A Quick Guide to Using Excel 2007’s Regression Analysis Tool
Table of Contents

The Analysis Toolpak .............................................................................................................. 1
If the Add-In is ALREADY Installed ..................................................................................... 1
If the Add-In is NOT ALREADY Installed ........................................................................... 1

Run a Multiple Regression: An Example
   Introduction.................................................................................................................. 2
   Steps to Follow ............................................................................................................. 2
    1. Rearrange the Data as Necessary .......................................................... 2
    2. Open the Regression Analysis Tool ..................................................... 3
    3. Complete the “Regression” Dialog Box ............................................. 4
    4. Run the Regression Analysis and View the Results ......................... 5
The Regression Results Areas.................................................................................... 6
Run Another Regression Analysis ............................................................................ 6

9-11-2009
Using Excel 2007’s Regression Analysis Tool

The Analysis Toolpak
Excel 2007 has a built-in regression analysis tool that’s packaged as part of its “Analysis Toolpak”. The Analysis Toolpak is a standard component of Excel. Microsoft makes it available as an Excel add-in. If you loaded your copy of Excel on your computer, you may recall that you had a choice during the installation process whether or not to include add-ins such as the Analysis Toolpak.

If the Add-In is ALREADY Installed
If the Analysis Toolpak add-in is already installed, you’ll see an “Analysis” group in the “Data” tab of Excel’s ribbon. Click the “Data Analysis” button in that group to open the “Data Analysis” dialog. Then scroll down to find the “Regression” option.

The Analysis Toolpak add-in is installed in Excel 2007 on all FuquaNet MBA PCs.

If the Add-In is NOT ALREADY Installed
If you’re using your own computer and you do not yet have the add-in installed, follow these steps to install it:

1. Close Excel if you have it open.
2. Make your installation medium (probably a CD) available to your computer.
3. Restart the Excel installation routine. (Note that you need not reinstall Excel entirely.)
4. Find the option to modify the installation and choose the Analysis Toolpak as an add-in that should be activated.
5. Exit the installation process.
6. Open Excel and confirm that the “Data Analysis” dialog is available (see above).
Run a Multiple Regression: An Example

Introduction
This example walks you through how to use Excel 2007’s built-in regression tool to analyze whether information collected by the Mrs. Smyth’s Gourmet Frozen Fruit Pie Company (price, advertising, competitors’ pricing, etc.) can predict company pie sales. The worksheet with pie data is partially shown in the illustration below. There are ten columns of data with forty-eight records.

For this example regression operation:

We use as the predictor variables (the independent or X variables) the data for Price, Ads, Comp_Price, Income, Population, and Time Var.

We use as the predicted variable (the dependent or Y variable) the data for Unit Sales. Unit Sales holds actual sales data.

Steps to Follow When Running a Regression with Excel

1. Rearrange the Data as Necessary
With Excel’s regression tool the independent X variables you use in your analysis must be located together in the worksheet. There must be no blank columns or columns with non-relevant data interrupting the range of X variables. The dependent Y variable need not be located adjacent to the X variables, but all Y variable values must be in a single range. These location specifications are required by Excel’s built-in regression tool.
For the data in our sample, the data need not be rearranged. The X variables are together in Columns F through K and the Y variable values are all in Column E.

<table>
<thead>
<tr>
<th>City</th>
<th>Year/Str</th>
<th>Row</th>
<th>Unit Sales</th>
<th>Price</th>
<th>(S) Ads</th>
<th>(S) Comp. Price</th>
<th>(S) Income</th>
<th>Population</th>
<th>Time Var</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlanta, GA</td>
<td>2003-1</td>
<td>1</td>
<td>193,334</td>
<td>6.39</td>
<td>15,827</td>
<td>6.92</td>
<td>33,337</td>
<td>4,116,260</td>
<td>1</td>
</tr>
<tr>
<td>2003-2</td>
<td>2</td>
<td>170,041</td>
<td>7.21</td>
<td>20,819</td>
<td>4.84</td>
<td>33,390</td>
<td>4,140,338</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2003-3</td>
<td>3</td>
<td>247,709</td>
<td>5.75</td>
<td>14,922</td>
<td>5.26</td>
<td>33,699</td>
<td>4,218,965</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2003-4</td>
<td>4</td>
<td>283,259</td>
<td>6.75</td>
<td>16,973</td>
<td>6.17</td>
<td>33,797</td>
<td>4,229,070</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>2004-1</td>
<td>5</td>
<td>297,114</td>
<td>6.36</td>
<td>18,815</td>
<td>6.36</td>
<td>33,915</td>
<td>4,279,812</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

2. Open the Regression Analysis Tool
In the “Data” tab on Excel’s ribbon find the “Analysis” group and click the “Data Analysis” button.

Excel opens its “Data Analysis” dialog. Scroll down and find the “Regression” option. Click OK.
Excel 2007’s “Regression” dialog displays. This is the sole interface to Excel’s regression tool. All your interaction with Excel’s regression tool happens in this dialog.

3. Complete the “Regression” Dialog Box

- Specify “Unit Sales” as the “Input Y Range”. Include the column header cell (the cell that holds the text “Unit Sales”) in your range specification.

- Specify the range of data in Columns F through K as the “Input X Range”. Include only one row of column headers with your range specification (that is, include only the labels directly above the data).

- In the “Regression” dialog box make sure the Labels box is checked. With this option checked Excel will recognize that the first cell in each column of data is a data label.

1 In this selection operation be careful to select data and header only and not an entire column. Here’s an easy way to select just the contiguous data from a column of data: Click the top cell. Depress the SHIFT key. Tap the End key and then tap the down arrow key. Excel selects all data down the column until it reaches the first empty cell in the column.
- Direct Excel to put its results on a new worksheet by choosing the option New Worksheet Ply.

- For outputs, select Residuals, and Standardized Residuals.

With data ranges and choices specified, your dialog should now look like the one below.

4. Run the Regression Analysis and View the Results
With the dialog box options completed as described above, click OK to execute the regression.

The New Worksheet Ply default output option means that Excel locates the regression outputs on a new, separate worksheet that it creates and puts to the left of the worksheet that holds the original data.

Excel completes the regression analysis leaving a large range of the worksheet selected. Click any cell to turn off the range selection. Because the Regression tool generates many outputs on a single sheet, you may want to use Excel’s Zoom Control option to reduce the new worksheet size in order to get a view of how the outputs are arranged.
Widen the columns on the new worksheet in order to see all the text results properly. A quick way to do this is to select the blank square at the intersection of the row numbers and column headers (to select the entire worksheet), and then double-click the dividing line between any two column headers. Excel widens each column so the widest entry in that column displays completely.

**The Regression Results Areas**

Selecting the *Residuals*, and *Standardized Residuals* options in the “Regression” dialog produces the output partially shown below.

In this view columns have been widened to show the results completely.

![Image of regression output](Image)

Excel labels the output areas “SUMMARY OUTPUT”, “ANOVA”, and “RESIDUAL OUTPUT”.

**Run Another Regression Analysis**

Perhaps after viewing the results of the first analysis you decide you’d like to run another regression analysis using different variables or a different combination of variables. To do so, go through the same steps as above. Excel adds a new worksheet to hold the new results. The original results (from the first analysis) remain in place.

End