

Release notes for StatCrunch October 2015 updates

This document describes the October 2015 StatCrunch updates.

Major additions:

- The menu paths in the StatCrunch documentation can now be clicked to reveal the location of a procedure. See page 2 for details.
- User-specified polynomials and other functions can now be added to scatter plots. See page 3 for details.
- The “Between” tab for continuous distribution calculators now produces the middle region based on equal tail probabilities when only a probability is specified. See page 4 for details.
- Dialogs are now updated dynamically if columns are added, deleted, or renamed in the data table. See page 5 for details.

Minor fixes and enhancements:

- The most recent paint object can now be dragged to a new location after it is added to a StatCrunch graph.
- When working within a StatCrunch dialog window, ctrl/cmd key sequences are no longer passed to the data table. Text inputs also now process their own undo/redo events.
- Point sizes can now be specified within the Correlation by eye, Regression by eye, and Regression influence applets.
- Titles that are too long to be fully displayed on StatCrunch graphs are now left aligned, making them easier to read.
- The “sum” function has been added to the stats that can be computed and plotted for maps of U.S. states.
- **Data > Simulate** now runs much, much faster when many columns are created.
- A small fix was put in place to keep large tables from being displayed in a tiny fashion on iOS 9.

Clickable menu paths in StatCrunch help documentation

StatCrunch documentation has long provided menu paths that show the sequence of menus that a user can follow to reach a particular StatCrunch procedure. These menu paths are now clickable. When a user clicks on one of these paths, the related menus are unfolded showing the user exactly where a procedure is located within the StatCrunch menu structure. In the example below, the menu path for computing summary stats for columns (**Stat > Summary Stats > Columns**) is shown along with the menu structure for reaching this procedure, which is revealed when the path is clicked.

StatCrunch

Applets Edit Data Stat Graph Help

Calculators >
Summary Stats >
Tables >
Z Stats >
T Stats >
Proportion Stats >
Variance Stats >
Regression >
ANOVA >
Nonparametrics >
Goodness-of-fit >
Control Charts >
Resample >

Columns >
Rows >
Correlation >
Covariance >
Grouped/Binned data >

Summary stats for columns

Stat > Summary Stats > Columns

[View examples](#)

Provides descriptive statistics in tabular format for the column(s) selected. By default, the following statistics are selected: sample size (n), mean, variance, standard deviation (Std. Dev.), standard error (Std. Err.), median, range, minimum, maximum, first quartile (Q1) and third quartile (Q3). The following statistics can also be selected: sum, interquartile range (IQR), unadjusted variance, unadjusted standard deviation, coefficient of variation, skewness, kurtosis, and mode.

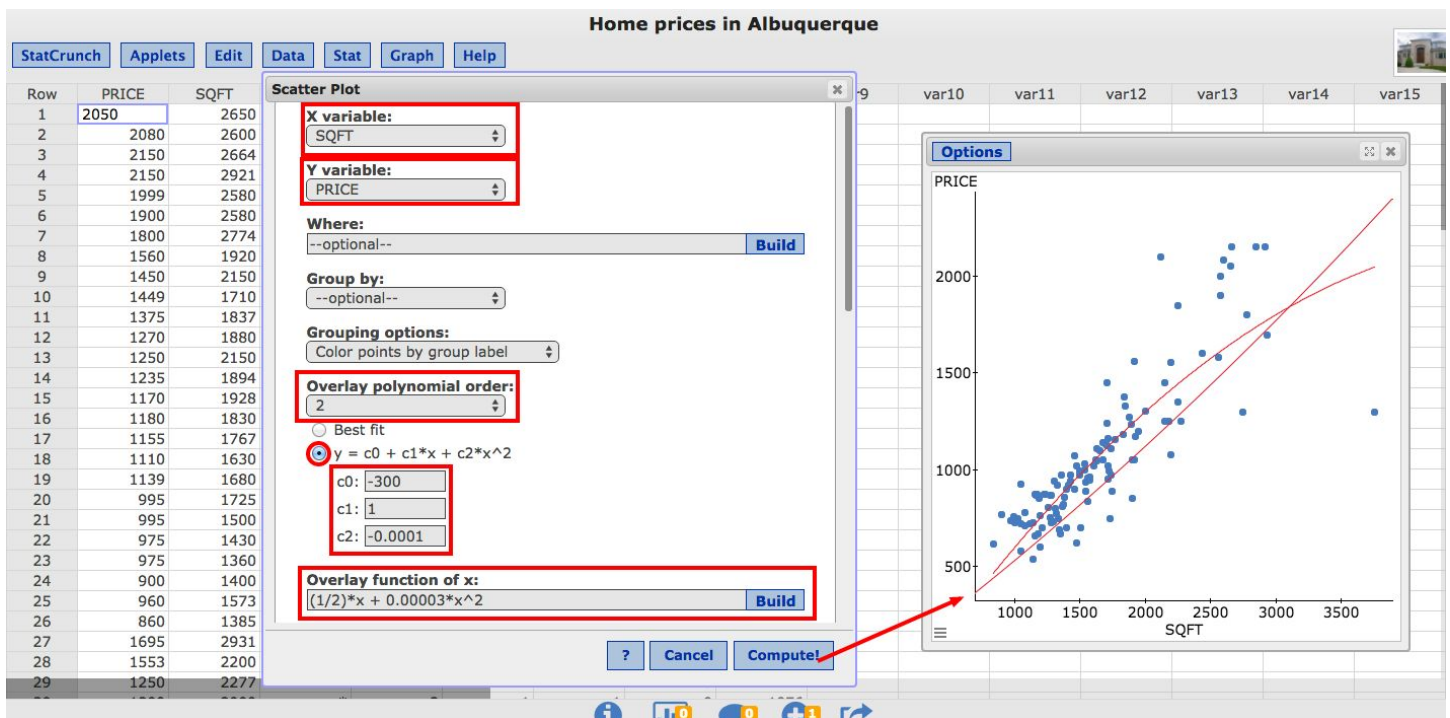
1. Select the columns for which summary statistics will be computed.
2. Enter an optional **Where** clause to specify the data rows to be included in the computation.
3. Select an optional **Group By** column to group results. If a **Group By** column is selected, choose whether to display the output in separate tables for each column selected or in separate tables for each group.
4. Select the summary statistics to be computed. The statistics will be displayed in the order in which they are selected (from right to left). Additional percentiles may also be entered as a space or comma delimited list.
5. Enter a custom **Other statistic** to be included with the results.
6. Check the **Store output in data table** option if the output is to be placed in the data table.
7. Click the **Compute!** button to view the results.

Help Home < >

User specified polynomials and other functions in scatterplot

The **Graph > Scatter Plot** procedure now offers the ability to add custom polynomials along with other functions to the resulting plot. The example below illustrates this new feature. In this example, a scatter plot is constructed using square footage to predict price for the [Home prices in Albuquerque](#) data set. In the dialog window, *SQFT* is selected for the **X variable** and *PRICE* is selected for the **Y variable**. The **Overlay polynomial order** is set to 2 for adding a quadratic function to the scatter plot. After making this selection, new options appear for adding the least squares best fit polynomial or for specifying the polynomial coefficients. A custom polynomial is specified below with coefficients matching the equation $y = -300 + 1*x - 0.0001*x^2$.

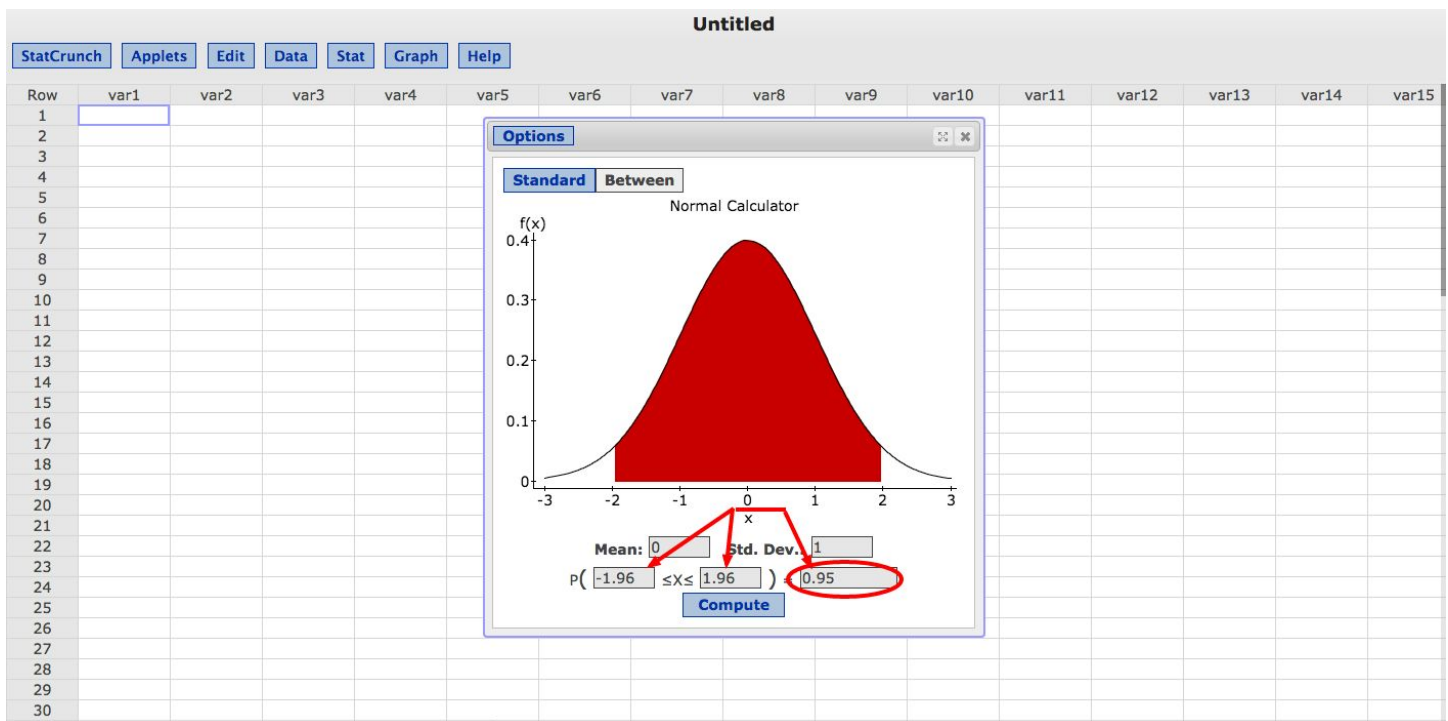
Additionally, a new feature for specifying an arbitrary function to the scatter plot has been added. With the **Overlay function of x** input, the user may specify any function in terms of an x variable in a StatCrunch expression. In the example below, $(1/2)*x + 0.00003*x^2$ was entered to add another quadratic equation to the plot. The scatter plot to the right of the dialog window shows the data points along with both functions overlaid in red. Hovering over one of the red curves with the mouse will result in the display of the corresponding equation in an info bubble



Distribution calculators: finding the middle region with equal tail probabilities

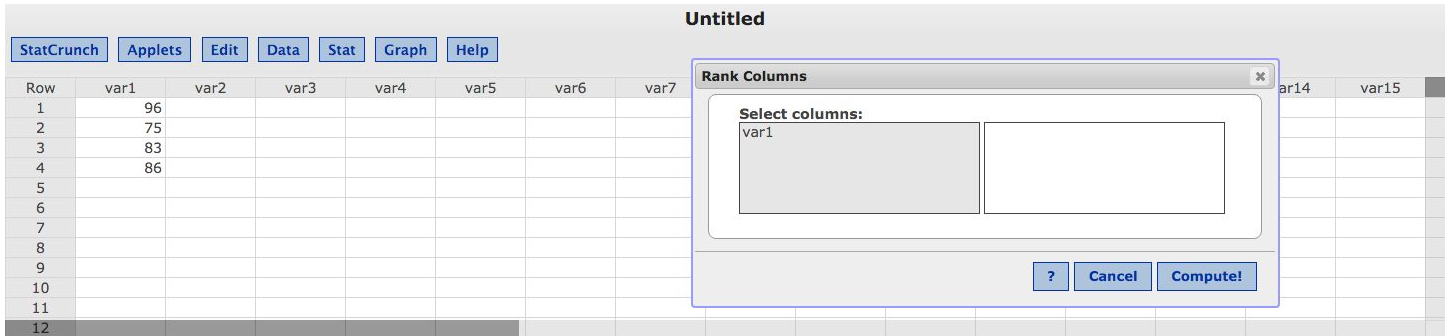
The continuous distribution calculators have been upgraded so that “between” calculations will now compute the middle region if only the middle probability is specified. This functionality is helpful when finding critical values for confidence intervals and hypothesis tests.

The calculator below was created using the **Stat > Calculators > Normal** menu option. The *Between* tab is selected and 0.95 was entered for the probability input. By default, both the lower and upper reference values are blanked out when the probability is specified. Clicking **Compute** with both reference values empty will now produce the reference values corresponding to a middle region with equal tail probabilities on each side. In the example below, 95% of the area under the standard normal distribution is shown to be between -1.96 and 1.96 with 2.5% of the area left over in each of the regions below and above this range.



Dynamically updated dialogs

When a column is added, deleted, or renamed in the data table, that change will immediately be reflected in any existing StatCrunch dialogs. For instance, the dialog below for the **Data > Rank** procedure shows the column *var1* from the data table.



If a new column, *var2*, is added, the dialog updates accordingly to reflect the new column as shown below. (In the past, the dialog would not have been updated when the data table changed.) Such updates will be applied to any dialogs that correspond to results listed under **StatCrunch > Session Results**.

