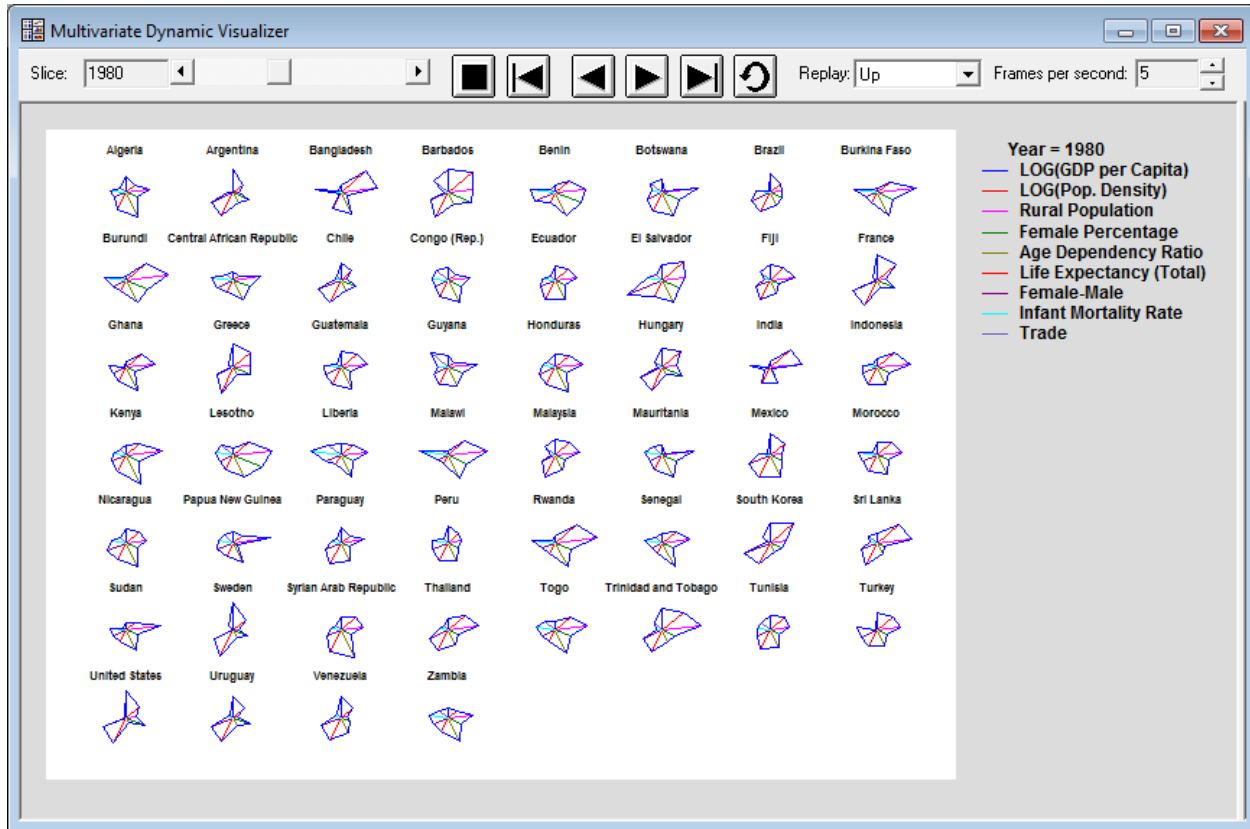
Profile plots:



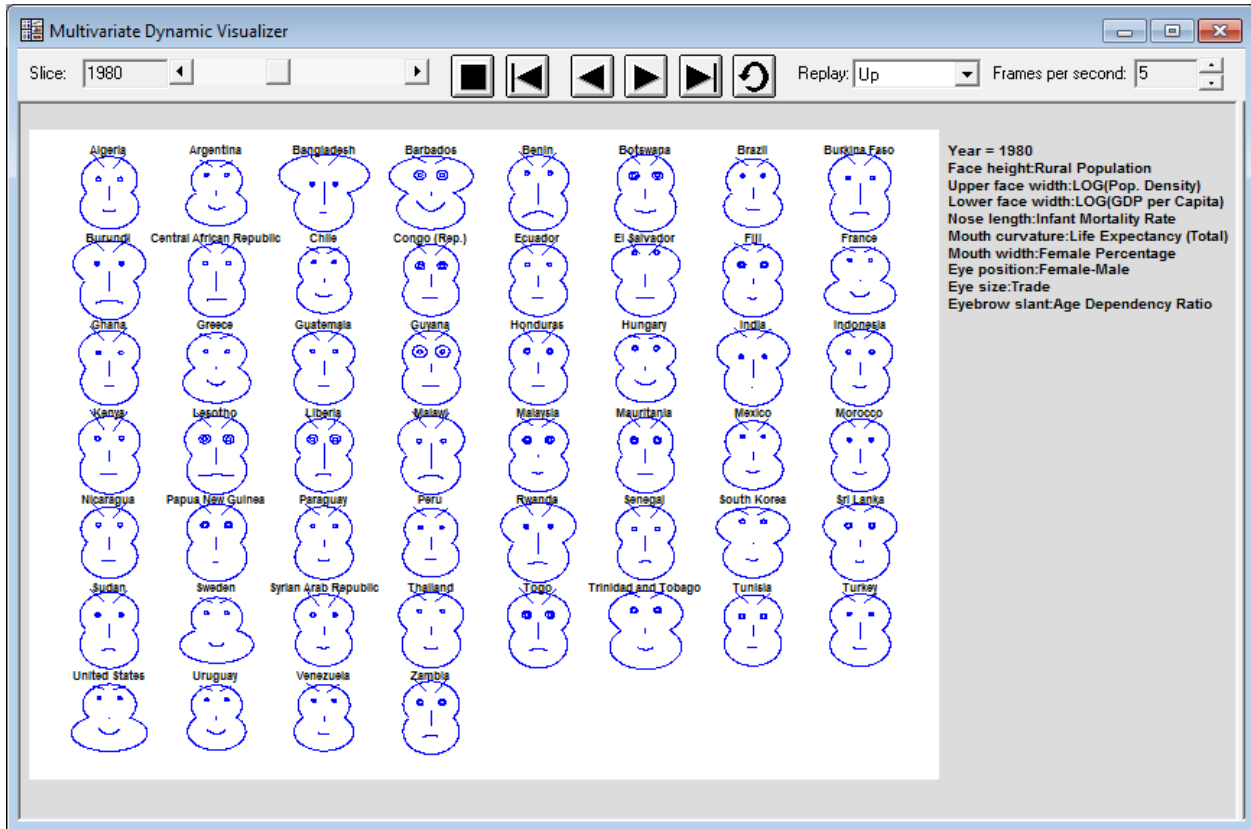
Strip plots:



Star plots:

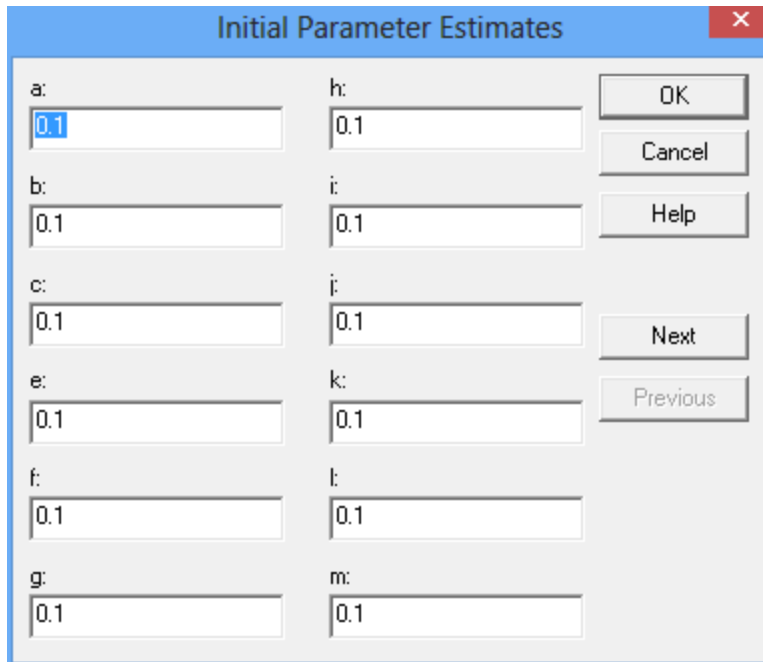


Chernoff faces:



Nonlinear Regression

The maximum number of unknown parameters in the model to be estimated has been increased from 12 to 36. Additional buttons labeled *Next* and *Previous* have been added to the dialog box for specifying the initial parameter estimates:



The image shows a dialog box titled "Initial Parameter Estimates" with a close button (X) in the top right corner. The dialog box contains 12 input fields for parameters a through m, arranged in two columns. Each field contains the value "0.1". The first field (a) is highlighted with a blue selection box. To the right of the input fields are five buttons: "OK", "Cancel", "Help", "Next", and "Previous".

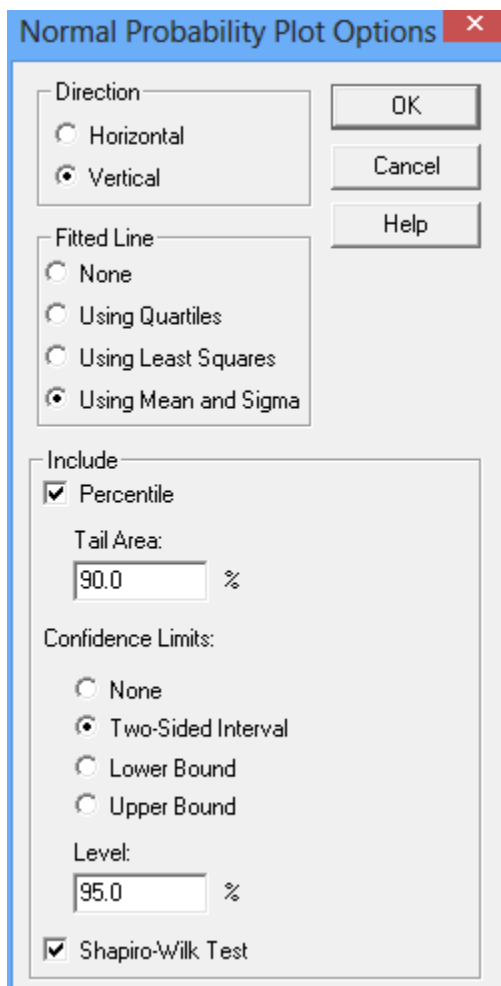
Parameter	Value
a:	0.1
b:	0.1
c:	0.1
e:	0.1
f:	0.1
g:	0.1
h:	0.1
i:	0.1
j:	0.1
k:	0.1
l:	0.1
m:	0.1

Buttons: OK, Cancel, Help, Next, Previous

Normal Probability Plot

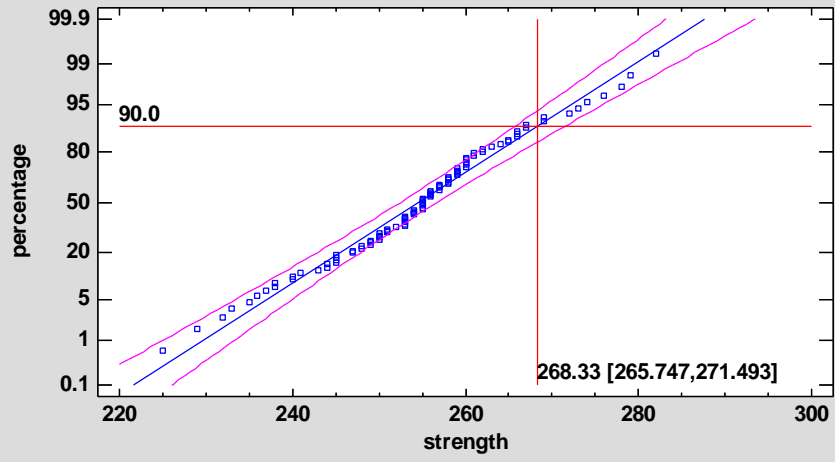
Several major enhancements have been made to the Normal Probability Plot procedure:

1. A third method of fitting a line based on the sample mean and sigma has been added.
2. Crosshairs may be drawn at a selected percentile based on the fitted line.
3. If the line is estimated using the sample mean and sigma, confidence limits may be drawn around the line.
4. The sample mean, standard deviation, and results of a Shapiro-Wilk test for normality may be drawn in the margins of the plot.



A plot containing all of the new options is shown below:

Normal Probability Plot
with 95% limits

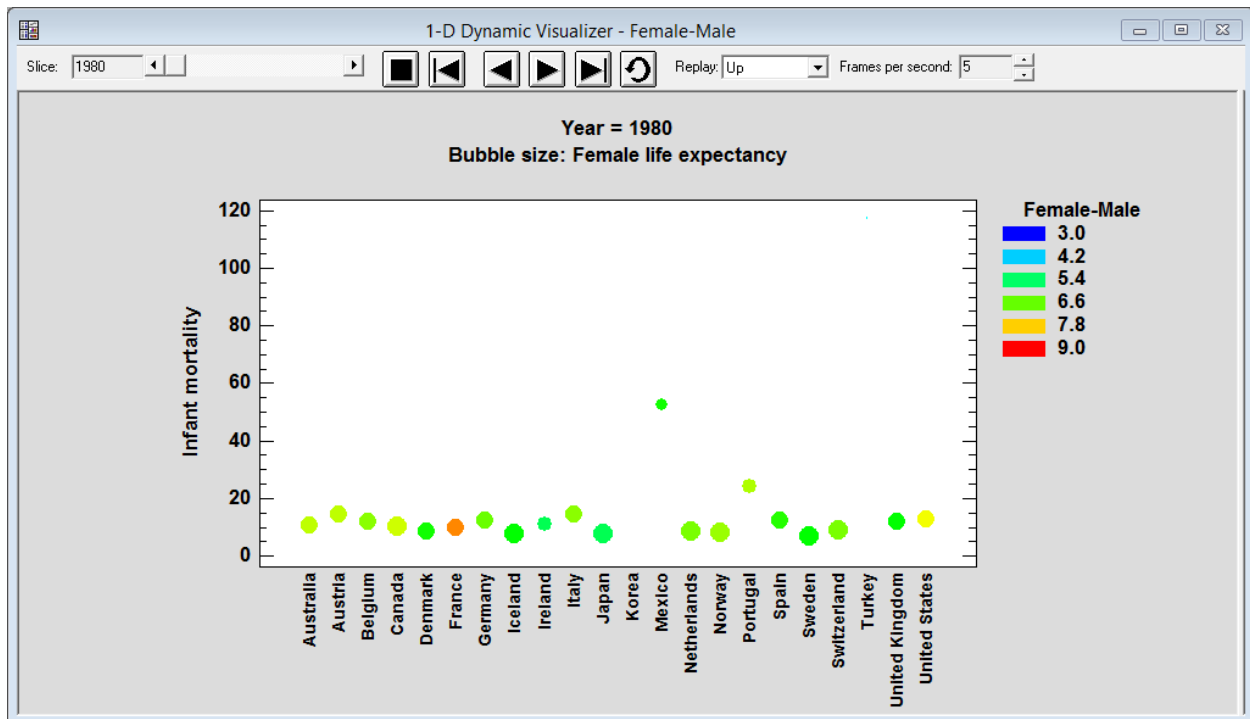


One-Dimensional Visualizer

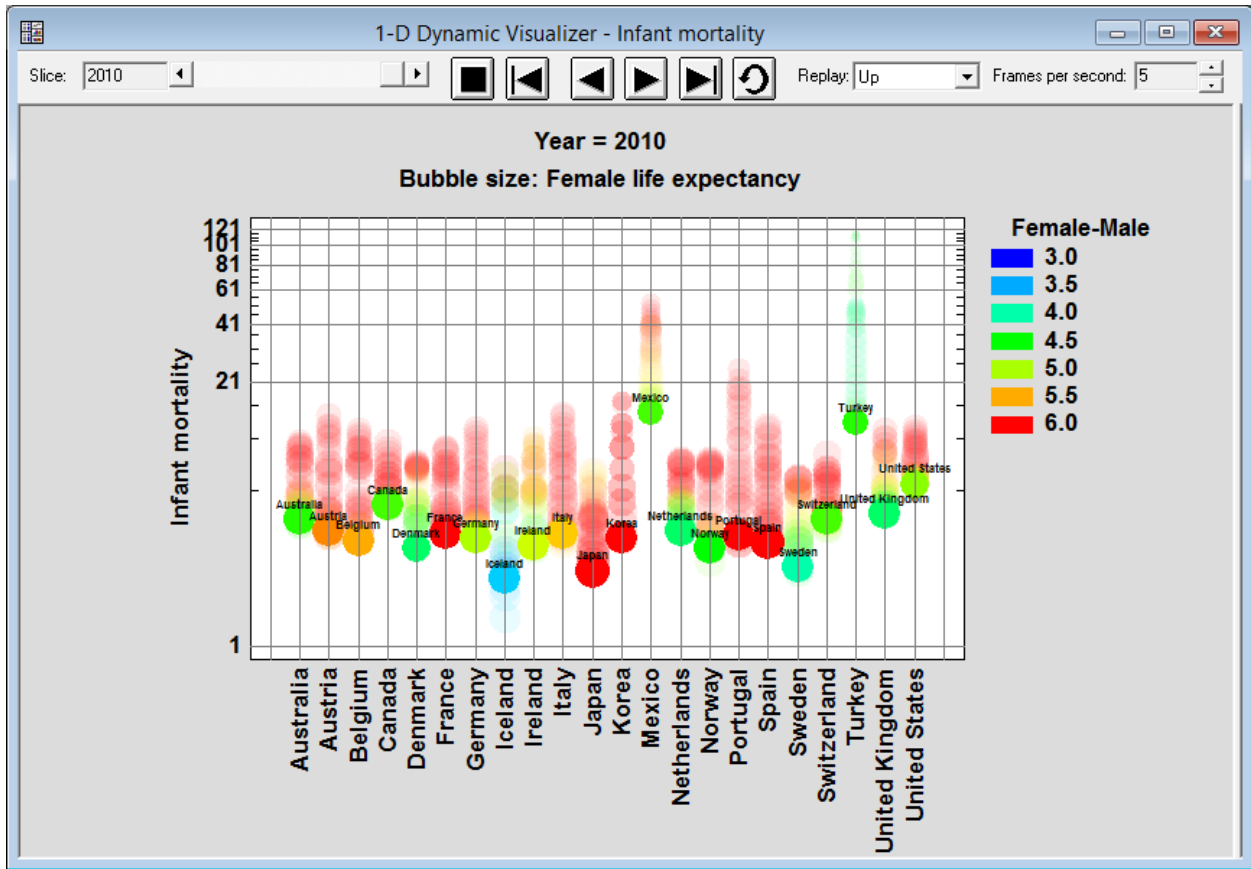
This new Statlet is designed to plot multiple time series in a manner that helps users visualize the changes in multiple variables over time. Given n time series observed over p time periods, the program generates a dynamic display that illustrates how each of the variables has changed over time. Typical applications include plotting:

1. Yearly demographic variables for different countries.
2. Quarterly sales figures for multiple divisions within a company.
3. Monthly economic indices on a state-by-state basis.
4. Daily closing stock prices for multiple equities within a portfolio.

The basic plot shows bubbles plotted on an X-Y display. The position along the Y axis represents the value of the primary data variable. The size and color of the bubbles may be used to illustrate other variables. As time increases, the analyst can follow changes in all of the variables. Various options are offered for smoothing the data and for dealing with missing values.

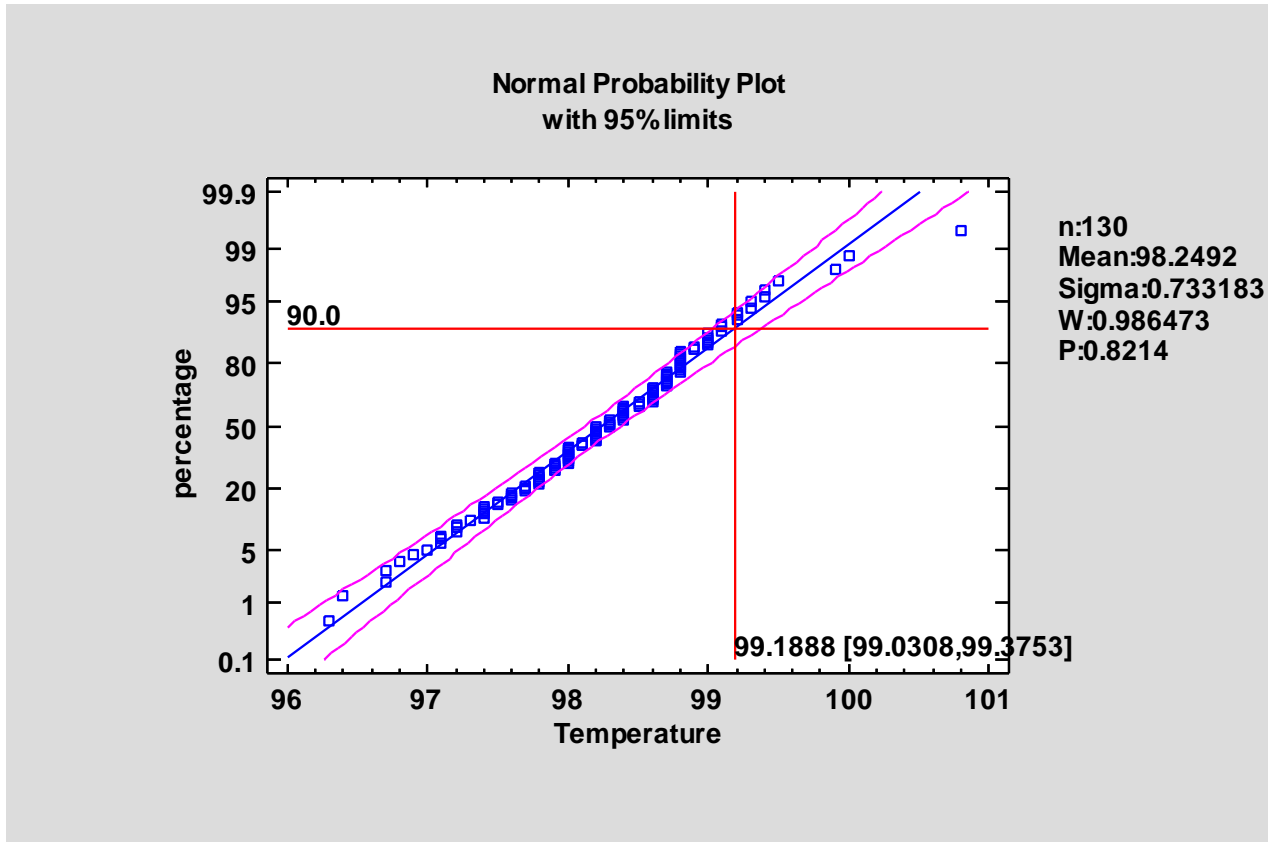


Breadcrumbs may be added to display past values:



One Variable Analysis

The new options for the *Normal Probability Plot* have also been added to this procedure:



Oneway ANOVA

A new method has been added for simultaneously comparing several means. Called the *Games-Howell* procedure, it is similar to the Tukey procedure, except that it does not assume equal variances within each group. This is a good choice if the *Variance Check* indicates significant differences amongst the group variances.

Multiple Range Tests for strength by material			
Method: 95.0 percent Games-Howell			
<i>material</i>	<i>Count</i>	<i>Mean</i>	<i>Homogeneous Groups</i>
D	8	20.5	X
C	8	22.625	X
B	8	31.875	XX
A	8	43.125	X

<i>Contrast</i>	<i>Sig.</i>	<i>Difference</i>	<i>+/- Limits</i>
A - B		11.25	12.8992
A - C	*	20.5	15.5273
A - D	*	22.625	13.1754
B - C		9.25	15.5023
B - D		11.375	13.1407
C - D		2.125	15.6958

* denotes a statistically significant difference.

The output of the Kruskal-Wallis test has also been expanded to include Bonferroni intervals for the difference between each pair of medians:

Kruskal-Wallis Test for strength by material		
<i>material</i>	<i>Sample Size</i>	<i>Average Rank</i>
A	8	26.125
B	8	17.6875
C	8	12.0625
D	8	10.125

Test statistic = 14.0787 P-Value = 0.00279989

95.0 percent Bonferroni intervals

<i>Contrast</i>	<i>Sig.</i>	<i>Difference</i>	<i>+/- Limits</i>
A - B		8.4375	12.3745
A - C	*	14.0625	12.3745
A - D	*	16.0	12.3745
B - C		5.625	12.3745
B - D		7.5625	12.3745
C - D		1.9375	12.3745

* denotes a statistically significant difference.

Bonferroni intervals permit the simultaneous comparison of many differences at a specified family confidence level.

Oneway Repeated Measures ANOVA

A new procedure has been added to simplify the analysis of oneway repeated measures designs. The need for a special procedure stems from potential dependence between observations made on a given subject. Included in the procedure is Mauchly's test for *sphericity*, which is a condition requiring that the variance between any two estimated treatment means be the same. Corrections to the standard F-test are included for cases when sphericity is not present. Alternative MANOVA tests are also included.

Sphericity Tests and Adjustments						
Mauchly's Sphericity Test						
<i>W</i>	<i>Chi-square</i>	<i>D.f.</i>	<i>P-value</i>			
0.712325	1.9411	5.0	0.857233			
Epsilon						
<i>Huynh-Feldt</i>	<i>Greenhouse-Geisser</i>	<i>Lower-bound</i>				
1.0	0.804865	0.333333				
Tests of Within-Patient Effects						
<i>Source</i>	<i>Sphericity Correction</i>	<i>Sum of Squares</i>	<i>Df</i>	<i>Mean Square</i>	<i>F-Ratio</i>	<i>P-Value</i>
Time	None	667.594	3.0	222.531	39.05	0.0000
	Huynh-Feldt	667.594	3.0	222.531	39.05	0.0000
	Greenhouse-Geisser	667.594	2.41459	276.483	39.05	0.0000
	Lower-bound	667.594	1.0	667.594	39.05	0.0004
Error(Time)	None	119.656	21.0	5.69792		
	Huynh-Feldt	119.656	21.0	5.69792		
	Greenhouse-Geisser	119.656	16.9022	7.07935		
	Lower-bound	119.656	7.0	17.0937		
Multivariate Tests						
<i>Test</i>	<i>Value</i>	<i>F</i>	<i>Hypothesis D.F.</i>	<i>Error D.F.</i>	<i>P-Value</i>	
Wilks' lambda	0.0387258	41.371	3.0	5.0	0.0006	
Pillai trace	0.961274	41.371	3.0	5.0	0.0006	
Hotelling-Lawley trace	24.8226	41.371	3.0	5.0	0.0006	
Roy's largest root	24.8226	41.371	3.0	5.0	0.0006	

Open-High-Low-Close Candlestick Plot Statlet

This new Statlet is designed to plot security prices in a manner often used by stock traders. It shows the opening price for each trading session, high and low prices during the session, and the closing price using a graphical image often referred to as a candlestick. Trading volumes may also be displayed as bars along the bottom of the plot.

In addition to the raw data, a smoother may be added to the plot using either a simple moving average or an exponentially weighted moving average. Trading bands may be plotted around the smoothed line at either a fixed percentage or using the method developed by Bollinger. The %b statistic is calculated to measure where the closing price lies with respect to the bands. A bandwidth is also calculated to measure volatility.



Interesting features of the Statlet include:

1. The bars for each session extend from the opening price to the closing price and are colored red if the price fell and green if the price rose. The lines extend to the low and high prices observed during the session.
2. Data for the session corresponding to the location of the vertical cursor are displayed in the right margin of the plot. The slider bar may be used to dynamically change the time period at which the cursor is positioned.
3. Users may dynamically change the smoothing parameters and the sigma multiple to visualize the effect of those values on the smooth.
4. Alerts may be generated for various types of events:

StatFolio Alerts

General | Control Charts | Capability Analyses | Open-High-Low-Close

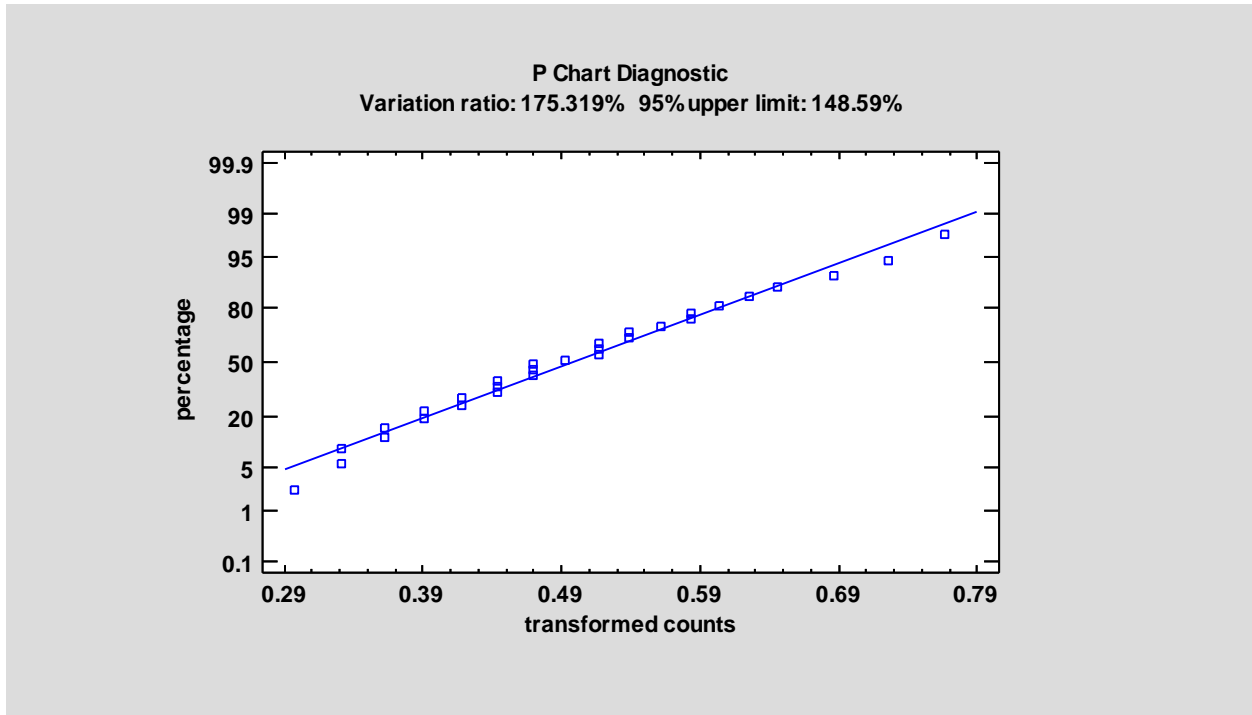
<input type="checkbox"/> Low %b: 0 or less	<input type="checkbox"/> Low BW: 0.1 or less
<input type="checkbox"/> High %b: 1.0 or greater	<input type="checkbox"/> High BW: 0.3 or greater
<input type="checkbox"/> Increase in close: 5.0 or greater	<input type="checkbox"/> Decrease in close: -5.0 or greater
<input type="checkbox"/> Increase in close: 5.0 % or greater	<input type="checkbox"/> Decrease in close: -5.0 % or greater
<input type="checkbox"/> High minus low: 5.0 or greater	
<input type="checkbox"/> High minus low: 5.0 % or greater	

OK Cancel Show XML Help

When an alert occurs, the actions that are automatically taken include sending an e-mail to a specified address.

P Chart

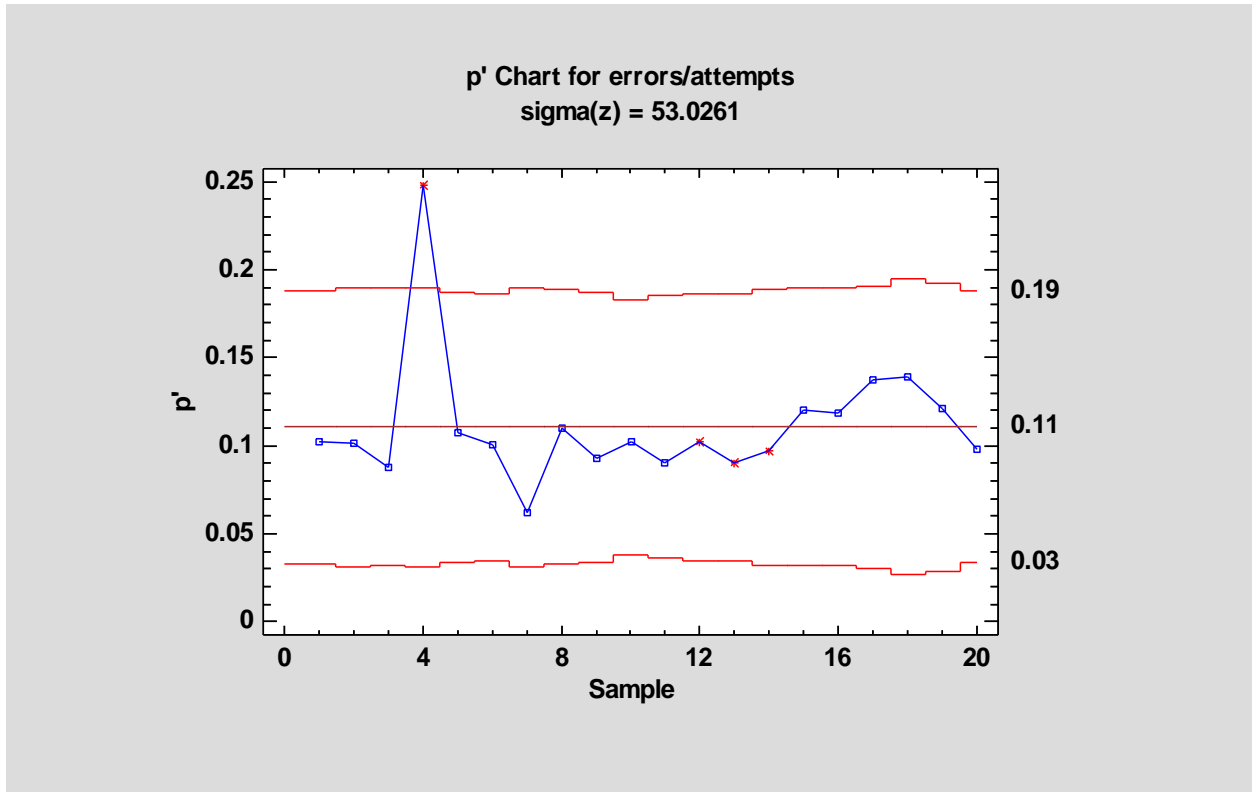
New diagnostic tests have been added to test the assumption that the counts charted follow a binomial distribution. The diagnostic plot displays a variance ratio, which compares the observed variance of the counts to their expected variation. A variance ratio in excess of the displayed upper limit or less than 60% indicates that the Poisson assumption is not valid.



For data that are overdispersed (the variance exceeds the mean), a new Laney P' chart has been added as an alternative to the P chart.

P' Chart

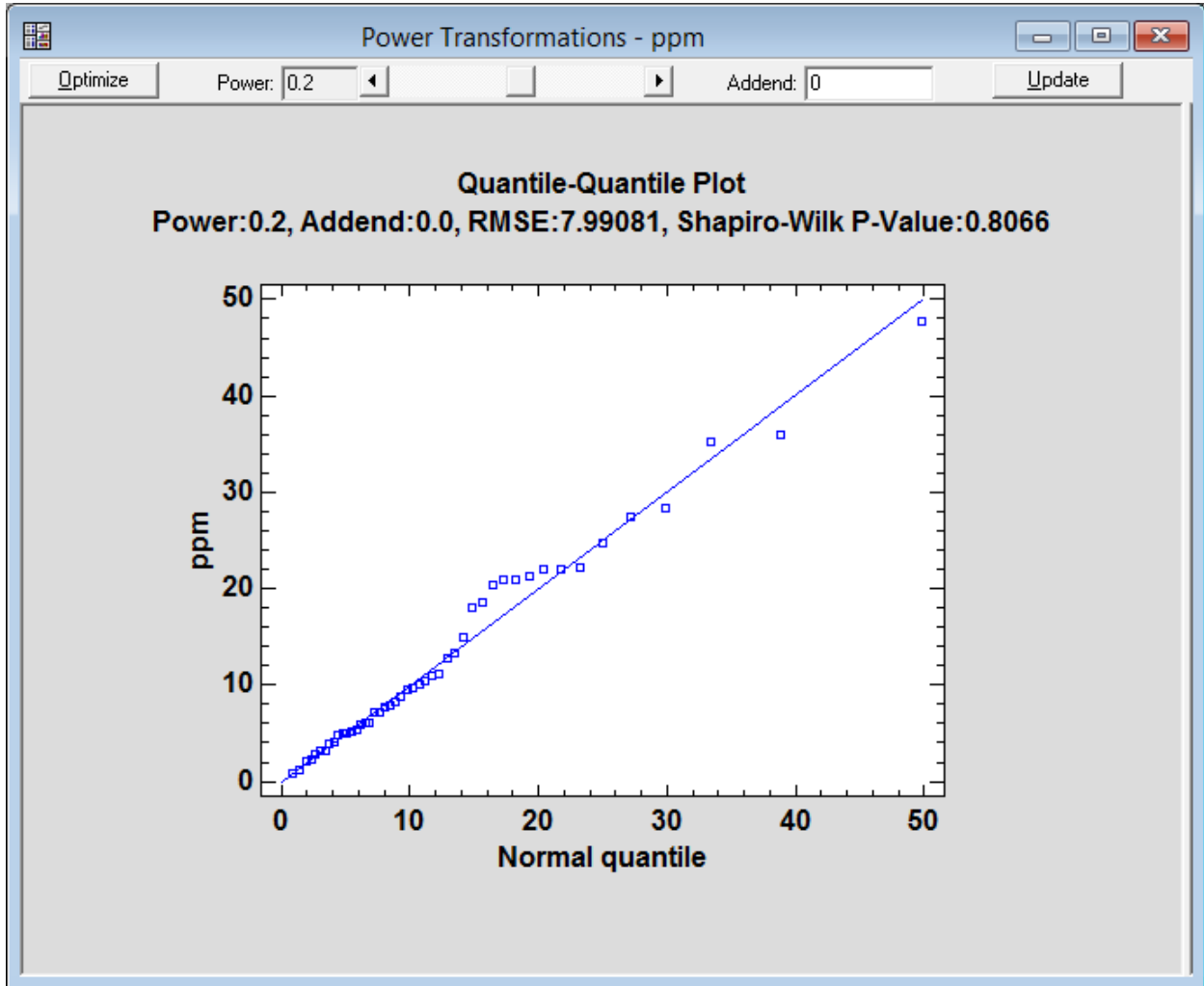
A new control chart has been added to the program which may be used to plot proportions which are overdispersed compared to a binomial distribution. It can prevent a high rate of false alarms, particularly for large sample sizes.



The graph also displays $\hat{\sigma}_z$, which is an estimate of the relative amount of process variation not explained by the binomial distribution alone.

Power Transformations Statlet

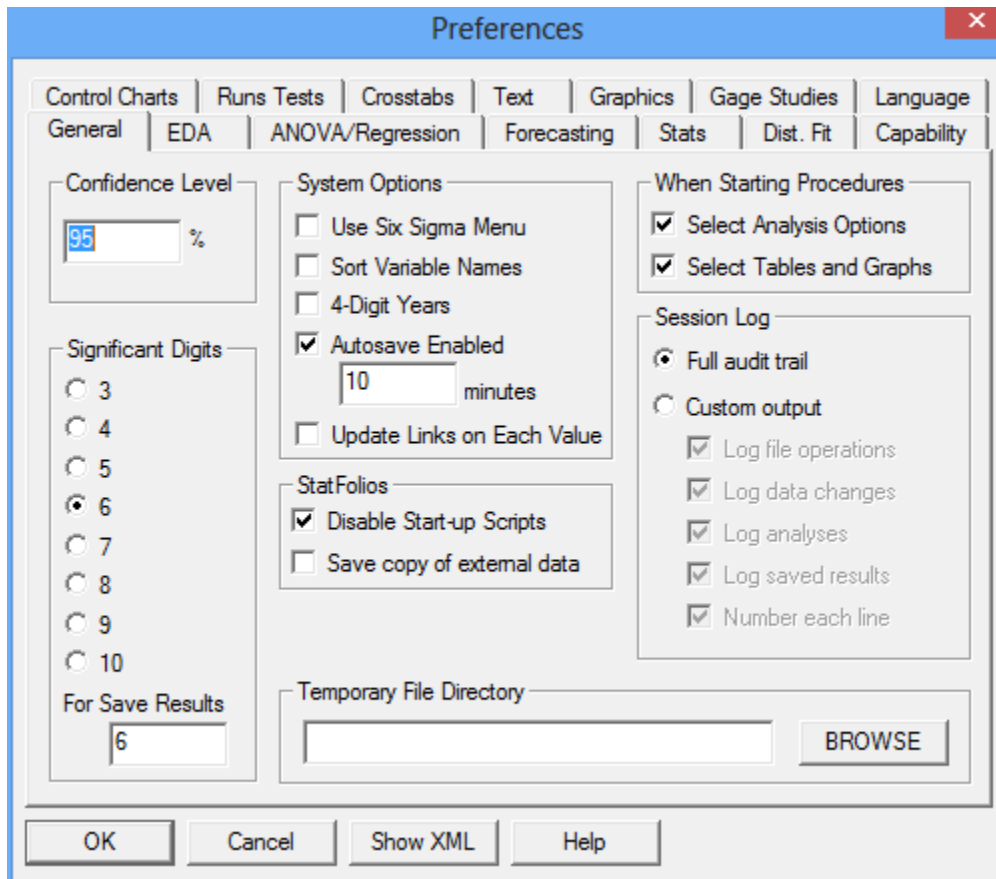
This new Statlet may be used to explore the effect of applying various power transformations to a column of numeric data. It may be used to find the transformation that makes the transformed data most closely characterized by a normal distribution. The controls on the toolbar allow the user to interactively change the power. Alternatively, the Box-Cox approach may be used to find an optimal power.



The output includes the root mean squared error (RMSE) and the results of a Shapiro-Wilk test for normality, performed on the transformed data. As the user moves the slider on the Statlet toolbar, the values of these statistics are updated dynamically.

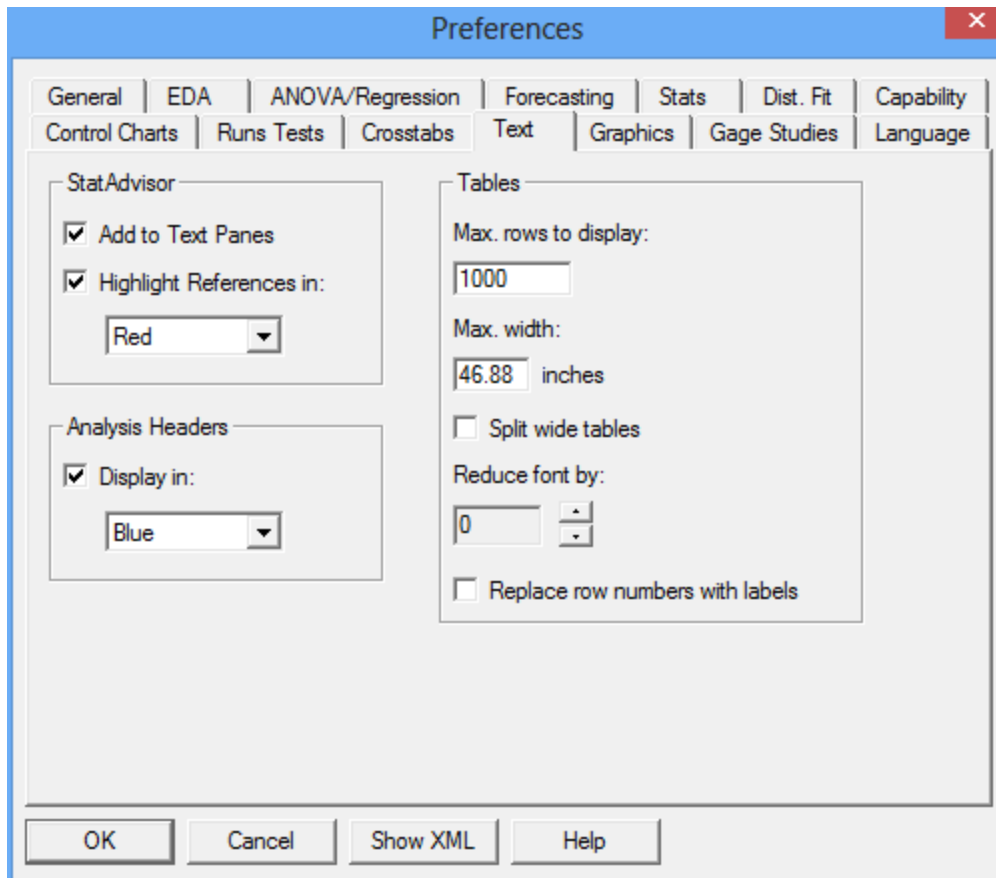
Preferences

The *Preferences* dialog box available on the *Edit* menu has been modified to contain new options.



Specific changes include:

1. *General* tab – addition of options to control what is added to the session log and whether StatFolios should save a copy of any external data.
2. *EDA* tab – additional options for creating normal probability plots.
3. *Graphics* tab – new option for the orientation of ternary plots.
4. *Text* tab – new tab with options that control text output, including the option to replace row numbers in tables with row labels:



Note two new options for text not available in the previous version:

1. The ability to reduce the size of text within tables relative to normal text.
2. The ability to replace row numbers in tables with corresponding row labels.

Principal Components Analysis

New tests have been added to the *Principal Components Analysis* procedure to determine whether it is useful to perform a PCA on a set of variables. The factorability tests include the Kaiser-Meyer-Olsen (KMO) measure of sampling adequacy and Bartlett's test of sphericity:

Factorability Tests

Kaiser-Meyer-Olkin Measure of Sampling Adequacy
KMO = 0.920192

Bartlett's Test of Sphericity
Chi-Square = 1299.83
D.F. = 55
P-Value = 0.0

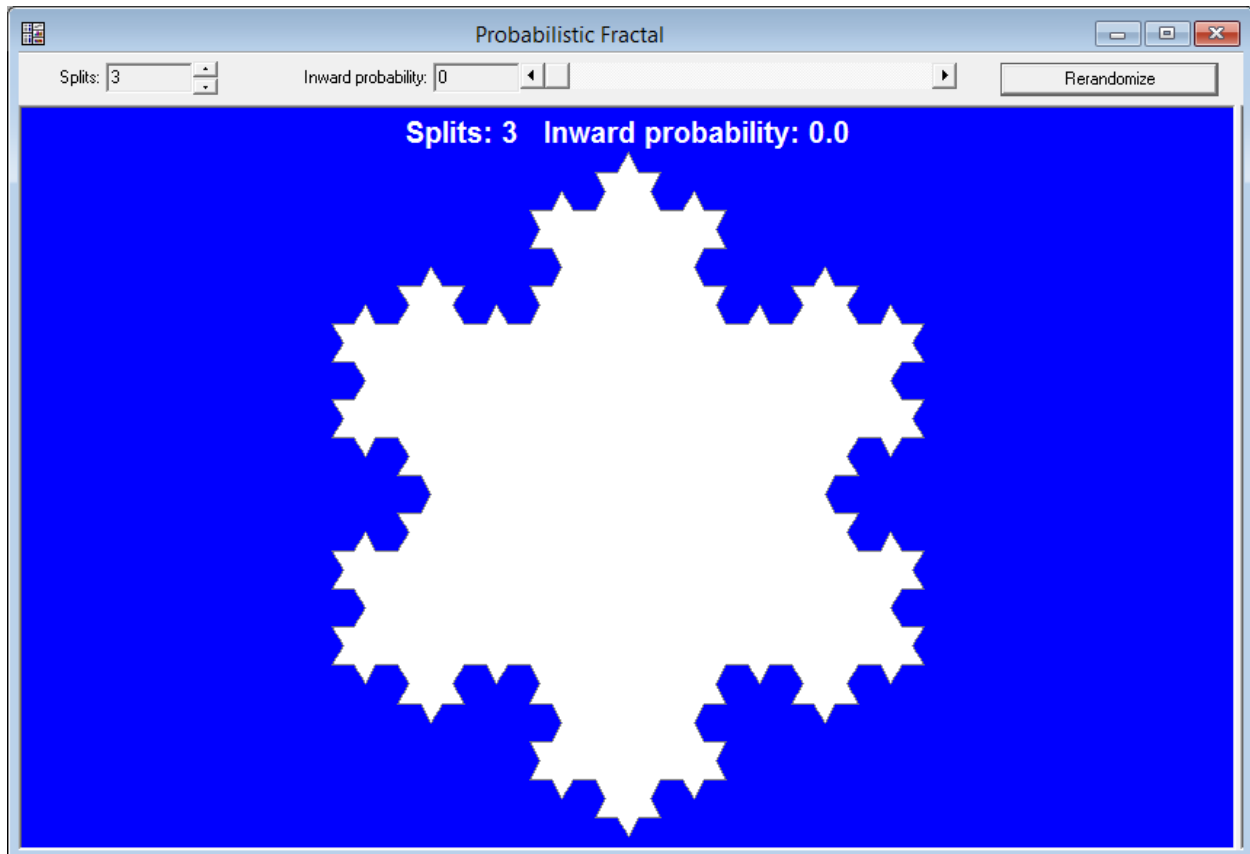
The StatAdvisor

The factorability tests provide indications of whether or not it is likely to be worthwhile attempting to extract factors from a set of variables. The KMO statistic provides an indication of how much common variance is present. For factorization to be worthwhile, KMO should normally be at least 0.6. Since KMO = 0.920192, factorization is likely to provide interesting information about any underlying factors.

Bartlett's test for sphericity tests the hypothesis that the correlation matrix amongst the variables is an identity matrix, indicating that they share no common variance. Since the P-value is < 0.05 , that hypothesis is rejected. Note: Bartlett's test is very sensitive and is usually ignored unless the number of samples per variable is no more than 5. In this case, the number of samples per variable equals 8.45455.

Probabilistic Fractal

This new Statlet is designed to illustrate the concepts of randomness and uncertainty. It is based on the famous Snowflake Fractal. To create the fractal, an equilateral triangle is first drawn. A number of iterations are then performed during which a new equilateral triangle is added to the exterior of each side of the current figure, with each side having one-third the length of the side on which it is placed.



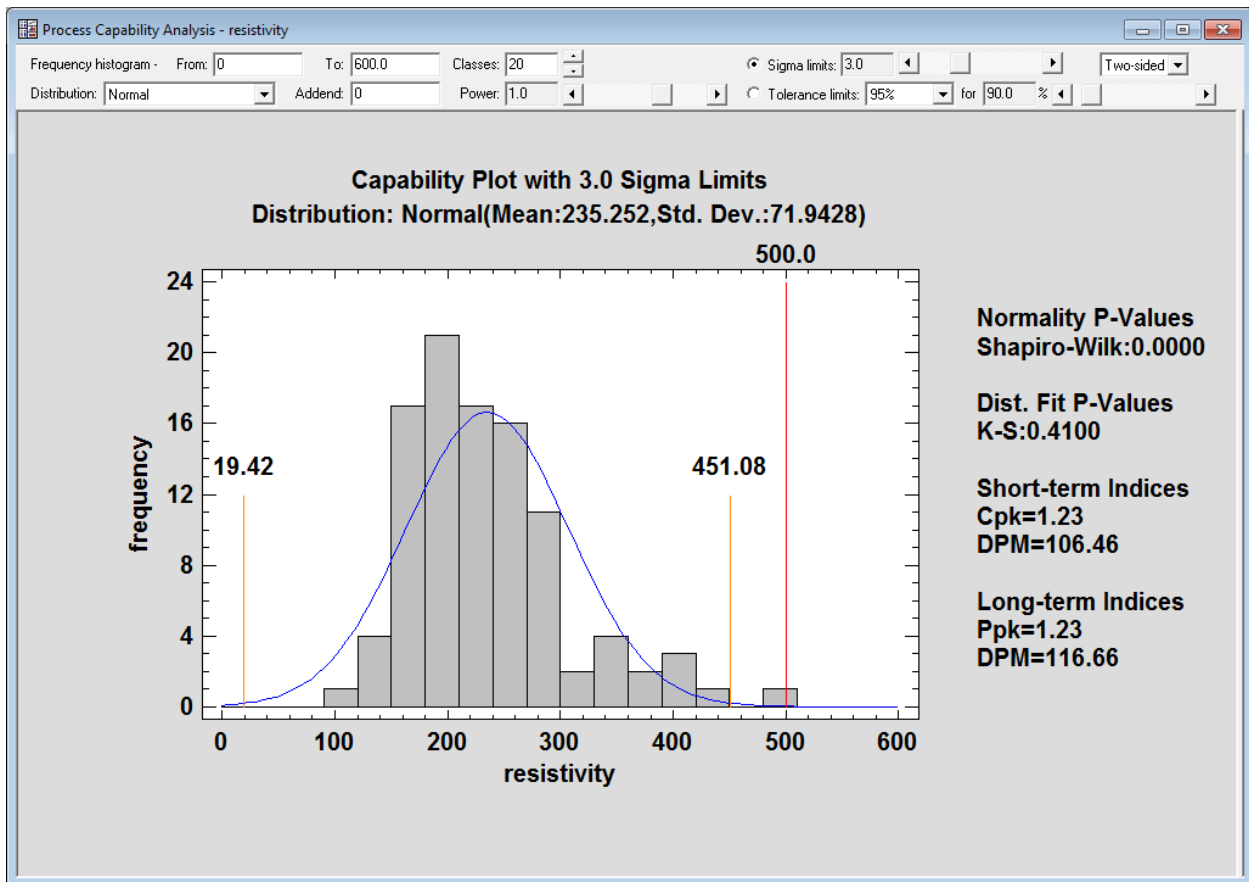
To make the fractal probabilistic, you may set a probability p that, when a new triangle is added to the figure, it is oriented inward rather than outward, thus removing part of the figure rather than adding to it. By pressing the *Rerandomize* button, the fractal may be redrawn using a new set of random numbers. It will be apparent that the amount of "chaos" in the figure is a function of p .

Process Capability Analysis Statlet

This new Statlet performs a capability analysis on measurement data. The data may be collected one at a time (individuals data) or in groups. Any of 21 probability distributions may be selected, although the normal distribution is used by default. The Statlet will calculate:

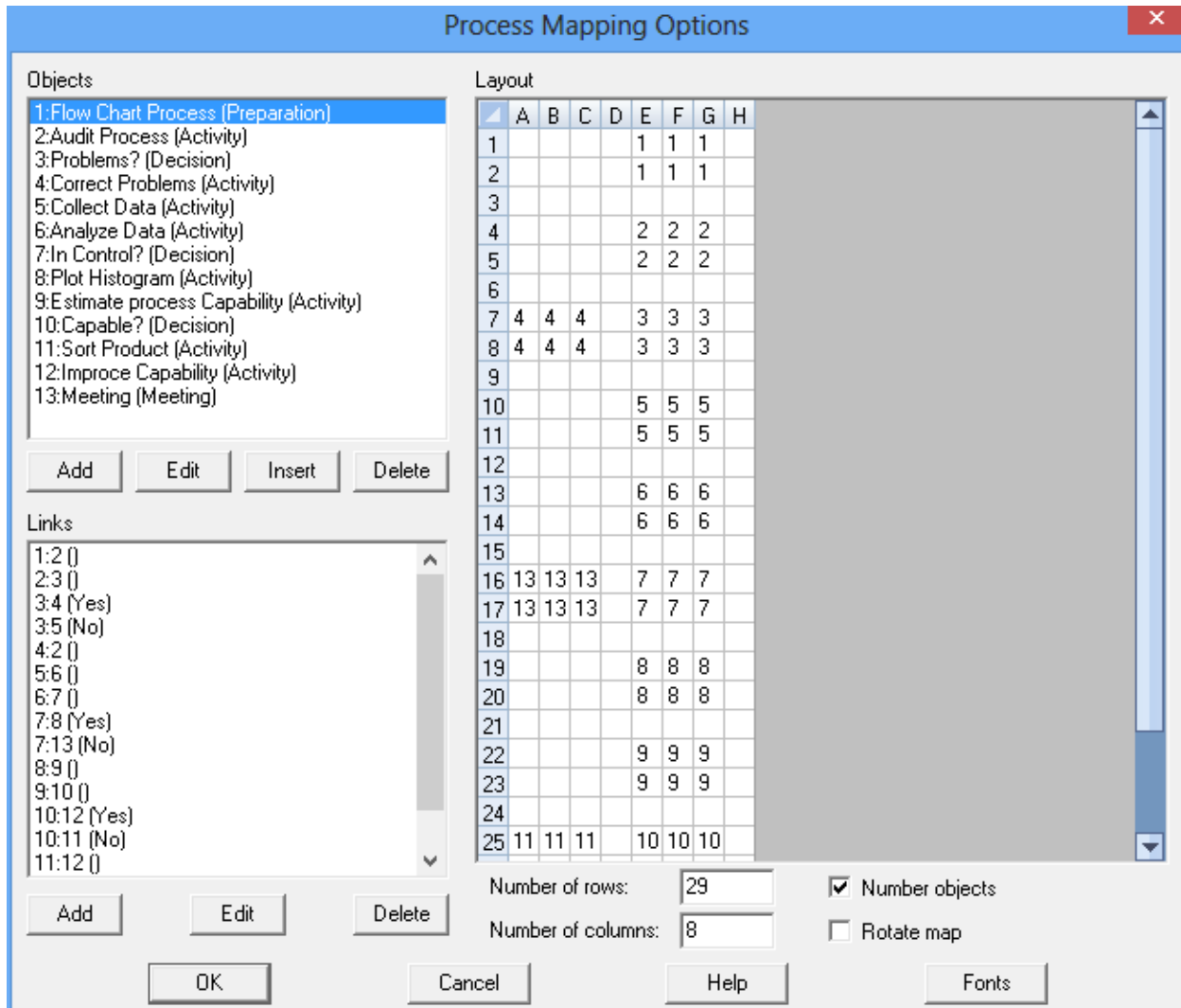
1. Capability indices (both long-term and short-term).
2. DPM (defects per million).
3. Goodness-of-fit tests for the selected distribution.
4. Statistical tolerance limits (for many of the available distributions).

The data may also be transformed using a standard power transformation.



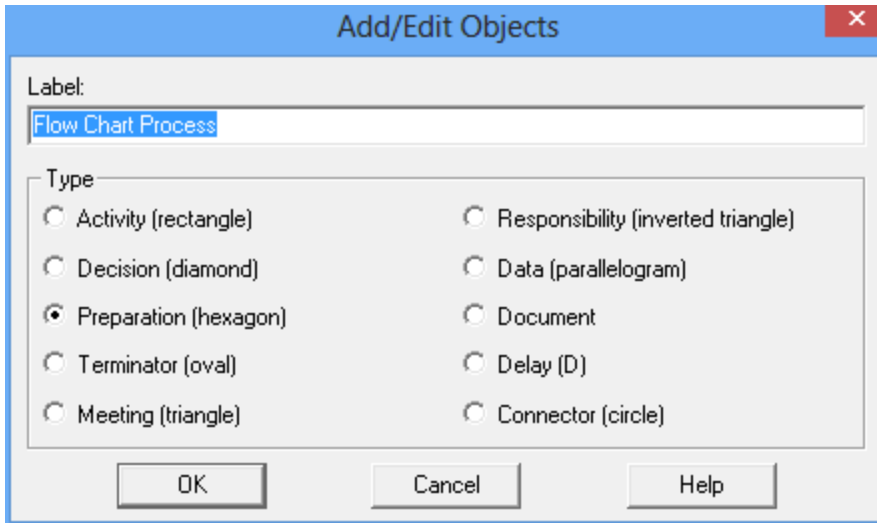
Process Map

The *Process Map* procedure has been significantly reworked. The most evident change is the layout of the *Analysis Options* dialog box:



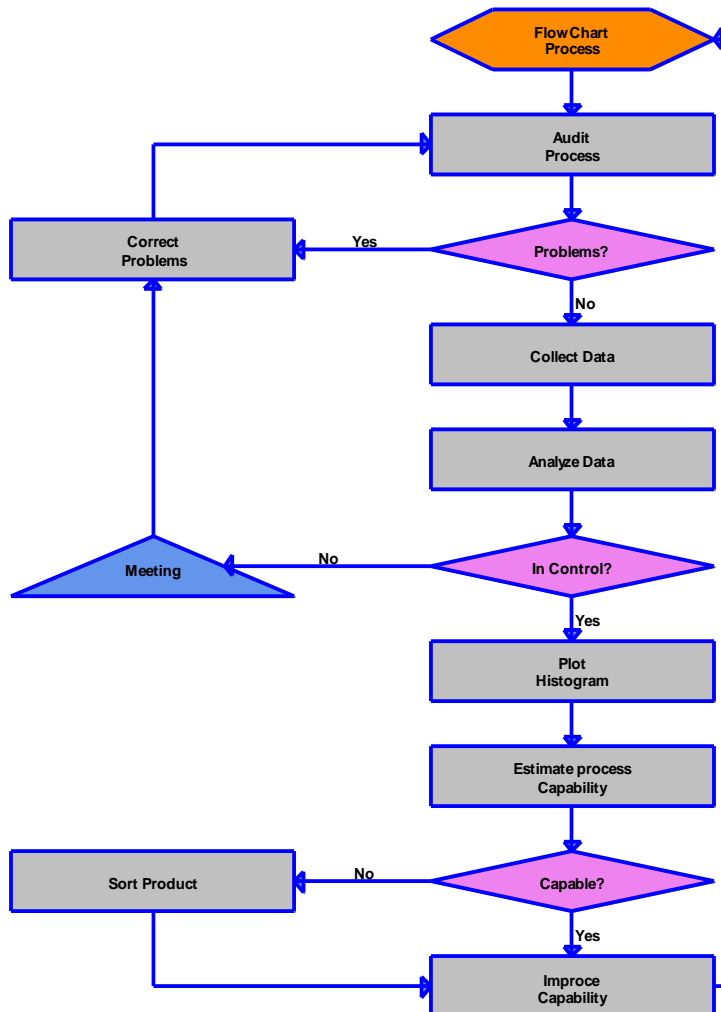
Specifically:

1. A new field has been added displaying all of the links.
2. The definition of the map is stored in the StatFolio rather than in a separate data file.
3. Objects and links are defined through a properties dialog box which contains more types of objects than were available previously:



- Changes made to objects and links are automatically reflected on the map without closing the *Analysis Options* dialog box.

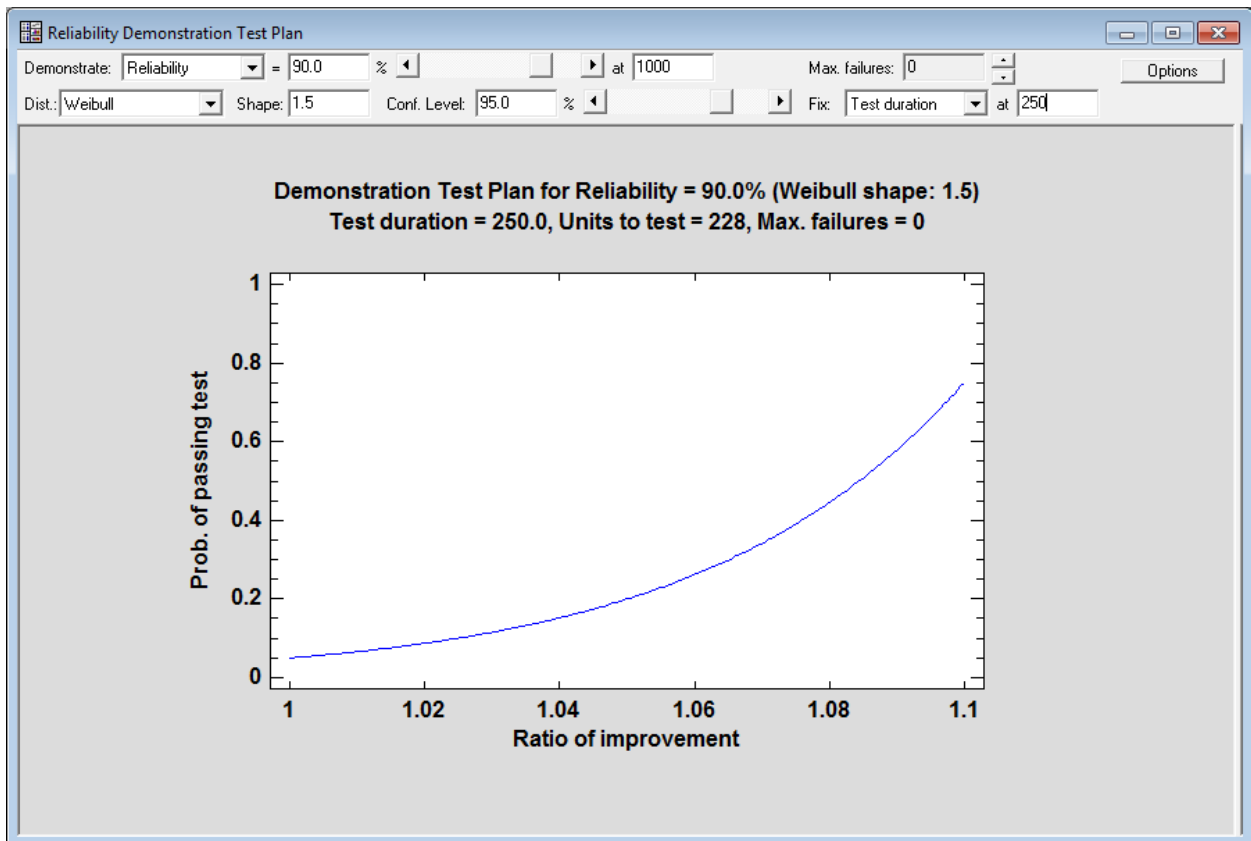
A typical map is shown below:



Reliability Demonstration Test Plan

This new Statlet creates test plans to demonstrate that a failure time distribution satisfies stated conditions. For example, it may be desired to show with 95% confidence that the reliability of a product equals or exceeds 90% at the end of the warranty period. During the demonstration, n units will be tested for a duration equal to t . The demonstration will be considered successful if no more than f units fail during the test.

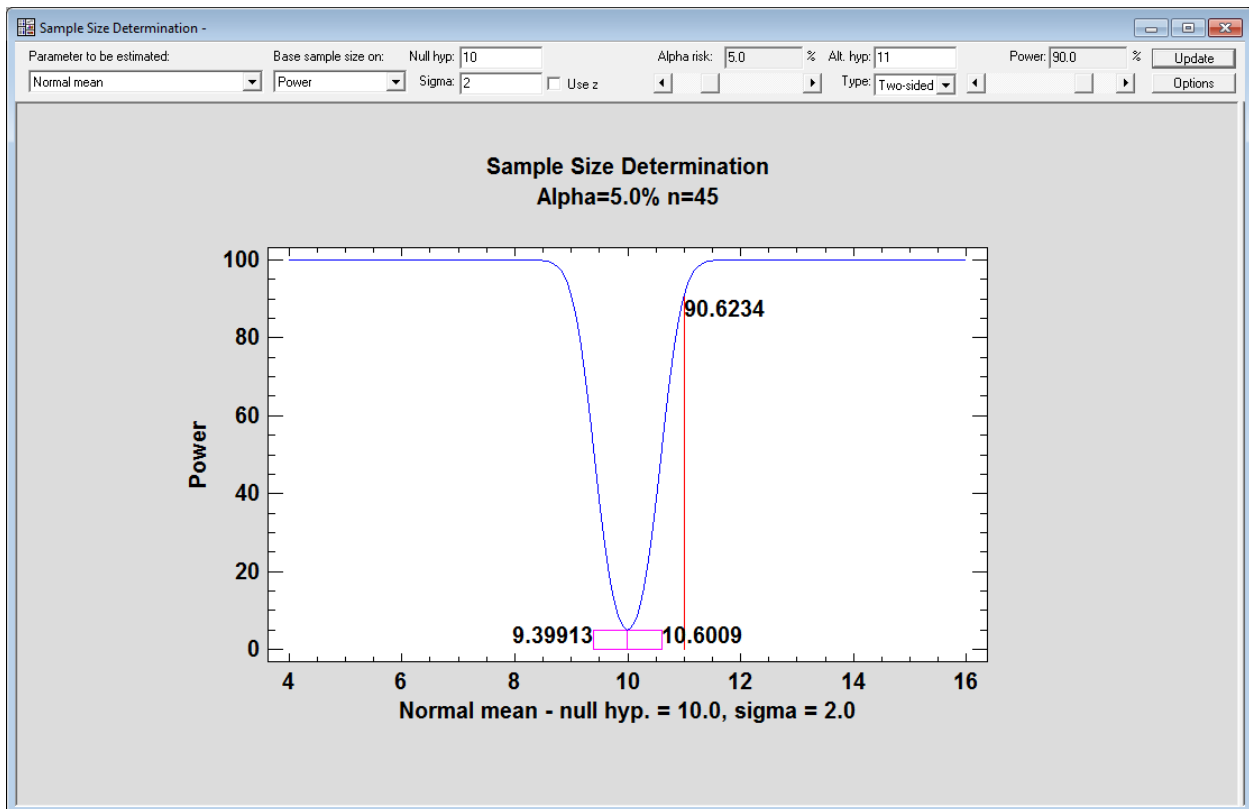
The user specifies either the number of units to be tested or the duration of the test. The procedure solves for the other quantity.



Sample Size Determination Statlet

This new Statlet determines the sample size needed to estimate or test values of various parameters. The size may be based on either the width of a confidence interval or the power of a hypothesis test. Parameters for which sample sizes may be determined are:

1. The mean μ of a normal distribution.
2. The standard deviation σ of a normal distribution.
3. The proportion of successes θ in a binomial distribution.
4. The rate of events λ in a Poisson distribution.
5. The difference between the means of 2 normal distributions.
6. The ratio of the standard deviations of 2 normal distributions.
7. The difference between the parameters of 2 binomial distributions.
8. The difference between the parameters of 2 Poisson distributions.
9. The maximum difference between the means of 3 or more normal distributions.
10. The Pearson correlation coefficient.
11. The capability index C_p .
12. The capability index C_{pk} .
13. The capability index C_{pm} .



The Statlet window displays:

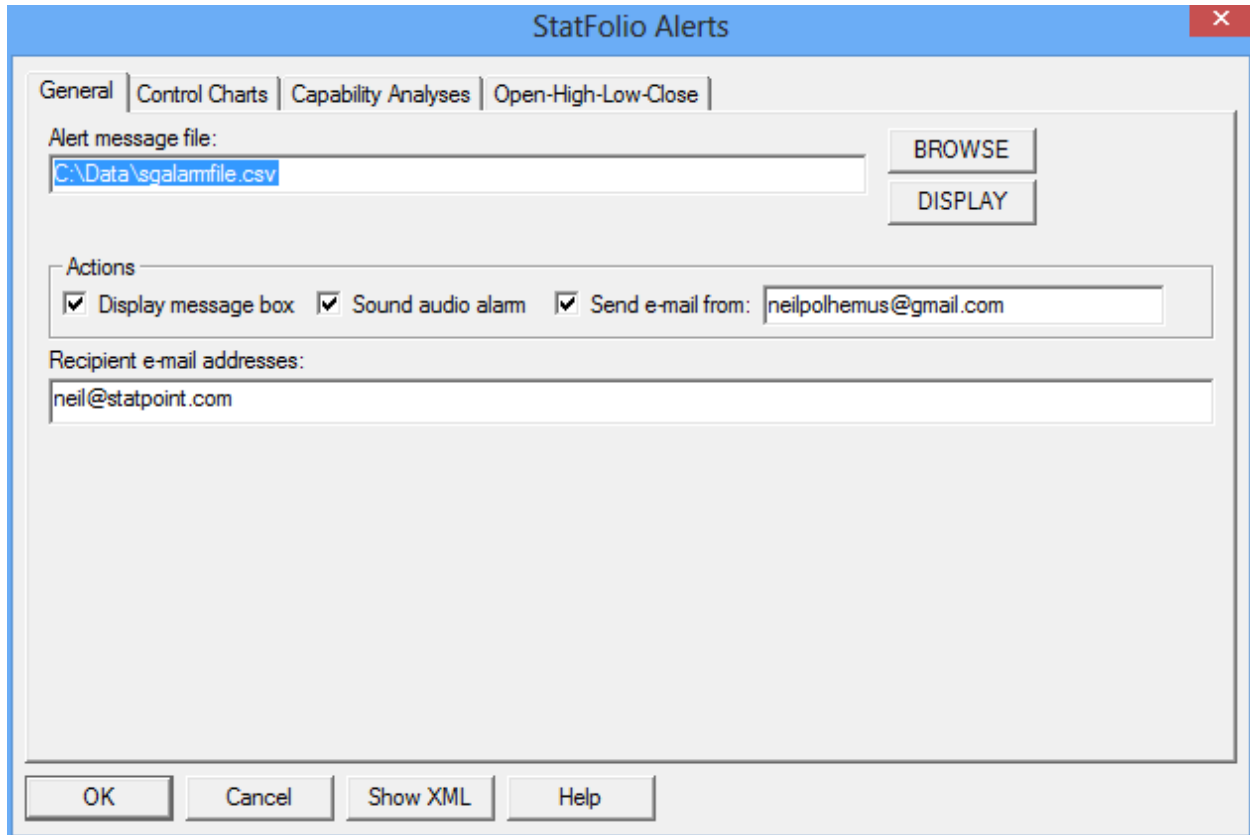
1. The sample size required to perform a test with the specified errors.
2. A power curve.

3. A confidence interval centered at the null hypothesis.
4. The calculated power at the alternative hypothesis.

As users change input values on the toolbar, the plot updates dynamically.

StatFolio Alerts

A new tabbed dialog box has been created to handle StatFolio alerts:



New alerts are provided for large changes in the *Open-High-Low-Close* Statlet:

StatFolio Alerts ✕

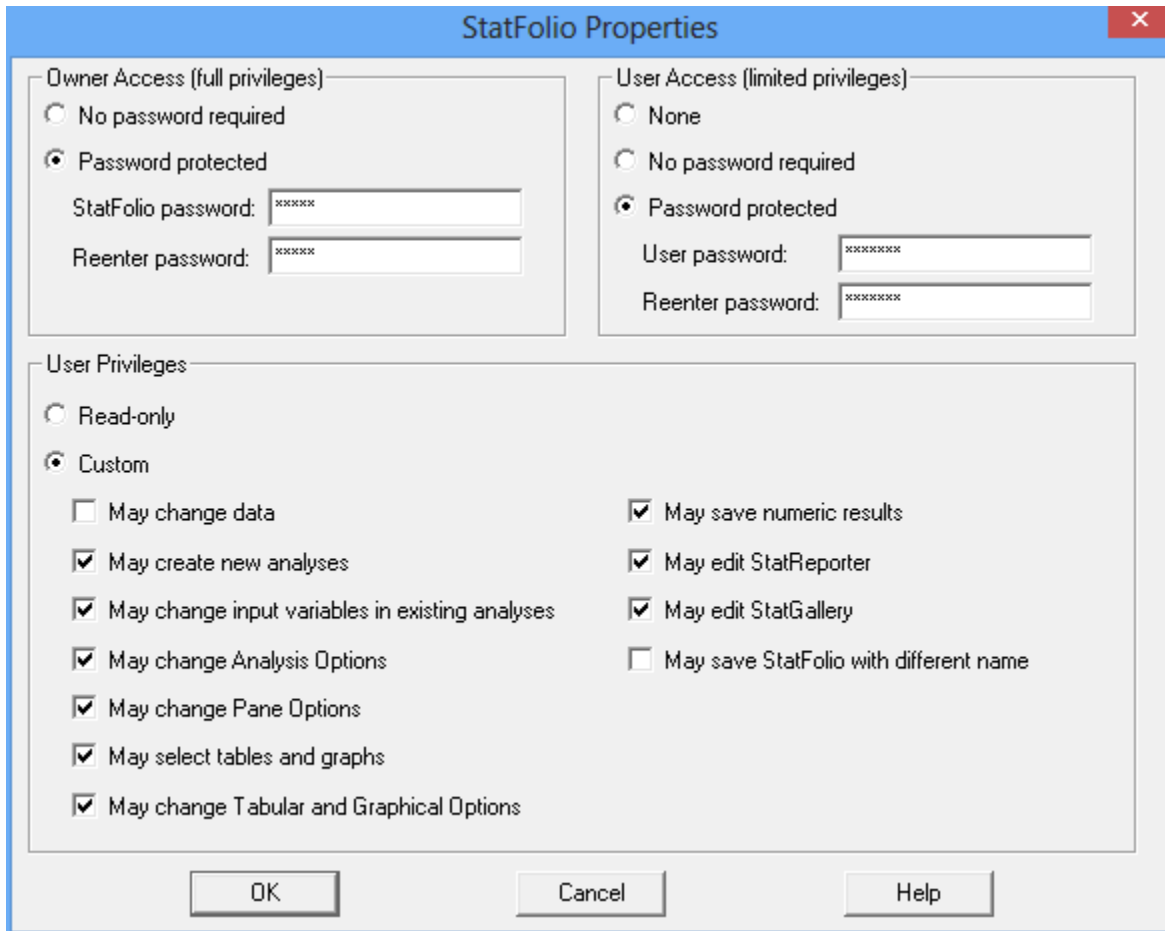
General | Control Charts | Capability Analyses | **Open-High-Low-Close**

<input type="checkbox"/> Low %b: <input type="text" value="0"/> or less	<input type="checkbox"/> Low BW: <input type="text" value="0.1"/> or less
<input type="checkbox"/> High %b: <input type="text" value="1.0"/> or greater	<input type="checkbox"/> High BW: <input type="text" value="0.3"/> or greater
<input checked="" type="checkbox"/> Increase in close: <input type="text" value="5.0"/> or greater	<input type="checkbox"/> Decrease in close: <input type="text" value="-5.0"/> or greater
<input type="checkbox"/> Increase in close: <input type="text" value="5.0"/> % or greater	<input type="checkbox"/> Decrease in close: <input type="text" value="-5.0"/> % or greater
<input type="checkbox"/> High minus low: <input type="text" value="5.0"/> or greater	
<input type="checkbox"/> High minus low: <input type="text" value="5.0"/> % or greater	

As in Version 16, alerts are also generated by the control charts and capability analysis procedures.

StatFolio Passwords

Both owner and user passwords may now be added to a StatFolio. Selecting *Edit – StatFolio Properties* from the main menu displays the following dialog box:

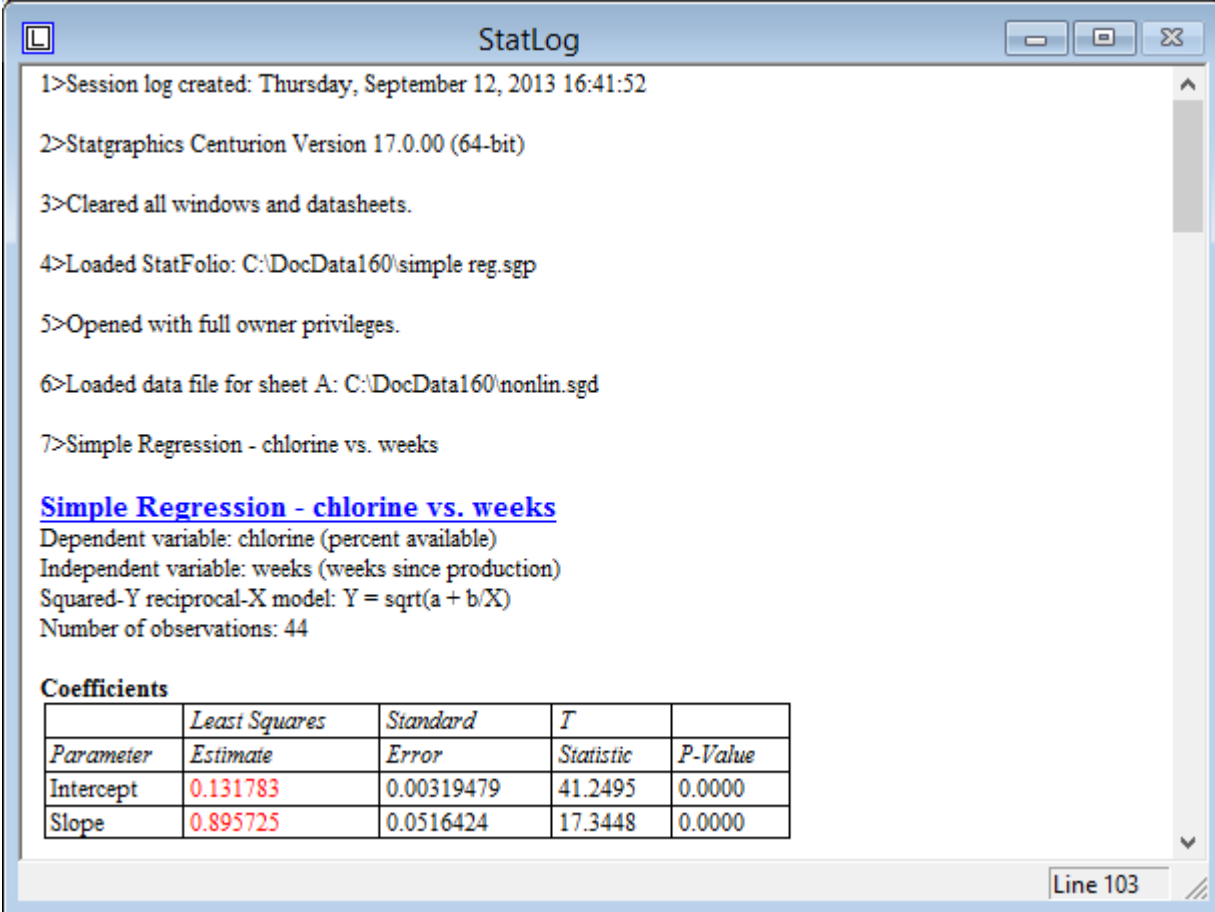


The dialog box, titled "StatFolio Properties", is divided into three main sections. The top-left section, "Owner Access (full privileges)", has two radio buttons: "No password required" (unselected) and "Password protected" (selected). Below are two text boxes for "StatFolio password:" and "Reenter password:", both containing "xxxxxx". The top-right section, "User Access (limited privileges)", has three radio buttons: "None" (unselected), "No password required" (unselected), and "Password protected" (selected). Below are two text boxes for "User password:" and "Reenter password:", both containing "xxxxxxx". The bottom section, "User Privileges", has two radio buttons: "Read-only" (unselected) and "Custom" (selected). Under "Custom", there are ten checkboxes: "May change data" (unselected), "May create new analyses" (checked), "May change input variables in existing analyses" (checked), "May change Analysis Options" (checked), "May change Pane Options" (checked), "May select tables and graphs" (checked), "May change Tabular and Graphical Options" (checked), "May save numeric results" (checked), "May edit StatReporter" (checked), "May edit StatGallery" (checked), and "May save StatFolio with different name" (unselected). At the bottom are three buttons: "OK", "Cancel", and "Help".

Owners have unlimited privileges and may change a StatFolio in any manner. Users are limited to the privileges specified on the dialog box.

StatLog

A new window called the StatLog appears by default whenever Statgraphics is loaded. The StatLog stores information about the current session:



The screenshot shows the StatLog window with the following content:


```
1>Session log created: Thursday, September 12, 2013 16:41:52
2>Statgraphics Centurion Version 17.0.00 (64-bit)
3>Cleared all windows and datasheets.
4>Loaded StatFolio: C:\DocData160\simple reg.sgp
5>Opened with full owner privileges.
6>Loaded data file for sheet A: C:\DocData160\nonlin.sgd
7>Simple Regression - chlorine vs. weeks
```

Simple Regression - chlorine vs. weeks
Dependent variable: chlorine (percent available)
Independent variable: weeks (weeks since production)
Squared-Y reciprocal-X model: $Y = \sqrt{a + b/X}$
Number of observations: 44

Coefficients

	<i>Least Squares</i>	<i>Standard</i>	<i>T</i>	
<i>Parameter</i>	<i>Estimate</i>	<i>Error</i>	<i>Statistic</i>	<i>P-Value</i>
Intercept	0.131783	0.00319479	41.2495	0.0000
Slope	0.895725	0.0516424	17.3448	0.0000

Line 103

Certain information is always included, such as the opening and closing of files and the creation of analysis windows. Other information, such as the contents of statistical tables and graphs, are only included if specified on the *Preferences* dialog box. At any time, the *Log* button  on the analysis toolbar may be used to copy the current analysis to the StatLog.

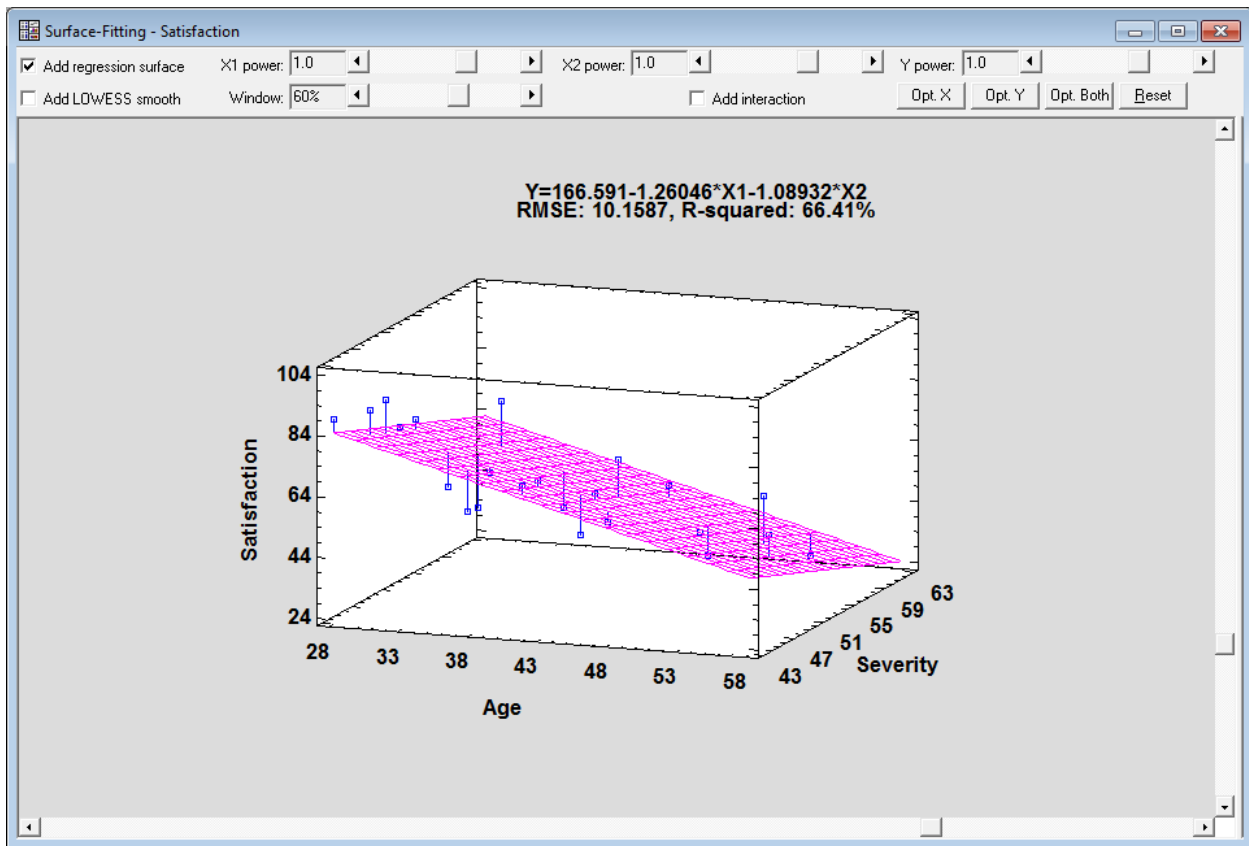
The contents of the StatLog may be saved in an RTF file at any time. These files may be opened in Microsoft Word.

Surface-Fitting Statlet

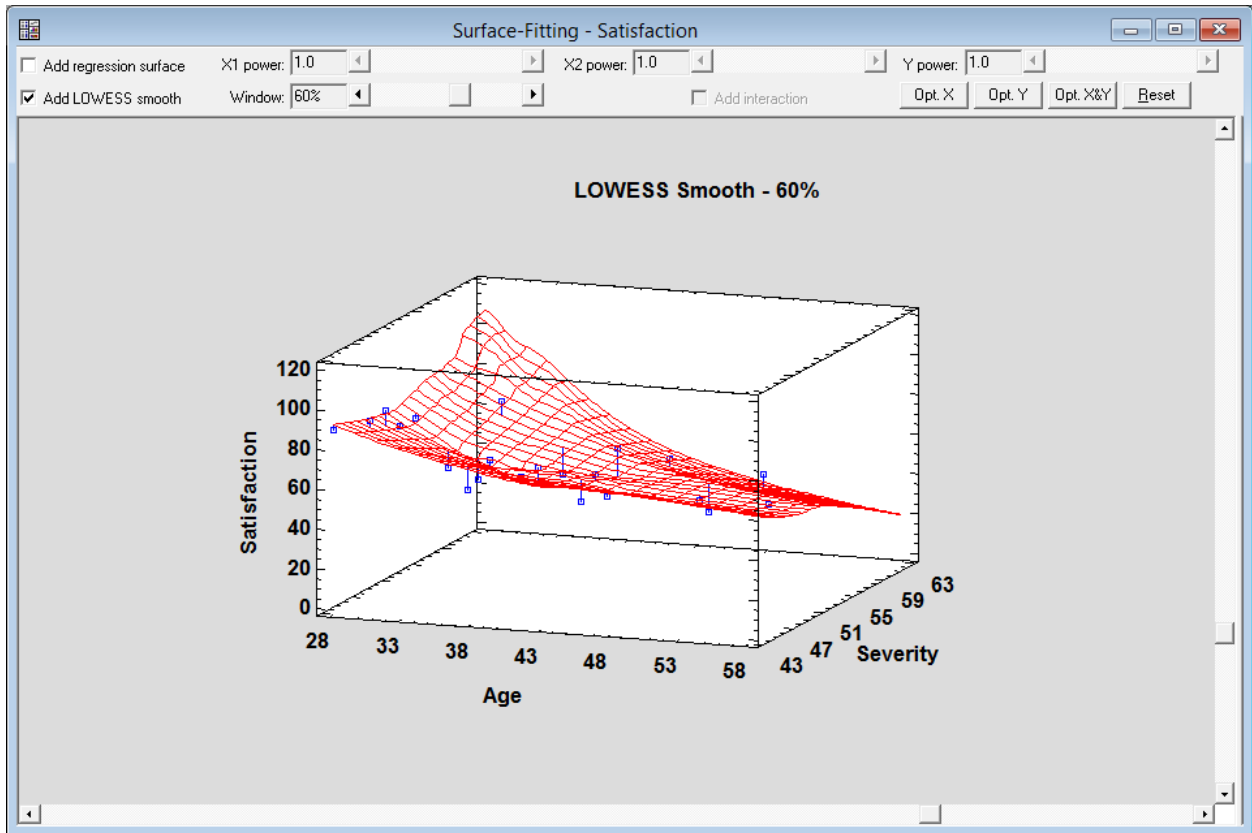
This new procedure fits linear and nonlinear regression models involving a dependent variable Y and two independent variables X₁ and X₂. The general model that is fit contains each variable raised to a power plus an interaction term involving the two independent variables:

$$Y^p = a + bX_1^{q_1} + cX_2^{q_2} + dX_1^{q_1} X_2^{q_2}$$

Using the Statlet control bar, you may interactively visualize the effect of transforming one or more of the variables.



The values of the powers may be changed interactively using the sliders on the toolbar, or the buttons may be used to numerically optimize the powers. A LOWESS smooth may also be added to the plot to compare it with the fitted model:



The *Window* slider may be used to interactively change the width of the smoothing window.

T Chart

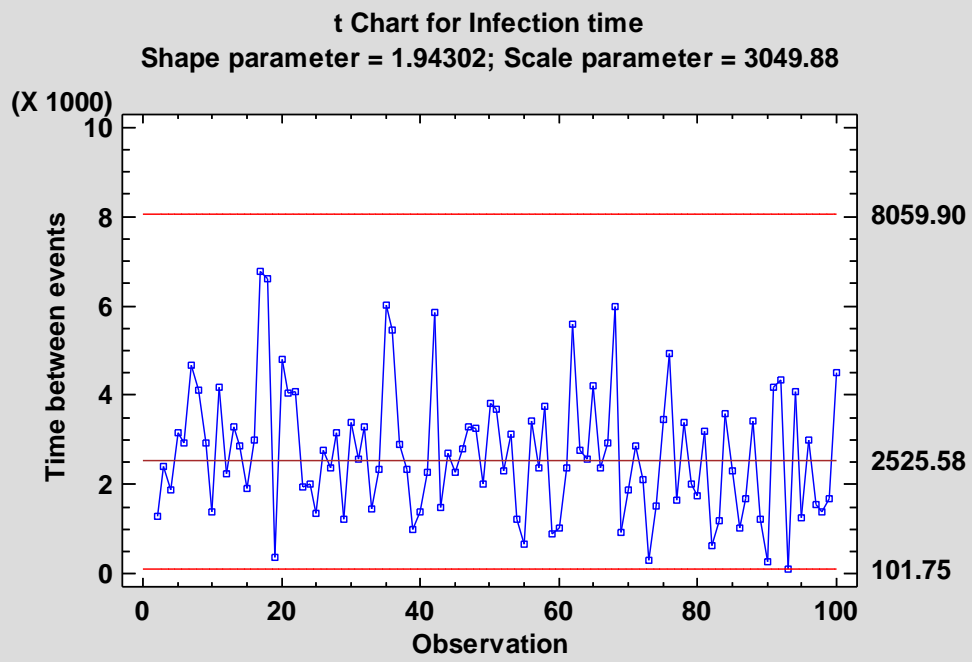
This new procedure creates a control chart for the length of time between the occurrence of rare events. It is based on the Weibull distribution, defined by shape and scale parameters. This type of chart is used frequently in the health sciences to monitor the occurrence of events such as post-surgical infections.

Input data may consist of either the times at which events occurred or the length of time between events:

The screenshot shows the 't Chart' dialog box. The title bar is 't Chart' with a close button. The main area is divided into a large text input field on the left labeled 'Infection time' and a control panel on the right. The control panel includes: a 'Data:' section with a play button and a text box containing 'Infection time'; a 'Type of data' section with two radio buttons, 'Time of occurrence' (selected) and 'Time between occurrences'; a '(Labels:)' section with a play button and a text box; three specification limit sections: '(LSL:)', '(Nominal:)', and '(USL:)', each with a play button and a text box; and a '(Select:)' section with a play button and a text box. At the bottom left is a checkbox 'Sort column names'. At the bottom are five buttons: 'OK', 'Cancel', 'Delete', 'Transform...', and 'Help'.

Specification limits may also be added to the chart.

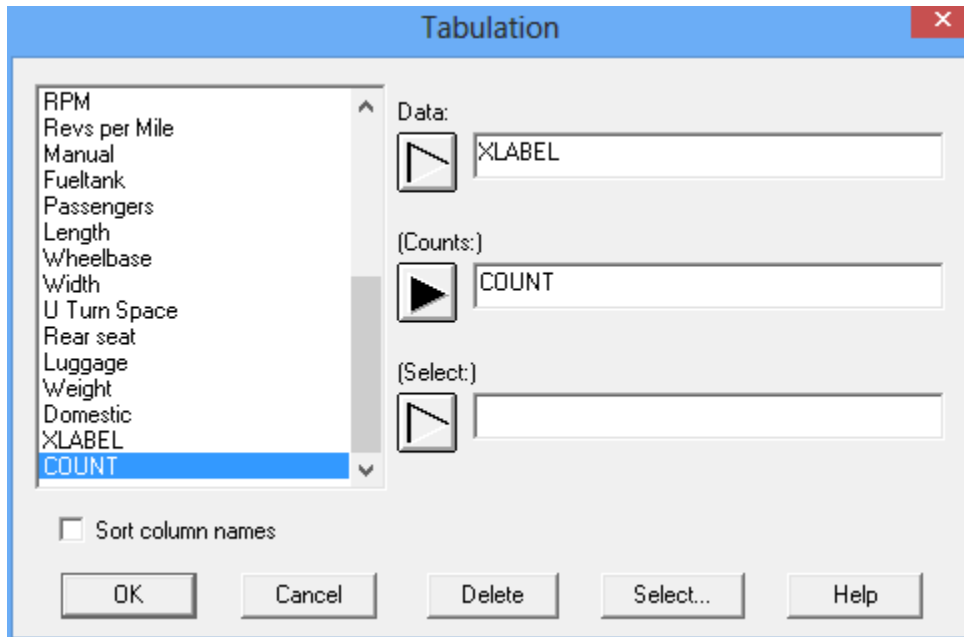
The control chart is similar to others generated by Statgraphics:



The graph also displays the shape and scale parameters.

Tabulation

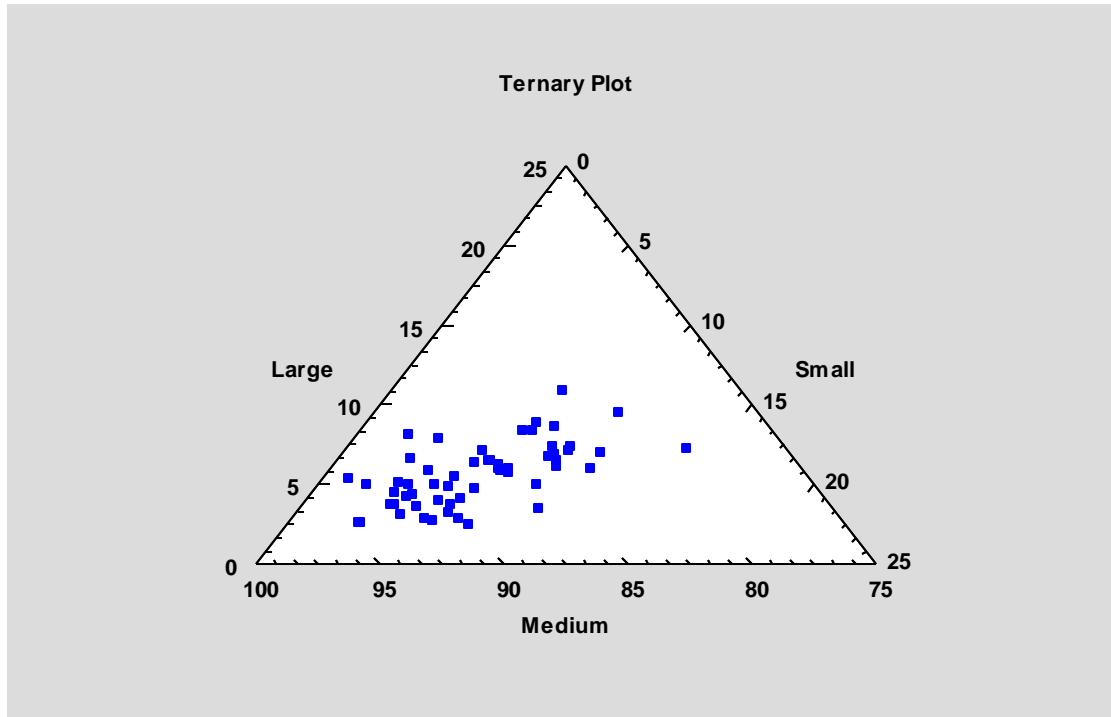
The data input dialog box for the *Tabulation* procedure has been modified to allow for specification of a *Counts* column:



If used, the *Counts* column specifies the frequency associated with each value in the *Data* column. If not used, each frequency is assumed to equal 1 as in Version 16.

Ternary Plot

This new procedure may be used to create a scatterplot of 3 variables which always sum to a constant value. It shows the values of the variables on a plot shaped like an equilateral triangle. It is sometimes called a *simplex plot* or a *de Finetti diagram*.



The *Analysis Options* dialog box sets the minimum value for each axis:

Field	Value
Factor A minimum:	0
Factor B minimum:	75
Factor C minimum:	0
Number of tickmark divisions:	5

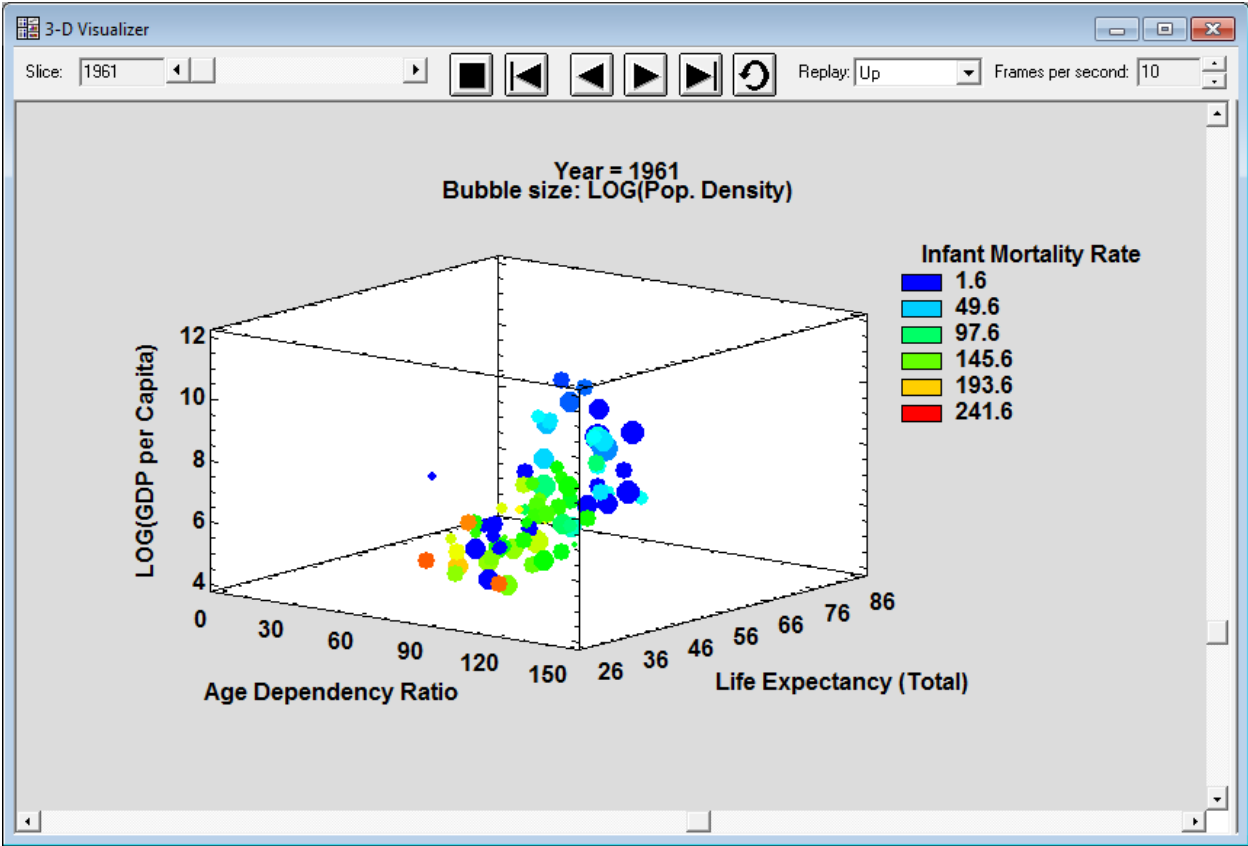
Points may be color coded according to the value of a fourth variable. The direction of the axes may be switched from clockwise to counterclockwise using the *Graphics* tab of the *Edit – Preferences* dialog box.

Three-Dimensional Visualizer

This new Statlet is designed to plot multiple time series in a manner that helps users visualize the changes in multiple variables over time. Given n time series observed over p time periods, the program generates a dynamic display that illustrates how each of the variables has changed over time. Typical applications include plotting:

1. Yearly demographic variables for different countries.
2. Quarterly sales figures for multiple divisions within a company.
3. Monthly economic indices on a state-by-state basis.
4. Daily closing stock prices for multiple equities within a portfolio.

The basic plot shows bubbles plotted on an X-Y-Z display. The positions along the X, Y and Z axes represent the values of three primary data variables. The size and color of the bubbles may be used to illustrate other variables. As time increases, the analyst can follow changes in all of the variables simultaneously. Various options are offered for smoothing the data and for dealing with missing values.

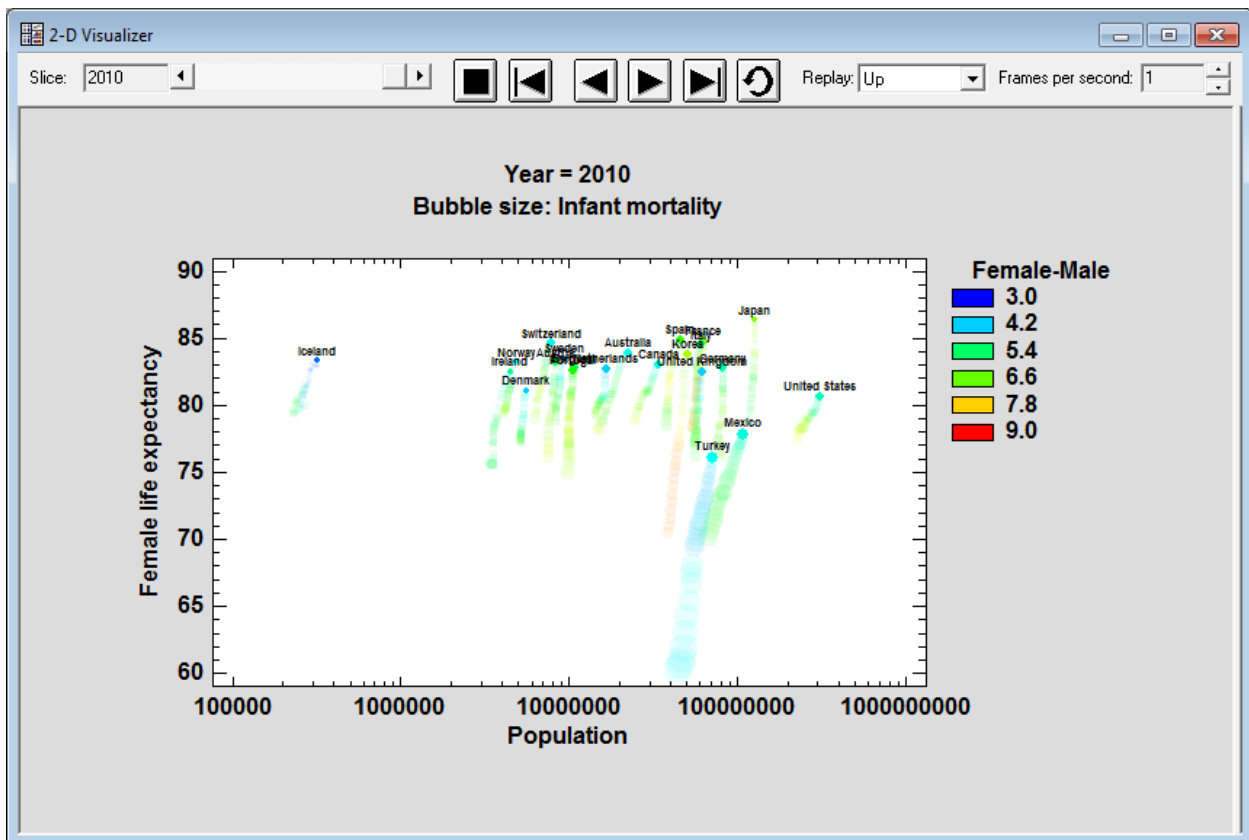


Two-Dimensional Visualizer

This new Statlet is designed to plot multiple time series in a manner that helps users visualize the changes in multiple variables over time. Given n time series observed over p time periods, the program generates a dynamic display that illustrates how each of the variables has changed over time. Typical applications include plotting:

1. Yearly demographic variables for different countries.
2. Quarterly sales figures for multiple divisions within a company.
3. Monthly economic indices on a state-by-state basis.
4. Daily closing stock prices for multiple equities within a portfolio.

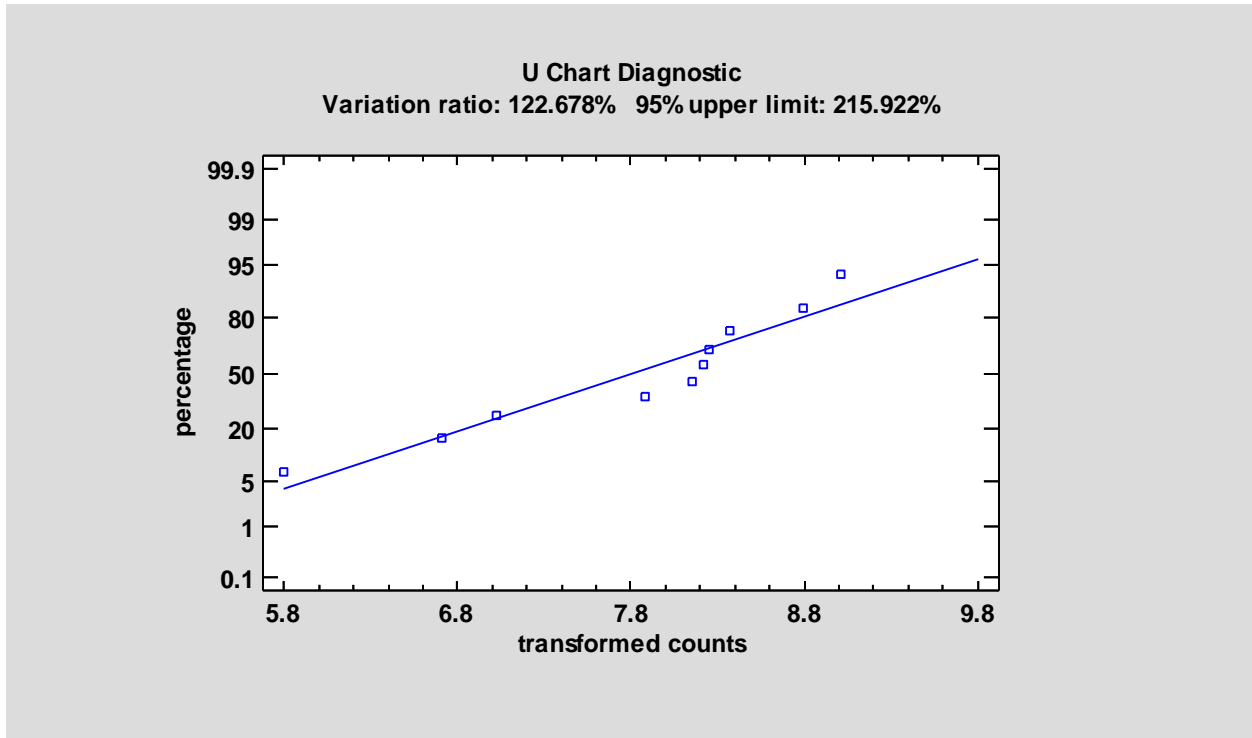
The basic plot shows bubbles plotted on an X-Y display. The positions along the X and Y axes represent the values of two primary data variables. The size and color of the bubbles may be used to illustrate other variables. As time increases, the analyst can follow changes in all of the variables. Various options are offered for smoothing the data and for dealing with missing values.



Breadcrumbs may be added to trace the movement of individual bubbles.

U Chart

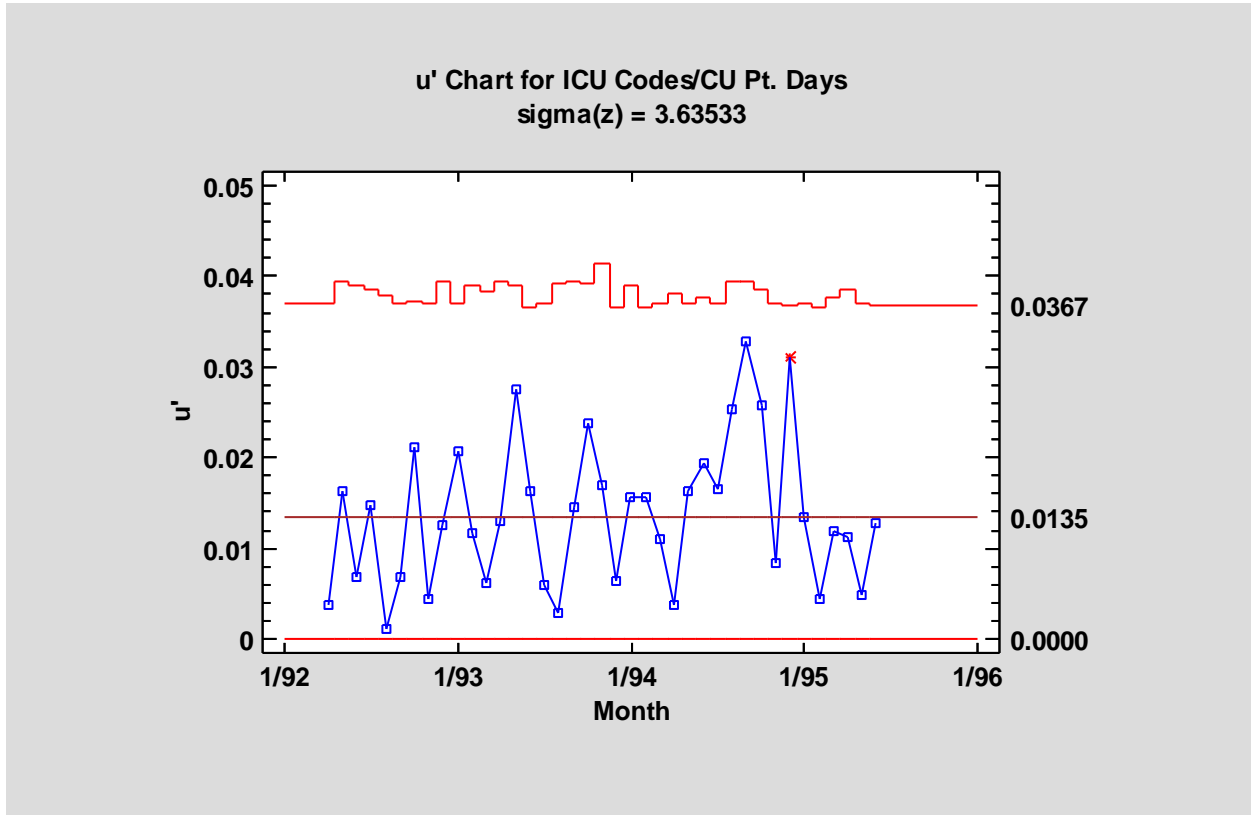
New diagnostic tests have been added to test the assumption that the counts charted follow a Poisson distribution. The diagnostic plot displays a variance ratio, which compares the observed variance of the counts to their expected variation. A variance ratio in excess of the displayed upper limit or less than 60% indicates that the Poisson assumption is not valid.



For data that are overdispersed (the variance exceeds the mean), a new Laney U' chart has been added as an alternative to the U chart.

U' Chart

A new control chart has been added to the program which may be used to plot counts which are overdispersed compared to a Poisson distribution. It can prevent a high rate of false alarms, particularly for large sample sizes.



The graph also displays $\hat{\sigma}_z$, which is an estimate of the relative amount of process variation not explained by the Poisson distribution alone.

Miscellaneous

Data

Users have the option in Version 17 of storing data in the StatFolio. This may be done in two ways:

1. If data is entered in a datasheet but not saved in an external data file, the data will automatically be saved in the StatFolio.
2. If the *Save copy of external data* option is selected on the *General* tab of the *Edit-Preferences* dialog box, a copy of all data in the DataBook will be saved in the StatFolio whenever the StatFolio is saved. If the original data source is not available when a StatFolio is loaded, the stored data will be retrieved instead.

Other data changes include:

1. Imported data files in formats such as Excel or delimited text are now saved on the recent data files list.
2. The maximum length of variable names has been expanded from 64 to 256 characters.

Statgraphics Operators

New operators have been added:

1. ACOSG, ASING, ATANG, COSG, SING, and TANG: for calculating trigonometric functions for angles specified in radians.
2. CONTAINS, ENDSWITH and STARTSWITH: for identifying rows containing a specified string. To be used in the "Select" field on data input dialog boxes.

Toolbar Shortcuts

The main toolbar now allows for 12 user-specified shortcuts rather than 11.