

MATH 124 – Find Regression Coefficients (and a lot more) using the Data Analysis Toolpak

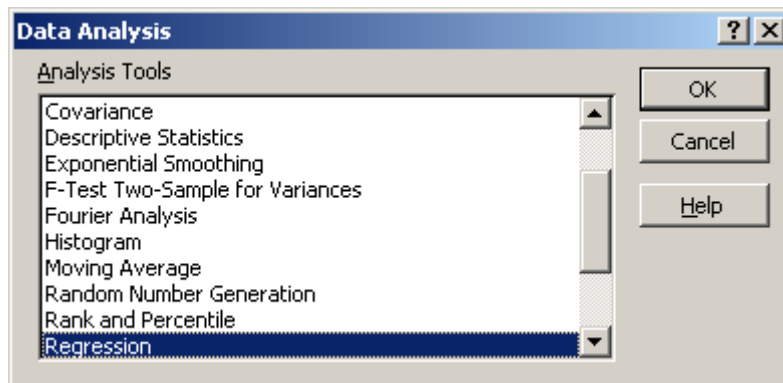
bolstad_math124@bmbolstad.com

<http://math124sfsu.bmbolstad.com>

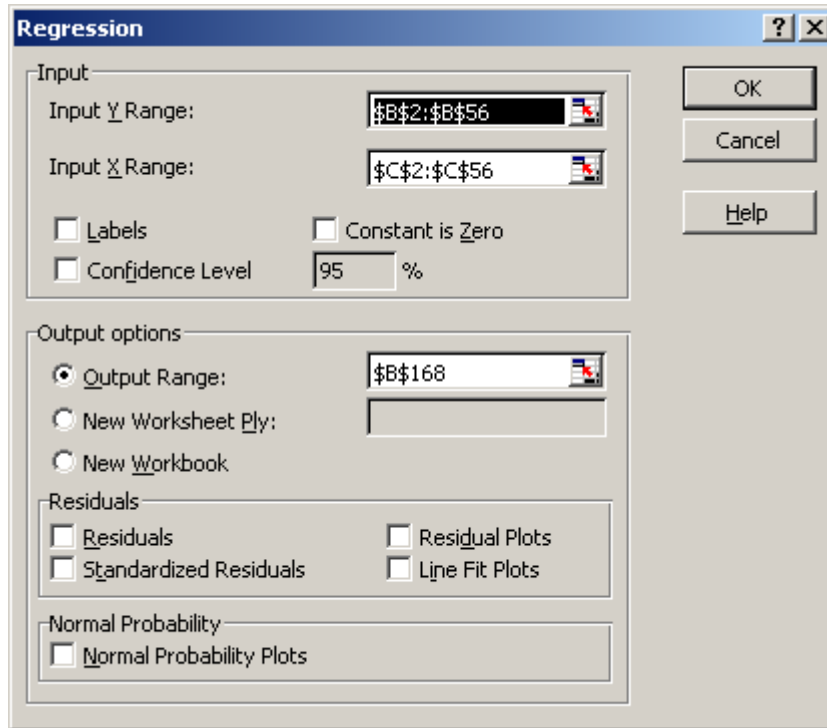
This document explains how to use Data Analysis toolpak to fit a simple linear regression. It is expected that you have already read the “Transform” and both the other two “Find Regression Coefficient” documents before you reached this point (if not now would be a good time to do so). We will again concentrate on the first dataset in the transform data file. A completely worked spreadsheet is available on the website.

Using the data analysis toolpak

1. It was explained in the Histogram document how to install the Analysis Toolpak. If you do not have it installed do so now.
2. Now it is time to fit the regression model. Go to the Tools menu and select the Data Analysis Option. Then choose “Regression” from the list and click ok.



3. Fill in the fields in the dialog box in a similar manner to the picture below. In particular the Input Y Range should be the set of cells containing the response variable (B2:B56), the Input X range should be the set of cells containing the explanatory variable and we will place the output back into the original spreadsheet at any empty location (B168). Make sure you have the Output Range radio button clicked. Click ok to finish.



4. You should now find output that looks something like the following appears in your spreadsheet. There are certainly a lot of numbers and in class we have only (and will only) discussed a subset of them.

SUMMARY OUTPUT

| <i>Regression Statistics</i> | |
|------------------------------|----------|
| Multiple R | 0.986553 |
| R Square | 0.973287 |
| Adjusted R Square | 0.972783 |
| Standard Error | 0.373172 |
| Observations | 55 |

ANOVA

| | <i>df</i> | <i>SS</i> | <i>MS</i> | <i>F</i> | <i>Significance F</i> |
|------------|-----------|-----------|-----------|----------|-----------------------|
| Regression | 1 | 268.9171 | 268.9171 | 1931.081 | 2.25E-43 |
| Residual | 53 | 7.380636 | 0.139257 | | |
| Total | 54 | 276.2978 | | | |

| | <i>Coefficients</i> | <i>Standard Error</i> | <i>t Stat</i> | <i>P-value</i> | <i>Lower 95%</i> | <i>Upper 95%</i> |
|--------------|---------------------|-----------------------|---------------|----------------|------------------|------------------|
| Intercept | 2.999511 | 0.378339 | 7.928098 | 1.46E-10 | 2.240659 | 3.758363 |
| X Variable 1 | 10.62795 | 0.241852 | 43.94407 | 2.25E-43 | 10.14285 | 11.11304 |

5. The most interesting things for us are the R-square value, the Standard Error value, the estimates of the intercept and slope, their respective standard errors, t

statistics and corresponding P-values (these may be discussed in some of the lectures). We will not discuss the remaining items.

6. The excel spreadsheet online has Data Analysis Toolpak regression output for the other datasets.