

# Release notes for StatCrunch April 2015 updates

This document describes the April 2015 StatCrunch updates.

## Major additions:

- Markers and dividers have been added to simulation, resampling, and hypothesis testing applets. See page 2 for details.
- Suggested values have been added to the expression builder. See page 3 for details.
- Color scheme creation/editing has been significantly overhauled. See page 4 for details.

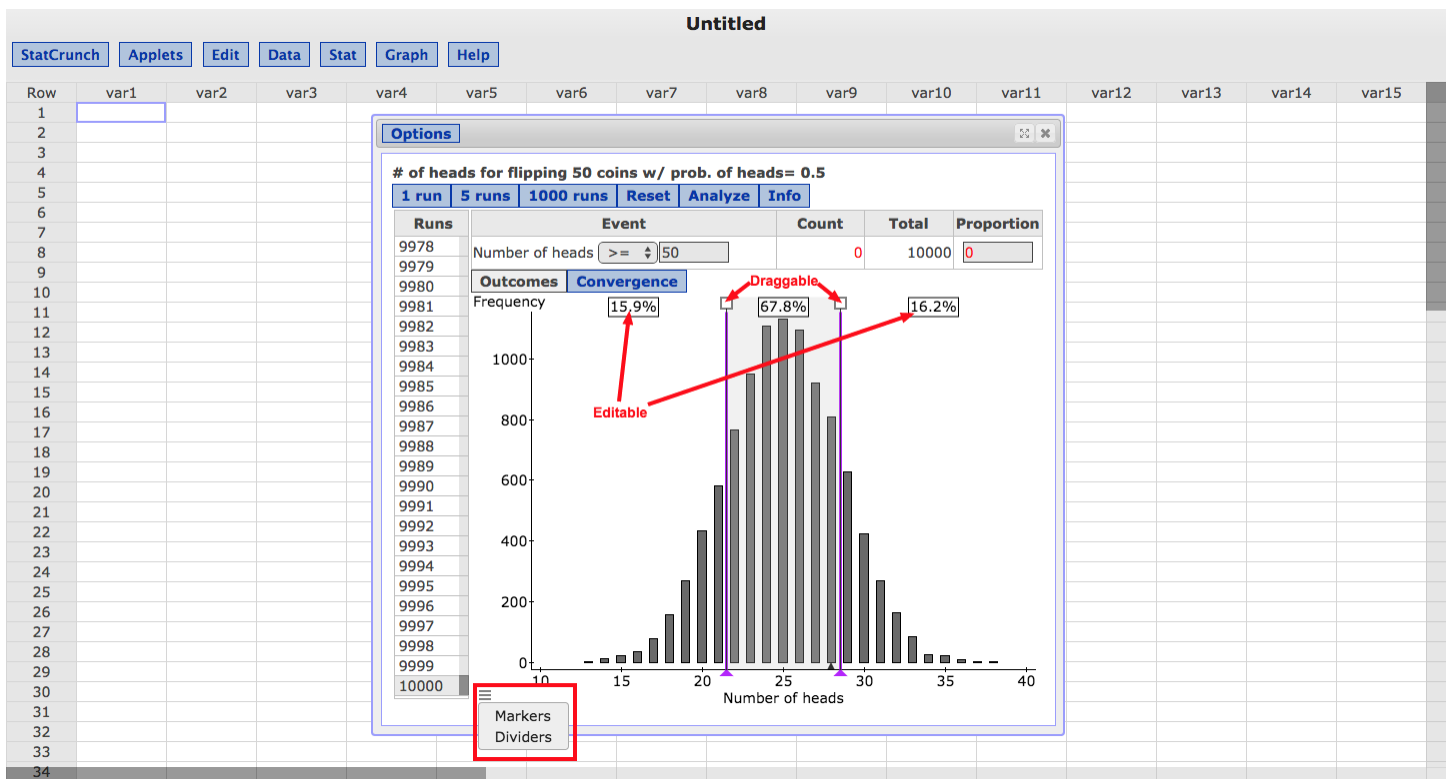
## Minor fixes and enhancements:

- The repeated measures procedure now offers a Tukey HSD option.
- Implemented a minor bug fix and an interface upgrade for the distribution calculators.
- Implemented a bug fix when an overlay and dividers are both added to a histogram.
- Fixed a bug in the hypothesis testing applets with the tallying of results.
- Dotplot now automatically sizes to minimize the empty white space when created.
- An issue has been fixed with the expression builder in resampling applets.
- The urn sampling simulation applet has been updated to initialize faster.
- Double-clicking capability provided for adding columns and functions in the expression builder.
- Tallying issue fixed regarding the use of proportions with simulation applets.
- More digits of accuracy have been added to the observed value in the resampling applets to allow for consistent results with the new dividers and results saved to the data table.
- Dividers have been updated so that the middle region is inclusive of the edges.
- The data table has been updated with Improved parsing of double quoted values.
- Numerous bug fixes were implemented for the simulation applets when used for nonstandard calculations.
- Contingency table procedures fixed to provide the correct cell format when chi-square residuals are used.
- An interface update was implemented for the power/sample size calculators to improve their ease of use.
- The simple linear regression routine was updated for faster results with larger data sets.
- The data table was upgraded for much faster creation of large user-generated data structures.

# Markers and Dividers for Applets

Markers and dividers are tools that allow users to overlay and extract more information from graphs. A marker is represented by a triangle anchor on a graph's axis with a line that extends from it. The location of the anchor can be calculated based on characteristics of the data (e.g. the mean or the median of the data) or it can be set to a specific numeric value. To use data characteristics, an appropriate expression must be entered where  $x$  is used to generically represent the data (e.g.  $mean(x)$  and/or  $median(x)$ ). The color of the anchor and its corresponding line can also be set by the user using a basic color chooser. An unlimited number of custom markers can be added to a graph with the mean and median provided as default options. Dividers are a tool that allow data values in a graph to be split into user-specified regions with the count/percent of the data in each region displayed. The regions are defined using two square buttons positioned at the top of the graph with lines that extend below them. The two split points can be set explicitly or interactively by the user by either double clicking on the button or by dragging the button to a specific location. The percentage falling in a particular region can also be specified using a double click on the associated label which will lead to a repositioning of the divider. Dividers are designed to work in unison with markers. If a divider button is dragged in the vicinity of a marker location, the value of the divider will "snap" to the value of the marker.

These new tools, which were previously made available for various distributional graphs in StatCrunch, are now available for simulation, resampling, and hypothesis testing applets. The markers will update their location and the dividers will update the information displayed as new observations are created. The example below uses the coin flipping simulation applet which was configured to repeatedly count the number of heads in 50 coin tosses. The three-line icon to the bottom left of the applet graphic provides options for adding markers and/or dividers to the graph. In this example, dividers are used in conjunction with two markers at  $mean(x) \pm std(x)$  to show that roughly 68% of the time the number of heads will fall within one standard deviation of the mean.



## Suggested values added to expression builder

The expression builder is the pop up tool connected to the **Build** buttons attached to StatCrunch inputs that support expressions. It supports standard keyboard entry, and it also allows users to interactively build expressions. The builder provides a numeric keypad with mathematical and logical operators, a list of columns available for the calculation and a complete list of all StatCrunch functions that can be used in the calculation. The builder also offers a listing of previously used expressions which can be used for the current expression. Both columns and functions can be added to an expression by double clicking on the appropriate name in the listings. When columns are selected in the listing, suggested values are now presented to the user so that they can be easily added to the expression. Up to 5,000 unique values for the selected column are available in the listing of values.

With this addition, the expression builder is now highly recommended as the tool for constructing all expressions in StatCrunch regardless of the user's experience level. The builder will automatically use double quotes to properly enclose column names or values that contain special characters such as spaces or mathematical operators. The builder also allows users to avoid any issues with case sensitive references to column names and text values. The expression builder also makes the construction of complex expressions much easier on mobile devices where keyboard input is more laborious.

The example below shows how to use the expression builder and the new suggested values feature to easily specify groups for a two sample t test. The data set used in the example contains the scores of 20 students randomly split into two groups of 10. One group takes the exam on yellow paper and the other group takes the exam on white paper. By clicking on the **Build** button attached to the **Where** input for each sample, the user can create the proper expression to identify each group very quickly and easily. In this case, using the tool results in no need to use the keyboard for typing.

The screenshot displays the StatCrunch interface for a two-sample t-test. The main window shows a data table with columns for Row, Score, and Color. The 'Two Sample T' dialog box is open, showing configuration for Sample 1 and Sample 2. The 'Where' field for Sample 1 is set to '"Color" = "Yellow"', and the 'Build' button is highlighted with a red arrow. The 'Expression' dialog box is also open, showing the same expression and a list of suggested values for the 'Color' column, including 'Yellow' and 'White'. A red box highlights the 'Values' field and the 'Add' button, with a red arrow pointing to it from a label 'New suggested values feature!'.

Row	Score	Color
1	95	Yellow
2	90	White
3	89	Yellow
4	85	White
5	81	Yellow
6	81	White
7	80	White
8	79	Yellow
9	79	Yellow
10	78	Yellow
11	77	White
12	75	Yellow
13	73	White
14	71	Yellow
15	65	Yellow
16	61	Yellow
17	60	White
18	59	White
19	55	White
20	50	White

## Edit > Color Schemes has been re-designed

The **Edit > Color Schemes** procedure now features a full swatch with all of the colors for each color scheme along with options for creating new standard or scaled color schemes. A standard color scheme cycles among the colors in the scheme repeatedly if necessary. For instance if the color scheme has 7 colors, excluding foreground and background colors, and a pie chart has 14 slices, then each color will be used for two slices. A scaled color scheme specifies only the starting and ending colors. For a graph, exactly enough intermediate colors are selected so that each color in the scaled color scheme is used only once.

Color schemes can also easily be edited within this procedure. After clicking the name of a color scheme to edit it, each color can be modified directly by simply selecting a color from a continuous palette or by specifying RGB values. Colors can also be inserted, deleted, and re-ordered within a color scheme by using the buttons adjacent to each color in the scheme. A color scheme can be specified as the default scheme to use in all StatCrunch graphs by checking the corresponding option. The user will then not need to explicitly specify the color scheme for each graph. The example below shows the editing of the newly added “Soft” color scheme which is a standard scheme with 9 colors. The palette is shown for modifying a color in the scheme, and the checkbox for making the scheme the default color scheme is also highlighted.

The screenshot displays the StatCrunch software interface. At the top, there are menu buttons: StatCrunch, Applets, Edit, Data, Stat, Graph, and Help. Below the menu is a grid with columns labeled var1 through var7. A 'Color Schemes' dialog box is open, showing a list of color schemes with their respective color swatches. The 'Soft - 9 colors' scheme is highlighted with a red box and labeled as '[DEFAULT]'. Below this list are buttons for 'Add standard' and 'Add scaled', and a 'Close' button. To the right, the 'Edit Color Scheme' dialog for 'Soft - 9 colors' is open. It features a 'Restore original colors' button and a checked checkbox labeled 'Make this scheme the default', which is also highlighted with a red box. The dialog lists 11 items: 1. Foreground (black), 2. Background (white), 3-9. Various colored squares, each with '+' and 'x' buttons. Item 10 is a color palette for selecting a color, showing a gradient from black to white and a vertical rainbow bar. The selected color is green, with the RGB value 'rgb(179, 222, 105)' displayed below it. Item 11 is a 'cancel choose' button. At the bottom of the dialog are '?', 'Cancel', and 'Update' buttons.