

GETTING STARTED WITH ANALYSIS¹

RegressIt is completely menu-driven and easy to use, it has very extensive built-documentation and teaching notes, and the documents on the program-features web pages and download pages at regressit.com provide detailed instructions. If you have some experience in regression analysis, you should find it to be more self-explanatory and more fun than whatever software you were previously using. If you're new to the subject or just a bit rusty, here's a list of the steps to follow to get started and do some data analysis.

1. Launch Excel and load your data: open the data file you wish to use, or else type or copy-and-paste data onto the first worksheet in a blank file. The data should be arranged columnwise with text labels in the first row that are to be used as variable names.

2. Launch RegressIt by opening the program file (RegressItPC, RegressItMac, RegressItMac2011, or RegressItLogistic). If you get a macro security prompt, choose "Trust all from publisher" if that button is present, or else choose "Enable". (You will not see the security prompt in the future if you choose the "Trust" option.) On a PC or on a Mac with Excel 2016, you should then see a RegressIt tab added to your menu at the top of the Excel window. In RegressItMac2011 (for Excel 2011), the RegressIt menu is a pop-up menu that is activated at any time by hitting Control-G. (This also works in the other versions.) Also please note:

- On a PC a newly downloaded program file must be "unblocked" before it will run. In the File Explorer, right-click on the program file and choose **Properties/Unblock/Apply/OK**. This only needs to be done once.
- On a Mac you will not see the menu appear until a data file has also been loaded.
- Excel's macro security level must be set to "disable all macros with notification" in order for RegressIt to run. This means that you will be prompted to enable the program when it is launched. Probably the macro security level is already set this way on your computer. If the program doesn't run, go to **File/Options/Trust Center/Trust Center Settings/Macro Settings** and check to see that the disable-with-notification option has been chosen there.
- If the program file will not download or will not run on a PC even after unblocking, this could be due to an additional layer of security on your computer. Download it in zip format instead, then unzip and unblock it.
- **On a PC you should close the File Explorer while running analyses**, because it may cause errors when RegressIt uses the clipboard for producing graphs or exporting text data.

3. Name your variables (first time only): hit the **SELECT DATA** button on the menu, which ought to select the whole data range automatically, and then hit the **CREATE NAMES** button. (This merely executes Excel's create-name-from-selection command.) You will see a pop-up box with four checkboxes for locations of the names, and **you should check ONLY the "Top row" box**. This will assign the text labels in row 1 as variable names (Excel range names) for the data underneath them. The following details apply to text data:

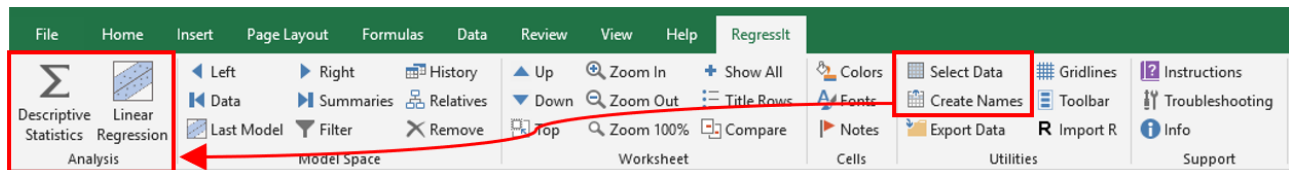
- The variable lists for descriptive statistics and regression models will show only variables that have one or more numerical values in their columns. **If some of your variables are encoded in text form, you won't see them in the dialog boxes**. Sometime dates, for example, are encoded as text rather

¹ July 29, 2020. Complete documentation and the program files and sample data files can be found at regressit.com.

than numerical dates ("2005Q1", "2005Q2"...). If this is the case and you want to treat them as numbers, you will need to recode them in number format (2005.0, 2005.25...) or Excel date format before trying to assign names.

- If some variables are text labels for categories, say, a variable called Sex with values "Male" and "Female", RegressIt will not automatically recognize them as such. However, you can use its **variable transformation tool** to instantly create dummy (0-1) variables for them with default names such as Sex.Eq.Male and Sex.Eq.Female, of which one or the other could be used as an independent variable. This tool will show you the full list of variables to choose from, even those that are all text. It can be launched from the descriptive stats or regression dialog boxes.

4. Save your file under the name you want to use for the master copy of your not-yet-analyzed data set.



Now for some analysis...

5. If you want to get summary statistics, correlations, and plots of your data, click the **DESCRIPTIVE STATISTICS** button on the RegressIt menu. This will open a dialog box with a list of your variables and options for descriptive analysis. This procedure will always generate a descriptive statistics table and a correlation matrix, and there are options for some additional statistics and various kinds of plots. **Select your variables, click the boxes of any optional outputs you would like, and hit Run.** A new worksheet will be created with the results, and initially it will be displayed in collapsed form with only header rows showing. **Hit the SHOW ALL button on the RegressIt menu, or click the plus signs next to the header rows, to open up the tables and display all the results.** If you toggle the **COLORS** button on the menu, you can turn on color-coding for the correlations according to their sign and magnitude. Try it and see! See the instruction manual or the Descriptive Statistics web page for pictures of these operations.

6. Save your file under a new name at this point. In general it is a good idea to save your file after each new analysis and also update the name periodically to indicate the general stage you have reached in your analysis. You should always keep a clean copy of your original data file, in addition to files with analysis.

7. If you want to fit a regression model, hit the **LINEAR REGRESSION** button on the menu. This will open up the linear regression dialog box. **Choose your variables and the optional outputs you would like to get, then hit Run.** The results will be generated on a new worksheet. If you do not choose any options, you will get minimal output: the summary statistics table (R-squared and all that) and the coefficient table. **It is highly recommended that you should select some chart output, minimally the "standard residual plots" option.** This will produce plots of residuals (errors) versus predicted values, residuals versus row number, and a residual histogram chart. The "actual and predicted vs. observation number" chart is also highly relevant if rows refer to consecutive time periods in your data set. If your model is a simple (1-variable) regression, you should also check the box for "line fit plot". **Click the plus signs next to the options in the regression dialog box for pop-up notes about their functions.** See the instruction manual or the Linear Regression web page for pictures of these operations.

8. Save your file (again)!

9. A novel feature of RegressIt is that when regression models are created, it keeps a **MODEL SUMMARIES WORKSHEET** that shows their error statistics and coefficient estimates side by side. This provides an audit trail for your work as well as a close-up view of what happens when you add or delete variables from regression models. Normally it is the far right worksheet, and you can jump directly to it by hitting the **Summaries button** on the menu. Hit the **Show All button** to see everything that is on it. Also click the **Colors button** to toggle color coding of coefficients by sign and significance.

10. Another novel feature of RegressIt is the **TEACHING NOTES** feature. If you check this option in the Linear Regression dialog box, teaching notes will be embedded in the output worksheet in the form of cell comments. Toggle the **Notes button** on the menu to turn on little red flags that indicate cells with notes in them. If you hover the mouse over a cell with a flag, the note will pop up.

11. After you've added some analysis worksheets to your file, you should **play with the other buttons on the RegressIt menu** to see how they can help you to navigate and study your output in a more systematic (and fun) way than you can do with most other software. For example, the **VCR buttons** allow you to move left or right one worksheet at a time or jump directly to the first sheet (data) or last sheet (model summaries). The **Last Model button** allows you to flip back and forth between the last two model worksheets you've looked at (very handy). The **History button** will show a list of all your analysis worksheets, including some of their statistics, and allows you to jump directly to any of them. The **Up and Down arrows** allow you to page up and down the worksheet by whole table or whole chart so that the topmost one is always aligned with the top of the screen. If you click on any cell in the coefficient table and then hit the **Filter (funnel) button**, you can sort the variables according to their P-values. You can also **zoom all sheets** at once, twiddle the **confidence level**, and show or hide title rows. **If you hover over a button, you will see a pop-up description of its function.**

