

Assignment 3: The Perils of Regression

To complete assignment 3, you will need to follow the steps below. It is recommended that you read through each step before starting the assignment. I apologize in advance for the inevitable typos.

Step 1: Download the Excel File from Blackboard

On the course blackboard site, in the section called “Assignments”, you will find a folder titled “Multiple Regression”. Inside that folder will be an Excel file titled “Data”. Download that spreadsheet and open it.

Step 2: Multiple Regression

Oh my! This has escalated quickly. Now the data table has five variables, and then three more columns of the regression predictions. Here is the breakdown:

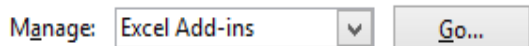
- » **GPA** is a student’s self-reported Grade Point Average.
- » **Age** is a student’s self-reported age in years.
- » **Height** is a student’s height in inches.
- » **Eye Color** is a variable which indicates self-reported eye color. The table all the way on the right contains information relating these numbers to eye color.
- » **Gender** is equal to 0 for male and 1 for female.
- » **GPA-hat** is the predicted GPA based on the regression coefficients in the “Your Model” table.
- » **e** is the set of residuals for this regression.
- » **e²** is the residual squared.

What you are going to do is estimate a *multiple* regression, where you are predicting GPA with several other variables, in this case age, height, eye color, and gender. Of course, to do this, you will need to use the advanced capabilities of Excel, and that means I need to try and show you how to do it with pictures. Unfortunately, I can only do this with the Windows version of Excel. Godspeed to those with a Mac.

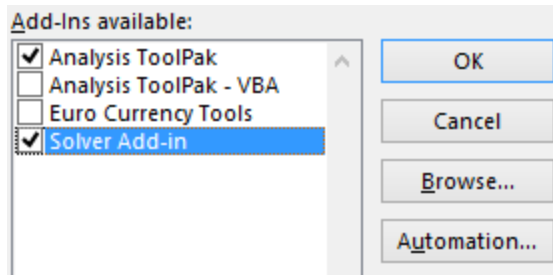
But first, an explanation. Linear Regression by Ordinary Least Squares finds the best possible coefficients for each variable by minimizing the SSE. It is will never get the SSE down to 0, but it will try to get it as close as possible. Excel’s *Solver* is capable of solving this problem by brute force. It will make millions of guesses for the beta coefficients in a split second, and stop only when it’s sure it found the minimum of the SSE. Here is how you do that.

Excel Solver

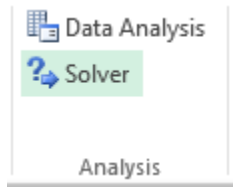
1. Enable Solver by going to File -> Options -> Add-Ins -> Go...



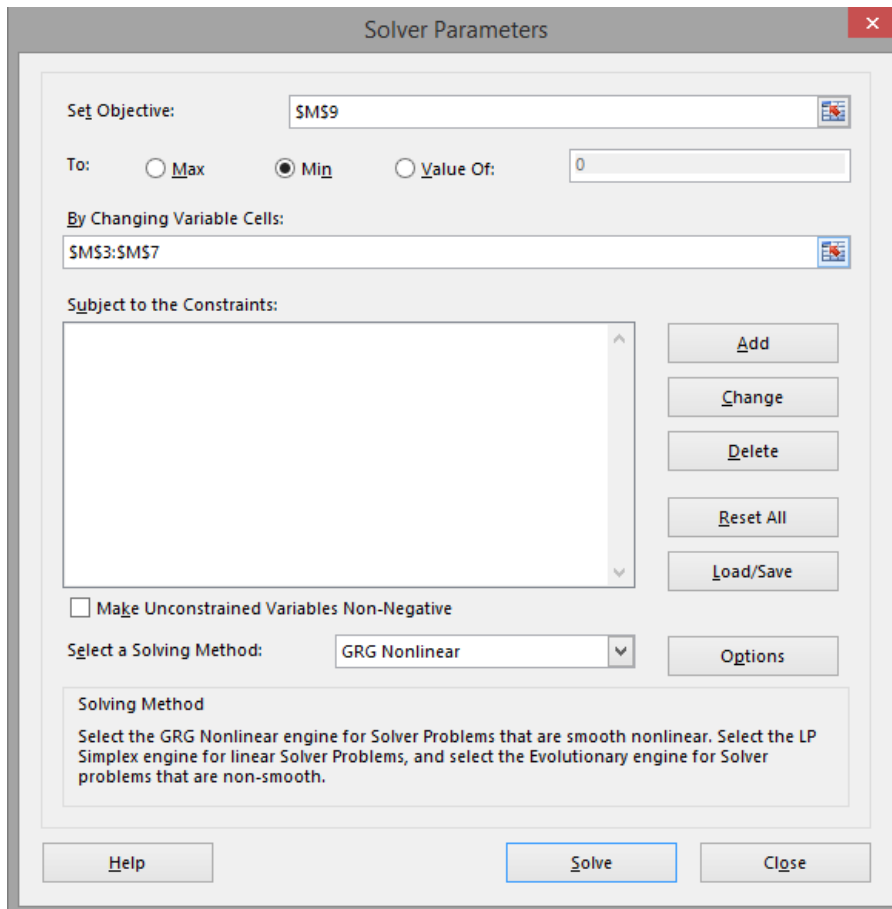
2. Make sure “Solver Add-in” is checked like so:



- Go back to the "Multiple Regression" spreadsheet and click over to the "Data" tab on the top ribbon. All the way on the right you will see the "Analysis" section, where you should now have the option of starting up "Solver".



- Click Solver and fill in the appropriate information.



- The object you are setting is to have cell M9, which is the SSE, minimized (which is selected in the "To:" section below that field).

6. It will minimize this value by changing the values in M3:M7, which are our beta coefficients that make up the linear regression.
7. Make sure the solving method is set to “GRG Nonlinear” and click “Solve”.
8. A “Solver Results” window will pop up. Just click “Ok”.
9. Take a look at the results now found in cells M3 to M7. These are your beta coefficients.

Once you have done this (and hopefully comprehended it), you will need to find the SSR and R^2 . SSR is easy, as I have already given you SST and SSE is being calculated by the Solver. For R^2 , you will need to use the following formula:

$$R^2 = \frac{SSR}{SST}$$

Step 3: Upload your Excel File through Blackboard

Save your completed Excel file and upload it to Blackboard using the “Assignment 9 Submit” link found in the “Assignment 9” folder. In the submission form, you will need to attach your spreadsheet and click submit. Students who fail to submit their Excel file will have their quiz score reduced to a 0.

Step 3: You’re Done!