

Tech Support¹

RegressIt has been extensively tested and ought to work fine on most computers as long as the directions are followed. When technical problems occur, they are usually due to one of a number of common issues that are avoidable or fixable as described below. If you encounter a problem that cannot be solved by one of those actions, you are welcome to send a tech support request to support@regressit.com. Be as specific as possible about the nature of the problem, what you were doing when it occurred, and what kind of computer you are using. Also let us know the country in which you are located in case a language or number format issue may be involved. However, please note that this is free software and it is offered "as is" as a public service. Before asking for help, you should read the material below. Also, visit the [download page](#) and follow the link for your program version to see if any updates have been posted there. If you are an instructor or TA, please read this document very closely to be able to respond quickly if students have questions, and also see the additional comments on the [advice page](#).

If you have any concerns about the accuracy of the calculations that are performed by RegressIt, just use its [R interface](#) to run an identical regression analysis in R (which takes just a few seconds and sends output back to Excel) and compare the results on the two worksheets.

Most common RegressIt technical questions and answers:

1. Excel macro security settings should be "disable with notification"
2. Program file must be unblocked before first use (PC only)
3. Smart card message when starting
4. Zipped version may be needed to avoid blocking by anti-virus software
5. Updating your version (replace old file with latest one and don't change the name--delete or rename the old file first)
6. Exit from File Explorer in Windows 10 before running analysis
7. Don't use Euro Currency Tools (RegressIt will turn it off anyway)
8. Language and number formats must use U.S. English standards during a session (e.g., "1,001.5", not "1.001,5")
9. Start a new project from clean columnwise data on the first sheet in a new xlsx file with variable names in the first row
10. Variable names should be self-descriptive with no funny characters in them (\$, %, #, €, etc.)
11. Variables won't appear in variable-selection dialog boxes if numbers are coded as text
12. Jiggle the worksheet on a Mac to close chart tabs if necessary
13. Don't edit model worksheets or the model summaries worksheet
14. Descriptive statistics output is hidden at first. Click "Show all" or "+" in left sidebar to display the tables and charts. The same is true on a model worksheet for the model equation in text form, the ANOVA table, the residual table, and the matrix of correlations of coefficients.
15. Be aware of missing values and how they are coded
16. R-squared equal to 1?
17. Model won't run because of multicollinearity among independent variables
18. Memory usage for low-res, high-res, and editable chart output (different on PC's and Macs)
19. Model size limitations: 200+ variables, 1M rows, 5M cells in total for a model on a PC, somewhat less on a Mac. Much larger models can be fitted via the R interface

¹ November 21, 2020, regressit.com

20. Running times for very large models
21. Conflicts with other software
22. Recovering from crashes

Details:

1. **MACRO SECURITY SETTINGS.** Your Excel macro security setting should be "disable all macros with notification", which means that macros will be disabled by default unless you choose "enable macros" at the security prompt when the program file is opened. On a PC, go to **File > Options > Trust Center > Trust Center Settings > Macro Security Settings** and check to see that the disable-with-notification option is chosen. (Probably it already is.) You can also choose the "trust all from publisher" option to avoid the prompt in the future. On a Mac, the command sequence is **Excel > Preferences > Security and Privacy** to get to the macro security menu. It is possible to add RegressIt to the official add-ins list so that it will load automatically with Excel, but it is recommended that you not do this until you've thoroughly tested it and are planning to use it on a regular basis.
2. **UNBLOCK A NEW PROGRAM FILE ON A PC.** On PC's, program files that have been downloaded from the internet are blocked from execution by default. Before the program will run, you need to go to the File Explorer, find the RegressIt program file, right-click on it, then click **Properties > Unblock > Apply > OK**. This only needs to be done once.
3. **SMART CARD MESSAGE WHEN STARTING.** If you get a message saying that you need to "connect a smart card" when trying to run the program on a PC, you may also need to change the Trust Center settings to make the RegressIt program folder a "trusted location". Click **File > Options > Trust Center > Trust Center Settings > Trusted Locations**. Then in the Trusted Locations dialog box, click "Add new location", then "Browse", then browse to the folder where the RegressIt program file is located, click on the folder, then hit "OK" twice. If you don't get the smart card message, you don't need this step.
4. **ZIPPED VERSION MAY BE NEEDED TO AVOID BLOCKING BY ANTI-VIRUS SOFTWARE.** In some cases anti-virus software may add a separate layer of protection against downloaded program files on either a PC or Mac. If nothing happens when you try to launch the program, even after checking your macro security level and unblocking (on a PC), download the program file in zipped form instead, then unzip it, and try again.
5. **HOW TO UPDATE YOUR VERSION.** If you install an updated version to replace an earlier one, you should first delete or rename the old program file and then save the new one in the same folder, so that you will only have one file named RegressItPC.xlam or RegressItMac.xlam or RegressItLogistic.xlam on your computer, and it will always be the newest one and always be in the same place. This is especially important if RegressIt has been added to your add-ins menu.
6. **CLOSE THE FILE EXPLORER.** In Windows 10, you should exit from the File Explorer before doing any analysis in a RegressIt session. Otherwise you may get one of the following errors when running an analysis: "cannot paste the data" or "paste-special method of worksheet class failed" or "application-defined or object-defined error." These errors occur because File Explorer interferes with use of the clipboard, which RegressIt needs for performing some calculations, converting graphs to picture format, and using text export options in the History feature and the R interface. Another symptom of a conflict with File Explorer is that nothing will appear when you try to paste text that has been stored on the clipboard. Similar errors can also be caused by some other conflicting software, as discussed below.
7. **EURO CURRENCY TOOLS.** The Euro Currency Tools add-in may cause errors in calculations in Visual Basic applications such as RegressIt, and RegressIt will turn it off by default when it is loaded. If you normally use Euro Currency Tools, you will need to go to the add-ins manager and turn it back on after your RegressIt session.

8. **LANGUAGE AND NUMBER FORMATS.** Language and number formats must use U.S. English standards during a session. Most importantly, numbers need to be formatted with a period as the decimal separator and a comma as the grouping separator. For example, the number "one thousand and one and one-half" should appear on the screen as "1001.5" or "1,001.5" depending on whether a grouping separator is used. RegressIt will generate a "type mismatch" error if the computer's default number format is one that uses a convention other than this. Also, the use of non-English characters in names of variables or files or directories will sometimes cause problems, even if the number format is OK. It is not hard to change your language or number format settings temporarily while using RegressIt, and it is usually possible to set up file paths that use English characters for work to be performed with RegressIt. The new format settings are only active temporarily during your analysis session and they are not hard-coded in the output worksheets. When you later change the settings back to their original values, worksheets created with RegressIt will show your familiar formats. For more details, see the language and number format instructions in the pdf document on the [download page](#) for your version of the program.
9. **START WITH CLEAN DATA ON A SINGLE WORKSHEET.** A new analysis project should start with an Excel file in which there is a single worksheet with data stored on it in value format in **columnwise orientation with variable names in the first row**. It is also OK to start from a text data file (e.g., CSV file), but you need to save it as an xlsx file before continuing. If the data is obtained from an Excel file in which modeling has already been performed with RegressIt or another add-in and/or which contains multiple worksheets, you should copy it to a new file before proceeding. However, you should not use the move-or-copy-worksheet command or the default paste-formulas option, nor should you merely delete worksheets other than the one that contains the data. Instead, hand-select the data range on the data worksheet in the original file, copy it to the clipboard, and use the **Edit/Paste-special/Values** command to paste it to the first worksheet in the new file. If the move-or-copy-sheet command or the paste-formulas option is used, hidden information that will interfere with future analysis could be transferred as well. If the very first analysis that you run does not have "Stats 1" or "Model 1" as the default name that first appears and/or if you are told that variable names already exist when you try to use the Create Names tool, you are not starting from clean data.
10. **USE DESCRIPTIVE VARIABLE NAMES WITH NO FUNNY CHARACTERS.** The variable names in the first row should contain only English-alphabet letters, numbers, spaces, underscores, or periods, not any other symbols (e.g., currency symbols) or punctuation marks, and you are strongly encouraged to use names that are self-descriptive because they will appear everywhere in the output in table and chart titles. At the beginning of the analysis, the names in the first row should be declared as variable names for the columns of data below them by clicking **Select Data** and **Create Names** on the RegressIt menu. This will select the whole data range and execute Excel's create-from-selection command to assign the names, and you should choose only the "Top row" option. (If column A contains text, Excel will try to create a named variable for every row as well as every column if the "Left column" box is checked.) Blanks in column headings will be converted to periods in variable names when the Create Names operation is performed.
11. **VARIABLES WON'T APPEAR IN VARIABLE-SELECTION DIALOG BOXES IF NUMBERS ARE CODED AS TEXT STRINGS.** It occasionally happens that the values in a data column for a numeric variable are coded as text strings, not as numbers. One indication of this is that the numbers may be left-justified rather than right-justified. In this case the variable will not show up in any of the dialog boxes except the one for creating dummy variables. One possible cause is that the string of digits in the cell is padded with spaces at the beginning or end. To convert the data to numeric form, select (highlight) the whole range of values (not including the variable name in row 1). You ought to see a yellow warning sign (<!)> pop up beside the first cell, and if you click on it you should see a drop-down menu with an option to convert text to numbers.
12. **JIGGLE THE WORKSHEET ON A MAC IF NECESSARY TO CLOSE CHART TABS.** On a Mac, if the output worksheet has been produced and you see a "finished" message in the lower left of the window but

the program appears frozen, jiggle the worksheet slightly with your finger on top of the mouse. It is possible that a chart tab has not closed and is waiting for a hand wave from you. If it is OK to move the mouse while an analysis is still running on a Mac, but you should NOT hit any keys on the keyboard, because that may interrupt processing.

13. **DON'T EDIT OUTPUT WORKSHEETS.** You should not edit regression model worksheets or the model summaries worksheet by deleting rows or columns or by editing the model name or worksheet name or any other text on a worksheet. Also do not hide any rows or columns. Take advantage of the opportunity to assign descriptive variable names and model names before running the analysis. RegressIt's workbook navigation tools and memory for model parameters will not work properly if the original names are not found where they are expected. If you want to edit a model worksheet or the model summaries worksheet for use in a report, you should first make a copy of the original worksheet, move it to the right of the model summaries worksheet, and edit the copy. It is OK to delete an entire model worksheet. It will remain part of the audit trail via its entry on the model summaries worksheet.
14. **DESCRIPTIVE STATISTICS OUTPUT IS HIDDEN AT FIRST.** On new descriptive statistics worksheets, the output will be hidden at first. This is to avoid overwhelming you if a huge amount of output has been produced. (There could be dozens or even hundreds of charts for large sets of variables.) Click the "Show all" button in the "Worksheet" group to show all the output. Alternatively, click the "+" signs in the left sidebar to open up one block of output at a time. **The same is true of the model equation in text form, the ANOVA table, the residual table, and the correlation matrix of coefficients on a regression model worksheet.** The ANOVA table is not very informative to look at in most cases since it involves numbers in squared units (also the F-stat is merely the square of the t-stat of the independent variable in a simple regression and their P-values are the same), and the residual table and correlation matrix could be huge, so these are all hidden at first and must be manually opened in the same way. **The model equation is at the top (row 6), the ANOVA table is located just below the coefficient table, the correlation matrix of coefficients is just below the error-distribution statistics table, and the residual table is at the very bottom.** Click the "Titles" button on the ribbon to explicitly show their locations on the model worksheet.
15. **CODING OF MISSING VALUES.** RegressIt will ignore cells on the data worksheet that are empty or which contain only text, treating them as missing values. If an entire column of data consists of empty cells and/or text strings, it will not show up on variable lists except for purposes of creating dummy variables. (For example, date variables and categorical variables are sometimes coded as text strings rather than numeric values.) When an analysis is run, only those rows of data where no values of any variables are missing will be used. If your variables contain any missing values, you should check the "#Fitted" statistic at the top of the output worksheet to see how many rows were used for a given descriptive analysis or regression model. Sometimes you will get very different results when adding or removing one variable (even for descriptive analysis) and this could be due to the variable having a lot of missing values or perhaps a few crazy values that are data errors. Because text strings are treated as missing data, it is OK for missing values to be coded as "---" or "N/A" or strings of space characters or some other convention **except** when the option to save the predictions and residuals to the data worksheet is used. In that case missing values must be coded as empty cells, or else the saved data may not be placed in the correct rows. To be on the safe side it is good to use find-and-replace to convert missing value codes to empty cells at the start of the analysis. Be sure that the cells are actually empty, not containing space characters.
16. **R-SQUARED EQUAL TO 1.** If you unexpectedly get an R-squared value almost exactly equal to 1, that could be a bad thing (circular logic), not a good thing. Your dependent variable may *by definition* be a perfect linear function of some of the independent variables (e.g., a sum or average), except perhaps for round-off error. This comes up more often than you might think if you have a large number of related variables and you are on a data-mining expedition in which you throw a lot of them into the model all at once to see what will happen. If R-squared turns out to be 1, don't rush to congratulate

yourself. Look carefully at the definitions of your variables and try to determine whether you may have just predicted the dependent variable from itself.

17. **MULTICOLLINEARITY ERROR OR VERY LARGE VARIANCE INFLATION FACTORS.** If one of your *independent* variables is a perfect linear function of others, or if a subset of variables adds up to a column of 1's, the model will not run and you will get a pop-up message saying that "multicollinearity" has been detected among a subset of variables and "the <...> variable is part of this group." *Remove the indicated variable and try again.* This may happen, for example, when your model includes dummy variables for a set of mutually exclusive events such as experimental treatments or months of the year and you fail to leave one out so as to avoid redundancy with the constant. In some situations there may be multicollinearity but the error message is not triggered, perhaps due to imprecision in the numbers. A quantitative measure of the extent of multicollinearity is provided by *variance inflation factors* (VIF's). The VIF of an independent variable is equal to 1 divided by 1 minus R-squared in a regression of itself on the others. Thus, a VIF greater than 10 [resp. 20, 100, 1000] indicates that an R-squared of greater than 90% [resp. 95%, 99%, 99.9%] would be obtained in such a regression. If you see variables with VIF's of this size or even much larger, you may want to consider removing one of them (or several in sequence) to see if an equally accurate and easier-to-understand model is obtained. When large VIF's are found, it is hard if not impossible to attach any meaning to the signs and magnitudes of individual coefficients.
18. **MOST EFFICIENT GRAPH FORMATS.** On a PC, you should use high-res rather than low-res picture format for graphs with small sample sizes up to a few hundred. High-res graphs remain sharp under enlargement but their memory usage increases with sample size because they contain the data that is plotted. Low-res graphs are bitmaps that use a fixed amount of memory and are more efficient for larger samples. (See page 16 in the [user manual](#) for examples of file sizes for different graph formats and sample sizes.) Editable graphs should be used when there is a need for editing or the output includes graphs which are interactive, particularly for logistic regression models. (As noted above, it is necessary to close the File Explorer on a PC if either the low-res or high-res option is used.) On a Mac there are only 2 graph format options: picture (which is the same as hi-res picture on a PC but takes up twice as much storage space) and editable. When running many models on a Mac with large sample sizes and lots of graphs, it is a good idea to monitor memory usage. If it gets too large, restarting Excel and reloading the last saved version of the file will generally free up memory. It is also OK to delete the worksheets of models that are no longer needed in order to conserve memory.
19. **MODEL SIZE LIMITATIONS.** On a PC a linear regression model may contain up to 200 variables and up to 1M rows of data subject to a practical limit of about 5M cells in the matrix of independent variables. (On a Mac those limits are 125 variables and 120K rows.) As those limits are approached, running times could be slow—on the order of minutes—and it may be best to not generate graphs. If graphs are generated for very large sample sizes, only the first 32K points will be plotted, and low-res picture format should be used. In such cases you may want to sort the rows in random order to get a representative sample in the plots. The logistic regression procedure is limited to about 16K rows of data. RegressIt's [menu interface for generating R code](#) can be used to run linear and logistic models with larger numbers of variables and larger sample sizes in much less time. This feature only requires pushing a few buttons, not learning the R language or writing any code of your own, and it does not require the full data set to reside in Excel, so it is not limited by the 1M row constraint. All that's needed is to have range names in Excel that match the variable names in the R environment, so that they will appear on variable-selection lists in RegressIt's dialog boxes. There is an option to export the model output from R back to Excel, and it shows up on a worksheet looking exactly as if RegressIt had produced it and having the same interactive features in terms of variable selection and color coding of cells. It will also be added to the model summaries worksheet.
20. **RUNNING TIMES.** RegressIt is written in Visual Basic, which is inherently slower to execute than programs with native compiled code, although this is usually more than compensated for by the fact that its output is very detailed and is designed for efficient reading and navigation and

presentation. Most descriptive analyses and regression analyses with small to medium data sets will take somewhere between 5 seconds and 30 seconds to run. Longer running times are usually due to the combination of a large input data matrix and/or a small amount of working memory and/or voluminous chart output and/or competition with other programs running at the same time. The normal quantile plot is relatively time-consuming to create if the sample size is very large (which is why it is not included in the standard chart set), and a scatterplot matrix is time-consuming to create if the sample size and number of variables are large. (On a PC it is possible to create a 50x50 scatterplot matrix consisting of 2500 separate Excel charts on a single worksheet, but this may take a while. Note that there is an option to produce only those scatterplots with a selected variable on the X or Y axis, for example, the dependent variable for a regression model or a common independent variable for many simple regressions.) When you run the descriptive analysis or linear regression procedures, you will see progress messages flash by in the lower left of the Excel window. On a Mac, there will be a brief delay and the colored wheel may spin for 5 or 10 seconds before you see any messages--be patient. If your data set is very large, there may be longer delays on both Macs and PC's before you see any progress messages--be patient. (Sometimes the Excel window will temporarily turn gray while a large input data matrix is being organized and checked. This does not necessarily mean a crash has occurred.) Occasionally an analysis that normally runs quickly will take a long time (say, a minute or more) to run. This could be due to a workspace-cleanup operation going on--be patient. If nothing has happened for a very long time (say, a couple of minutes) and the requirements of your analysis are not huge, it is possible that a crash has occurred, in which case you may need to use the Task Manager on a PC or the Activity Monitor on a Mac to close Excel. As noted above, for very large datasets you can run linear and logistic regression models very quickly in R by using RegressIt's [R interface](#).

21. **SOFTWARE CONFLICTS.** It is possible for other software to interfere with the running of analyses in RegressIt, particularly other software that actively uses the clipboard or spontaneously interrupts you. In particular, SPSS will block the operation of RegressIt, and it is possible that some social media applications could interrupt it. If you encounter technical problems that do not have simple fixes as described above, you should re-boot the computer and start a new RegressIt session with as few other programs running as is possible, just to be sure that a software conflict is not the cause.
22. **RECOVERING FROM CRASHES.** If an analysis crashes for any reason, this may or may not damage the file to such an extent that further analyses cannot be run in it. A crash has occurred if you get an error message and/or if the output worksheet is incomplete or has visible gridlines and/or if a "temp" (temporary) worksheet called "__running__" appears as the first sheet in the file (to the left of the data sheet). You should also go to the name manager (Formulas/Name Manager on a PC or Insert/Name/Apply on a Mac) and look for temp variables that shouldn't be there. These will have names that begin with text strings such as "__running__" or "__adj__". If you see the temp worksheet and/or temp variables and/or the output worksheet shows gridlines or other problems, you must delete them all before trying to continue using that file. Then save the file under a new name, close it, and re-open it, and try to resume your analysis. Another symptom of file damage is an error message saying "That name is already taken" when trying to run another model. (This condition is not fatal, but it means you will have to manually update the model name when running a new model, because a new default name is no longer being generated automatically.) If you continue to have problems, you should back up to an earlier saved version if you have one, and it is also good to restart Excel. (It is recommended that you save your files under new names periodically while working in order to have backup points.) If the damaged file included some models that were not backed up, it is not hard to re-run them in a new file, and this does not require remembering or re-selecting the variables. First, copy and paste the contents of the data worksheet in value format to the first worksheet in a new file, and use the Select Data and Create Names tools to re-assign variable names. (It is **very** important to use value format when pasting in this situation.) Then go to the model summaries worksheet in the damaged file and hit the Select Data button, which will automatically select the range that contains all the model information. Then copy and paste it in value format onto another new worksheet in the

new file and give it a name such as "old models". (Here too, it is very important to use value format when pasting.) You should notice that the cells directly above the model names contain long text strings that encode the variable lists and other descriptive information about the models. (This information is normally hidden in white font.) If you position the cursor on one of these cells and hit the Linear Regression button, the model name and variables from that text string will be pre-selected as the defaults for a new model. Just choose your desired output options and hit Run in order to reproduce the original model's results. A new model summaries worksheet will be created when the first model is run.