PERFORMING REGRESSION ANALYSIS USING MICROSOFT EXCEL

John O. Mason, Ph.D., CPA
Professor of Accountancy
Culverhouse School of Accountancy
The University of Alabama

Abstract:

This paper introduces you to the Excel regression tool. It shows you how this tool can be used to perform linear regression analysis (one independent variable) as well as multiple-regression analysis (two or more independent variables up to a maximum of 16). It also shows you how to interpret Excel's regression output.

For regression analysis involving only one independent variable, the paper shows you how to select the Line Fit Plots option so that Excel will create an XY (Scatter) chart at the time that Excel computes and writes the regression output to the workbook. Moreover, you will learn how to modify the XY (Scatter) chart so that it communicates better.

KEYWORDS: Linear regression, multiple regression, regression chart, intercept, coefficient of determination, coefficient of X in linear regression analysis, coefficients of X_1 and X_2 in multiple regression analysis, and residuals

REGRESSION ANALYSIS

Excel's regression analysis tool performs linear regression analysis, which fits a line through a set of observations using the "least squares" method. Regression is used in a wide variety of applications in finance and accounting to analyze how the value of a single dependent variable is affected by the values of one or more independent variables. You can then use the regression results to predict the value of the dependent variable based on values of the independent variable(s).

With Excel, you can perform both regression analysis (one independent variable) and multiple-regression analysis (two or more independent variables up to a maximum of 16). Excel also enables you to choose the Line Fit Plots option to create a Line Fit Plots column chartand the Residuals Plots option to create a residuals column chart. However, because Excel can create only two-dimensional XY charts, only one independent variable and one dependent variable can be represented in a regression chart or a residuals chart. Whenever you select the line fit plots option or the residuals plots option and there are two or more independent variables, Excel will create a single column chart or residuals column chart for twoor more independent variables. However, the XY (Scatter) chart, which is changed from the column chart,will not include the predicted value of the dependent variable with a straight line. The examples that follow include regression analysis using one independent variable, modifying a column chart created by the Line Fits Plot option so

Note: The Data Analysis command displays a Data Analysis dialog box that lists statistical and engineering tools. These tools are part of the Analysis ToolPak add-in that is included with Microsoft Excel. If you do not see the Data Analysis button in the Analysis group of the Data tab's Ribbon, then the Analysis ToolPak add-in was not installed. See the Appendix on installing the Analysis ToolPak.

Data Analysis

that it becomes a an XY (Scatter) chart, and multiple regression with two independent variables.

To perform regression analysis, click on the Data Analysis button in the Analysis group, which is on the far right of the Data tab's Ribbon. If you do not see the Data Analysis button or the Analysis group in the Data tab's Ribbon, it means that, although the Analysis Toolpakadd-in was placed on your hard drive when Microsoft Office 2010 or 2007 was installed, this add-in was not installedfor Microsoft Excel. Thus, you should review Appendix at the back of this paper to see how to install the Analysis Toolpakadd-in.

If you click on the Data Analysis button, you will then see a list of data analysis tools in the Analysis Tools list box of the Data Analysis dialog box. Scroll down the list until you see Regression, as shown in Figure 1. Click on Regression and then click on the OK button. Excel then displays a Regression dialog box, as shown in Figure 2. The required ranges and options in the Regression dialog box are as follows:

- **Input Y Range**. Allows you to enter the range of the dependent variable. The data for the input range must be in a column.
- **Input X Range**. Allows you to enter the range of the independent variable(s). The data for the Input X Range must be in columns. If two or more independent variables are present, they must be located in adjacent columns. Excel orders the independent variables in ascending order from left to right, using 1, 2, 3, and so on for the variable names in the summary output table.
- **Labels**. If the first row in the input range includes labels (i.e., text) that describe the independent variables, check the Labels box. Excel then generates the appropriate row titles for the summary output table. If your input range does not contain labels, then the Labels check box should be clear.
- **Constant is Zero**. Check the Constant Is Zero box only if you want to force the regression line to pass through the origin.
- **Confidence Level**. The default confidence level is 95%. If you want an additional level included in the summary output table, check the Confidence Level box, click on the text box to the right, and enter a confidence level (i.e., 99% or 90%) that you want Excel to apply to the regression in addition to the default 95% confidence level.
- **Output Options**. Excel writes the regression output, in the form of a summary output table, to (1) an output range in the active worksheet, (2) a separate worksheet within the active workbook, or (3) a separate workbook. Choose any one of the three options by clicking on its option box. If you choose the output range option, click on the text box at the right and enter the cell address of the upper-left corner where you want the summary output table to appear. When Excel writes the summary output table, it will consist of nine columns and eighteen or more rows, depending upon the other options selected.
- **Residuals**. Check the Residuals box if you want Excel to write the predicted values of Y and the residuals to a residual output table.
- **Standardized Residuals**. Check the Standardized Residuals box if you want Excel to write the predicted values of Y, the residuals, and the standardized residuals to a residuals output table.
- **Residual Plots**. Check the Residual Plots box if you want Excel to create a column chart for the independent variable(s) versus the residual. If you choose this option, Excel will also write the predicted values of Y and the residuals to a residual output table.

- **Line Fit Plots**. Check the Line Fit Plots box if you want Excel to create a column chart for predicted values versus the observed values of Y in relation to the observed values of X. If you choose this option, Excel will also write the predicted values of Y and the residuals to a residual output table.
- **Normal Probability Plots**. Check the Normal Probability Plots box if you want Microsoft Excel to write the observed values of Y and the related sample percentiles to a probability output table and to generate a chart plotting the observed values of Y and the sample percentiles.
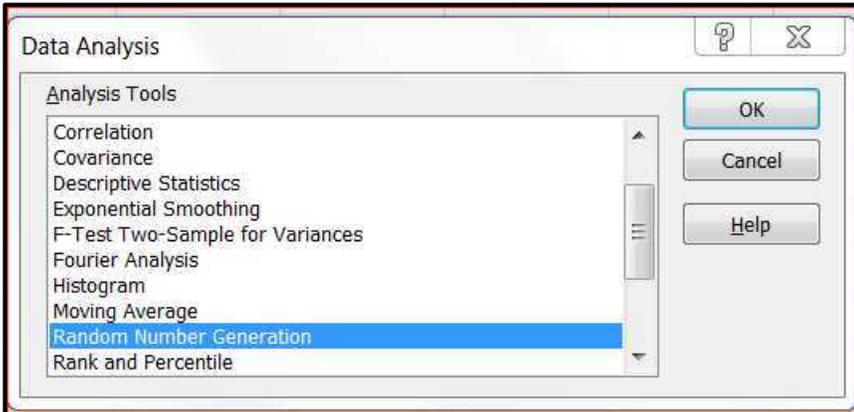

Figure 1
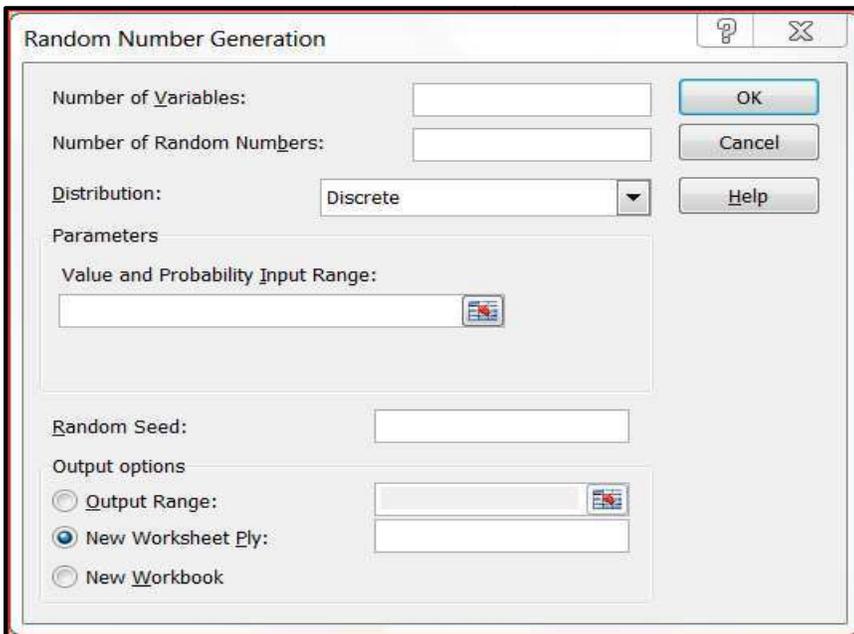*Selecting the Regression Toolin the Data Analysis Dialog Box*

Figure 2
Regression Dialog Box
Linear Regression Analysis

Using Excel

As an example of performing regression analysis with Excel, consider the materials-handling cost example. In this example, you are a new employee in the finance department of Fabuglass Plastics, Inc., a maker of fiberglass products for the automotive industry. A major component of Fabuglass' manufacturing costs is the cost of handling the materials used in the molding process. You were asked by the Vice President of Finance to look at the relationship between the cost of materials and the weight of materials handled each month. As a result, you created a new workbook, entered the cost and weight data for the year 2012 in Sheet1, changed the Sheet1 tab name to **Linear Regression**, and saved the workbook as Fabuglas. The Linear Regression worksheet, which is illustrated in Figure 3, documents the materials-handling costs associated with the weight

194

of materials used for the months January 2012 through December 2012. Each month represents a pair of observations of materials-handling costs and weight of materials, for a total of 12 pairs of observations for the year.

Figure 3

Materials-Handling Costs (Dependent Variable) and

Weight of Materials Used (Independent Variable) for the Year 2012

	A	B	C	D	E	F	G
1		FABUGLASS PRODUCTS, INC.					
2	MONTHLY MATERIALS-HANDLING COSTS						
3	AND WEIGHT OF MATERIALS HANDLED						
4							
5							
6	Year	Month	Weight of Materials	Materials-Handling Costs			
7	2012	Jan	4,608	$ 1,439,250			
8		Feb	3,580	1,227,500			
9		Mar	4,226	1,552,500			
10		Apr	4,838	1,505,000			
11		May	4,220	1,402,500			
12		Jun	5,464	1,710,000			
13		Jul	5,008	1,575,000			
14		Aug	4,830	1,702,500			
15		Sep	4,008	1,497,500			
16		Oct	3,220	1,295,000			
17		Nov	3,648	1,347,500			
18		Dec	3,912	1,423,450			
19							
20							
21							
22							

Linear Regression / Multiple Regression / Sheet3

Now open the workbook FabUGlas and select the Linear Regression worksheet. To perform regression analysis, click on the Data Analysis button in the Analysis group of the Data tab's Ribbon; choose Regression in the Data Analysis dialog box, and then click on OK. Select the following ranges and options in the Regression dialog box:

1. Click on the Input Y Range text box and either select the range D6:D18 with the mouse or enter the range at the keyboard.
2. Click on the Input X Range text box and either select the range C6:C18 with the mouse or enter the range at the keyboard.
3. Check the Labels option box.
4. Check the Confidence Level box, click on the related text box at the right, and then enter 90 (for a confidence level of 90%).
5. If necessary, click on the Output option*New Worksheet Ply:*.If you don't want the output to be in a new worksheet immediately to the left of the *Linear Regression* worksheet, then enter the new worksheet name in the option box to the right of *New Worksheet Ply:*.
6. Check the Line Fit Plots box (Selecting this box also causes Excel to write the predicted values of Y and the residuals to a residual output table below the summary output table. Therefore, if you select the Line Fit Plots box, it is not necessary to check the Residuals options box.).
7. If your Regression dialog box matches the one shown in Figure 4, click on OK to have Excel perform the regression analysis, write the summary and residual output tables to the output range, and create the line fit plots chart.
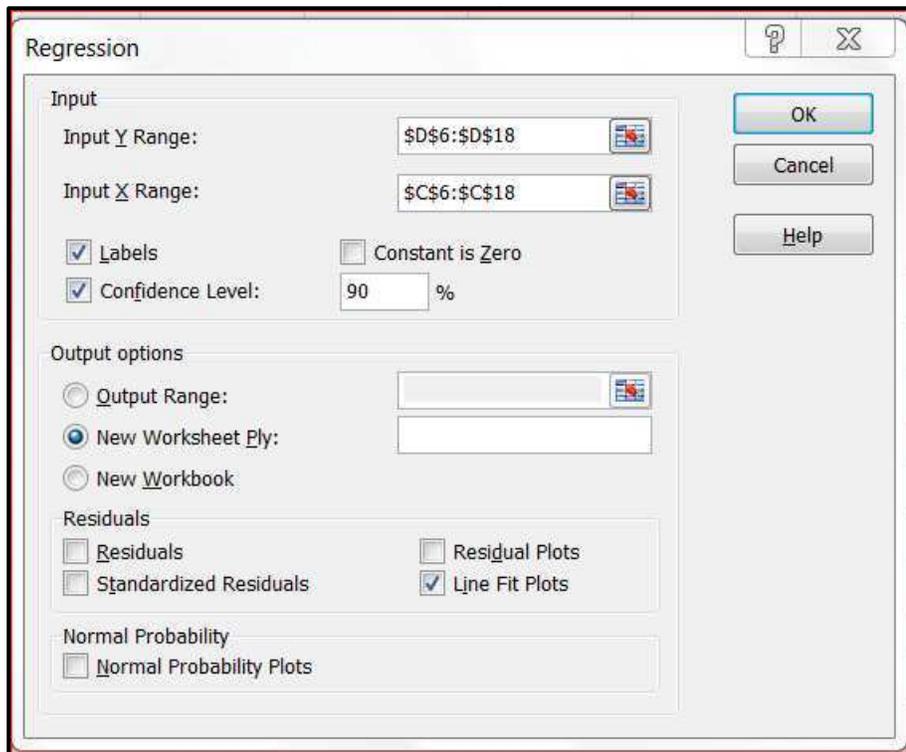

Figure 4
Entering Ranges and Selecting Options in Regression Dialog Box for Fabuglass Regression Example

After performing the regression analysis, Excel writes the summary output table, beginning in cell A1 (the upper-left corner of the output range), to column I (nine columns) and rows 1 through 18; and the residual output table to columns A through C and rows 24 through 36. In order to see all of the information in summary

output table, you will have to adjust the columns widths to fit the row and column titles. Because Excel highlights the columns that contain the summary output table when you click on OK, click on the Format button in the Cells group of the Home tab's Ribbon and then choose the AutoFit Column Width command to adjust the column widths to fit the text and values in the summary output and residual output tables. After adjusting column widths, your summary output and residual output tables should be similar to those which appear in Figure 5. The coefficient of determination (R^2) is in cell N25; the standard error of the Y estimate is in cell N27, the Y-intercept (or constant) is in cell N37; the t-statistic for the Y-intercept is in cell P37, the coefficient of Weight of Materials, the independent variable, is in cell N38, and the t-statistic for the coefficient of Weight of Materials is in cell P38.

The values in the Lower 95% and Upper 95% columns (columns R and S) can be used to develop interval estimates of the Y-intercept and the coefficient of X at the 95% confidence level. Similarly, the values in the Lower 90% and Upper 90% columns (columns T and U) can be used to develop interval estimates at the 90% confidence level. However, these procedures are beyond the scope of this paper.

As you can see in Figure 5, Excel provides quite a bit of useful regression output, including t-statistics for both the Y-intercept and the coefficient of X. Now review the regression output. The R^2 value of 0.74264004 in cell N25 indicates that approximately 75 percent of the change in the dependent variable Materials-Handling Costs is explained by the change in the independent variable Weight of Materials. How do you interpret a t-statistic? A rule of thumb is that, if the absolute value of a t-statistic is greater than or equal to 2, you can accept the related sample coefficient as being significantly different from 0; this indicates that a significant relationship exists between the independent variable and the dependent variable. As you can see in cells P37 and P38, the values of the t-statistic for the Y-intercept and the coefficient of Weight of Materials are approximately 4.1 and 5.4, respectively. Since each t-statistic is greater than 2, you can conclude that a significant relationship exists between the independent variable Weight of Materials and the dependent variable Materials-Handling Costs. Now rename the worksheet with the regression output as Linear Regression Out put

The standard formula for linear regression with one independent variable is:

Y(X) = a + bX

where:

Y(X) = Predicted value of the dependent variable

a = Y-intercept

b = Coefficient of the independent variable

X = Any value of the independent variable Weight of Materials

In the worksheet, the **a** (Y intercept) can be found in cell N37, and the **b** (coefficient of X) can be found in cell N38. Thus, using Excel terminology, the regression formula in this example could be written as:

=N37+N38*X

or as

=645980.1856+192.5*Weight of materials

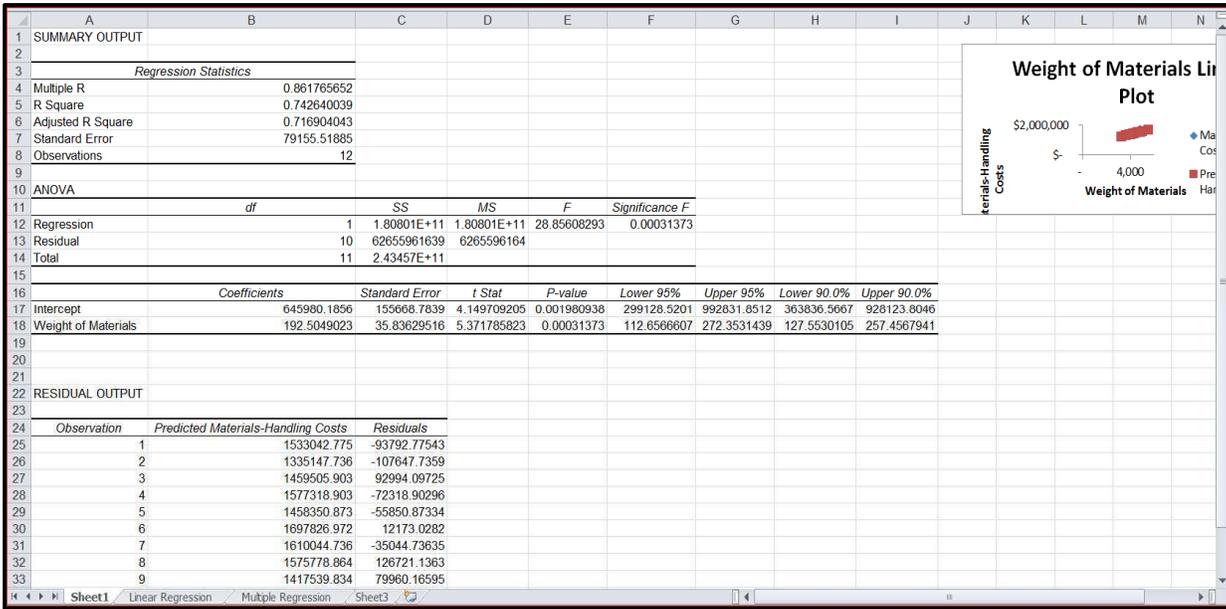|	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3		*Regression Statistics*							
4	Multiple R	0.861765652							
5	R Square	0.742640039							
6	Adjusted R Square	0.716904043							
7	Standard Error	79155.51885							
8	Observations	12							
9									
10	ANOVA								
11		*df*	*SS*	*MS*	*F*	*Significance F*			
12	Regression	1	1.80801E+11	1.80801E+11	28.85608293	0.00031373			
13	Residual	10	62655961639	6265596164					
14	Total	11	2.43457E+11						
15									
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 90.0%*	*Upper 90.0%*
17	Intercept	645980.1856	155668.7839	4.149709205	0.001980938	299128.5201	992831.8512	363836.5667	928123.8046
18	Weight of Materials	192.5049023	35.83629516	5.371785823	0.00031373	112.6566607	272.3531439	127.5530105	257.4567941
19									
20									
21									
22	RESIDUAL OUTPUT								
23									
24	*Observation*	*Predicted Materials-Handling Costs*	*Residuals*						
25	1	1533042.775	-93792.77543						
26	2	1335147.736	-107647.7359						
27	3	1459505.903	92994.09725						
28	4	1577318.903	-72318.90296						
29	5	1458350.873	-55850.87334						
30	6	1697826.972	12173.0282						
31	7	1610044.736	-35044.73635						
32	8	1575778.864	126721.1363						
33	9	1417539.834	79960.16595						

Sheet1 / Linear Regression / Multiple Regression / Sheet3

Figure 5 Regression Summary Output and Residual Output Tables

Whenever you check one of the Residuals options, such as the Line Fits Plots option, in the Regression dialog box, Excel creates a residual output table as part of the regression output. In creating the residual output table, Excel computed a predicted value of Y (i.e., a value on the regression line) for each observation of X and the residual, which is the difference between the observed and predicted values of Y for each observation of X. Rename the worksheet as Linear Regression Output

Modifying the Regression Chart

Because you selected the Line Fit Plots box, Excel created theLine Fit Plots chart to the right of the summary output table. However, in its present form, the chart is difficult to read and interpret. To facilitate changing the chart type and modifying the chart, select the chart by clicking on it. When you select a chart, round handles with holes in them appear along it borders. Next, copy the chart, paste it, and then increase its size as follows:

> TIP: Create a copy of the Line Fit Plots chart and modify the copy. Thus, should you damage the copy when modifying the chart, you can always recopy the original chart and modify the copy all over again.

1. Copy the chart to cell A22of the Linear Regression worksheet as follows:
- Copy the chart to the clipboard (e.g., press Ctrl + C). You should find the upper-left corner of the chart in cell V21.
- Select cell A22of the Linear Regression worksheet and then paste the chart from the clipboard (e.g., press Ctrl + V).

2. Increase the chart's size as follows:

- Position the mouse pointer on the black handle in the lower-right border of the chart until you see the double-headed arrow.
- Press the right mouse button, drag the lower-right border down and to right until the chart extends to the lower-right corner of cell J48, and then release the mouse button.
3. If necessary, click on the Full Screen button in the Worksheet Views group of the View tab's Ribbon in order to display rows 22 to 48 simultaneously on the screen.
   Next, modify the chart by completing the steps below:
- The square data markers (symbols) that appear to form along a diagonal in the chart represent the predicted Y data series. Change the data markers from symbols to a regression line as follows:
⇒ Click on any of the data markers and then choose the Format Data Series command from the shortcut menu. Excel will display a Format Data Series dialog box.

⇒ Click on the Marker Options tab in the dialog box and choose the None option, as shown in Figure 6.

⇒ Click on the Line Color tab and choose the Solid Line option.

⇒ Click on the arrow at the right end of the Color list box to display the color palette and choose the color Dark Red, which is in the left-most color under Standard Colors in the color palette, as illustrated in Figure 7.

⇒ Click on Close to close the dialog box. As a result, Excel will replace the data marker symbols with a dark red regression line.
- Click on the chart title:
⇒ Change the font and font size to Times New Roman and 12 points, respectively.

⇒ Edit the chart title so that it reads as follows: MONTHLY COST OF HANDLING MATERIALS RELATIVE TO WEIGHT OF MATERIALS FOR YEAR 2012. Insert a line immediately before "TO" and move that part of the title beginning with "TO" to the next line by pressing the Enter key.
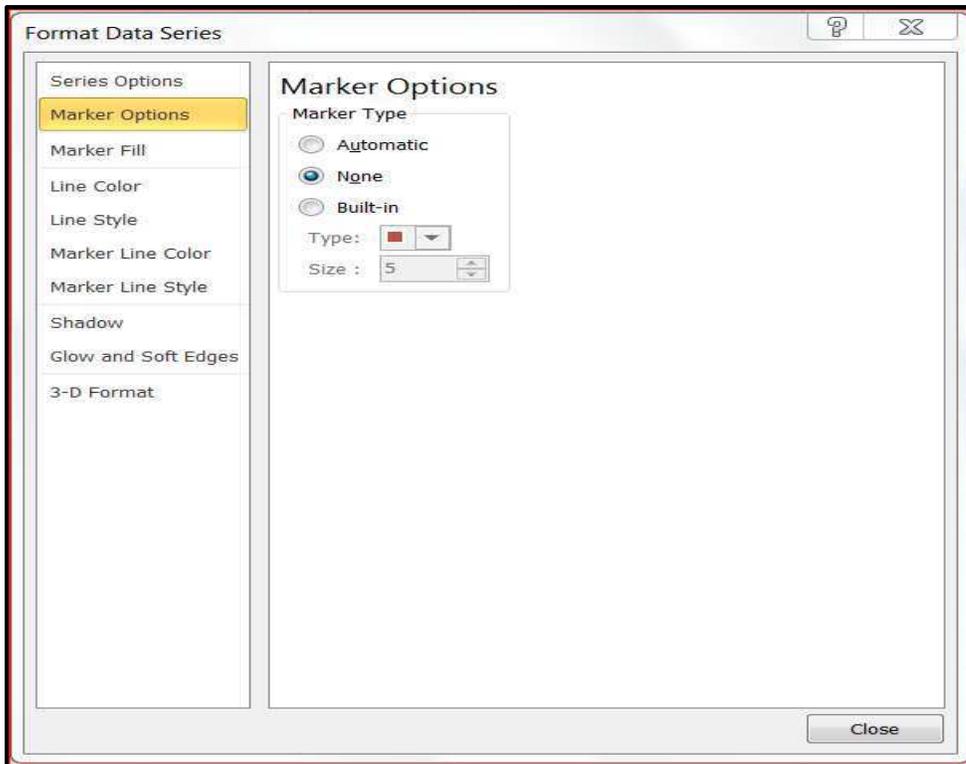
Figure 6
Marker Options Tab of the Format Data Series Dialog Box

- Click on the X-axis title:
- ✓ Change the font and font size to Times New Roman and 10 points, respectively.

- ✓ Edit the title so that it reads Weight of Materials (in pounds). Move "(in pounds)" to the next line by pressing the Enter key.
- Click on the Y-axis title:
- ✓ Change the font and font size to Times New Roman and 10 points, respectively.
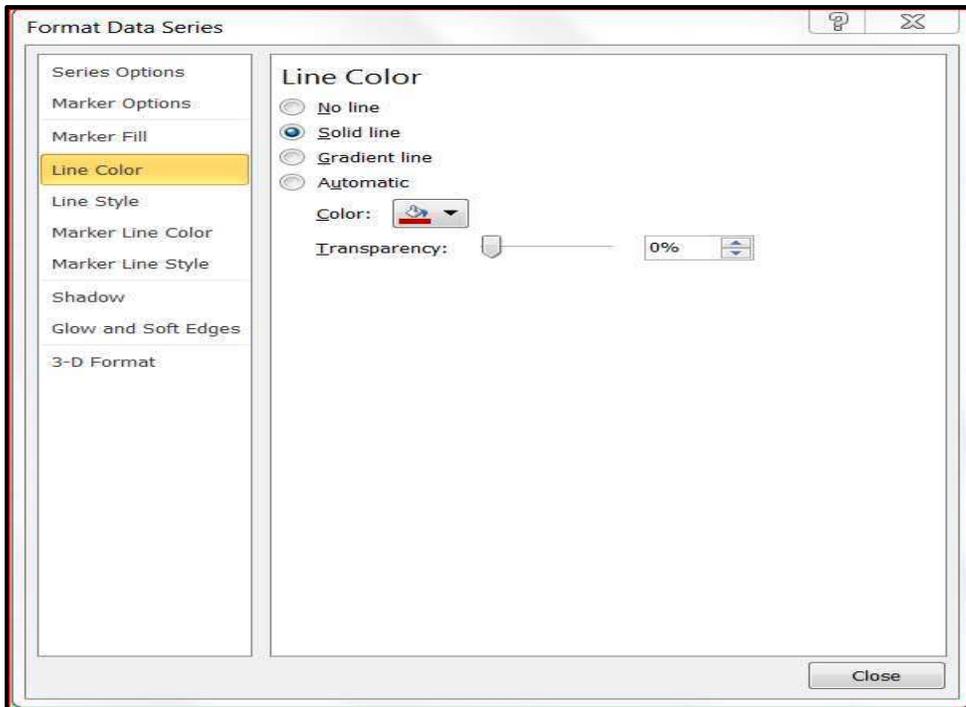
Figure 7
Line Color Tab of the Format Data Series Dialog box

- Click on any of the category names along the X axis:
- ✓ Change the font size of the category names to 9 points.

- ✓ Change the scale of the X axis by choosing the Format Axis command in the shortcut menu to display the Format Axis dialog box, clicking on the Axis Options tab, if necessary, and changing the following options in the Axis Options portion of the dialog box:

- ➢ Check the Fixed option in the Minimum box and then change the minimum value to 3000.

- ➢ Check the Fixed option in the Maximum box and then change the maximum value, if necessary, to 6000.

- ➢ Check the Fixed option in the Major unit box and then change the increment value, if necessary, to 1000.

- ➢ Click on Close to close the dialog box.

- Click on any of the values along the Y axis:
✓ Change the font size of the category names to 9 points.

✓ Change the scale of the X axis by choosing the Format Axis command in the shortcut menu to display the Format Axis dialog box, clicking on the Axis Options tab, if necessary, and changing the following options in the Axis Options portion of the dialog box:

➢ Check the Fixed option in the Minimum box and then change the minimum value to 1000000.
➢ Check the Fixed option in the Maximum box and then, if necessary, change the maximum value to 2000000.
➢ Check the Fixed option in the Major unit box and then, if necessary, change the increment value to 200000.
➢ Click on Close to close the dialog box.
- Click on the Legend:
✓ Change the font size to 9 points.

✓ Right-click on the legend and choose the Format Legend command from the shortcut menu to display the Format Legend dialog box.

➢ In the dialog box, click on the Border Color tab, click on the Solid Line option, click on the arrow on the right of the Color box, and then click on the fourth color from the right under Standard colors in the palette of colors.

➢ Next, click on the Border Styles tab and change the width to 1.25.

➢ After changing the width, click on the Close button to close the dialog box.
  After completing the above steps, view the modified regression chart. It should look like the one shown in Figure 8. If it does not, repeat one or more of the above steps in order to make the necessary changes. In your chart, the individual markers (i.e., symbols) represent the observed values of Materials-Handling Costs. The straight line represents the values along the regression line (i.e., the predicted values of Materials-Handling Costs).
  Since the regression analysis quantified the relationship between materials-handling costs and the weight of the materials handled and indicated that the relationship was significant, you should be able to observe this relationship in the chart. And in fact, the sample observations are tightly clustered around the regression line. Thus, the regression chart reinforces your observations from regression analysis. It is also a lot easier to explain the relationship between the dependent and independent variables to nonstatisticians by presenting the data in a regression chart than by trying to explain to them statistical terms such as R^2, the Y-intercept, the coefficient of X, and the t-statistic.
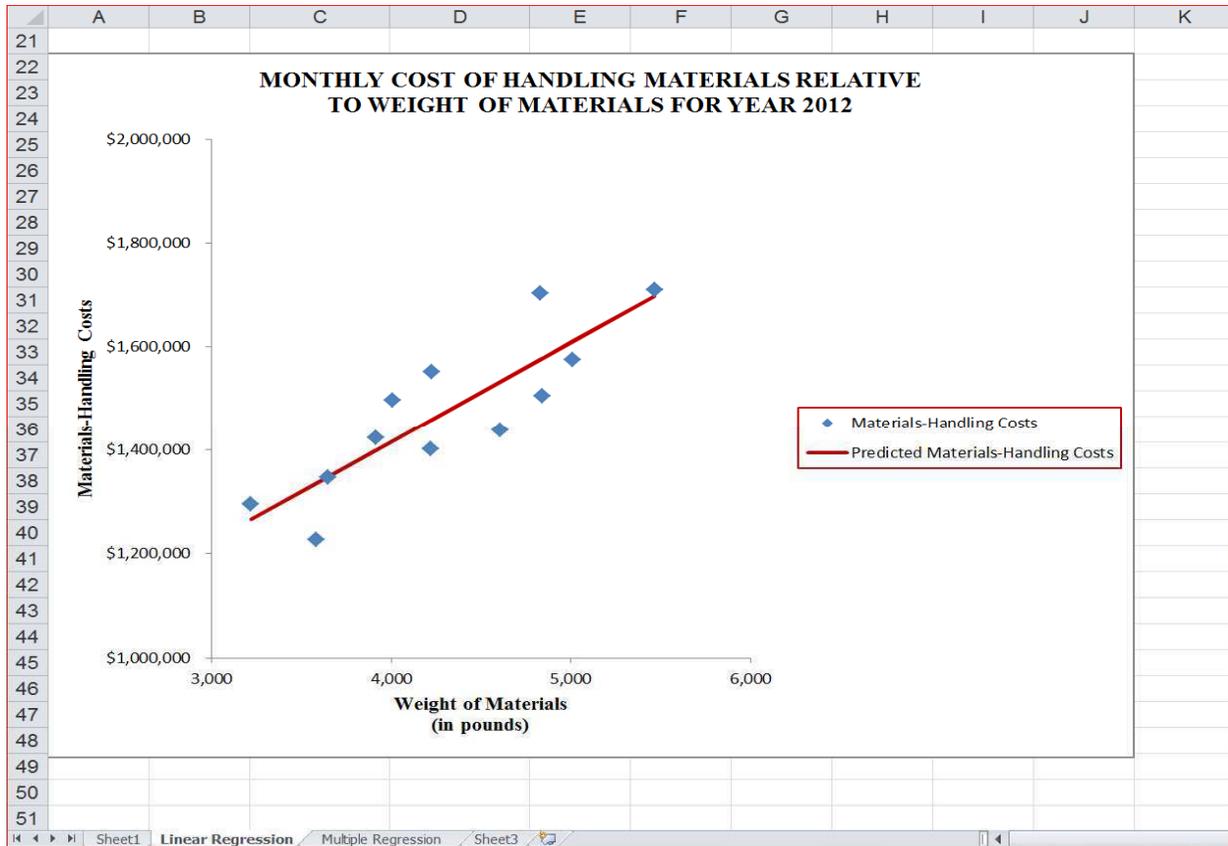
Figure 8
Modified
Regression
Chart

Multiple-Regression Analysis Using Excel

The preceding regression example included only one independent variable. But suppose that the Vice President of Finance met with the Vice President of Production to review the results of your Excel regression analysis and the regression chart. During the meeting, the Vice President of Finance suggested that the dollar value of the materials used might affect the cost of handling the materials, since higher-priced materials must be handled more carefully and consequently take longer to handle. After returning from the meeting, the Vice President of Finance provides you with information on the dollar value of materials used during the twelve-month period ending December 31, 2012. She asks you to add the dollar cost of materials used as a second independent variable in predicting materials-handling cost.

When you conduct multiple-regression analysis using Excel, the observations of the independent variables must be in the same rows and in adjacent columns. Now select the Multiple Regression worksheet in the Fabuglas workbook. It should match that shown in Figure 9. The steps you used in performing linear regression analysis with one independent variable are applicable to multiple-regression analysis, except that here you specify the Input Y Range as E6:E18 and the Input X-Range as C6:D18. Also, the Line Fit Plots box should be cleared if checked. Then check the Residuals box in order to have Excel write the predicted values of Y and the related residuals to the residual output table. If your Regression dialog box matches the one shown in Figure 10, click on OK to have Excel perform the multiple regression analysis and write the summary output and residual output tables.

	A	B	C	D	E	F	G
1		FABUGLASS PRODUCTS, INC.					
2		MONTHLY MATERIALS-HANDLING COSTS					
3		AND WEIGHT OF MATERIALS HANDLED					
4							
5							
6	Year	Month	Weight of Materials	Dollar Value of Materials	Materials-Handling Costs		
7	2012	Jan	4,608	$ 3,031,000	$ 1,439,250		
8		Feb	3,580	1,938,000	1,227,500		
9		Mar	4,226	3,566,000	1,552,500		
10		Apr	4,838	2,784,000	1,505,000		
11		May	4,220	3,001,000	1,402,500		
12		Jun	5,464	3,388,000	1,710,000		
13		Jul	5,008	3,305,000	1,575,000		
14		Aug	4,830	3,717,000	1,702,500		
15		Sep	4,008	3,025,000	1,497,500		
16		Oct	3,220	2,222,000	1,295,000		
17		Nov	3,648	2,817,000	1,347,500		
18		Dec	3,912	2,700,000	1,423,450		
19							
20							
21							
22							

Sheet1 / Linear Regression / **Multiple Regression** / Sheet3

Figure 9
Materials-Handling Costs (Dependent Variable) and Two Independent Variables:
Weight of Materials Used (Independent Variable 1) and Dollar Value of Materials Handled (Independent Variable 2)

Regression

Input
Input Y Range: $E$6:$E$18
Input X Range: $C$6:$D$18

☑ Labels   ☐ Constant is Zero
☑ Confidence Level: 90 %

Output options
☐ Output Range:
◉ New Worksheet Ply:
☐ New Workbook

Residuals
☑ Residuals   ☐ Residual Plots
☐ Standardized Residuals   ☐ Line Fit Plots

Normal Probability
☐ Normal Probability Plots

OK
Cancel
Help
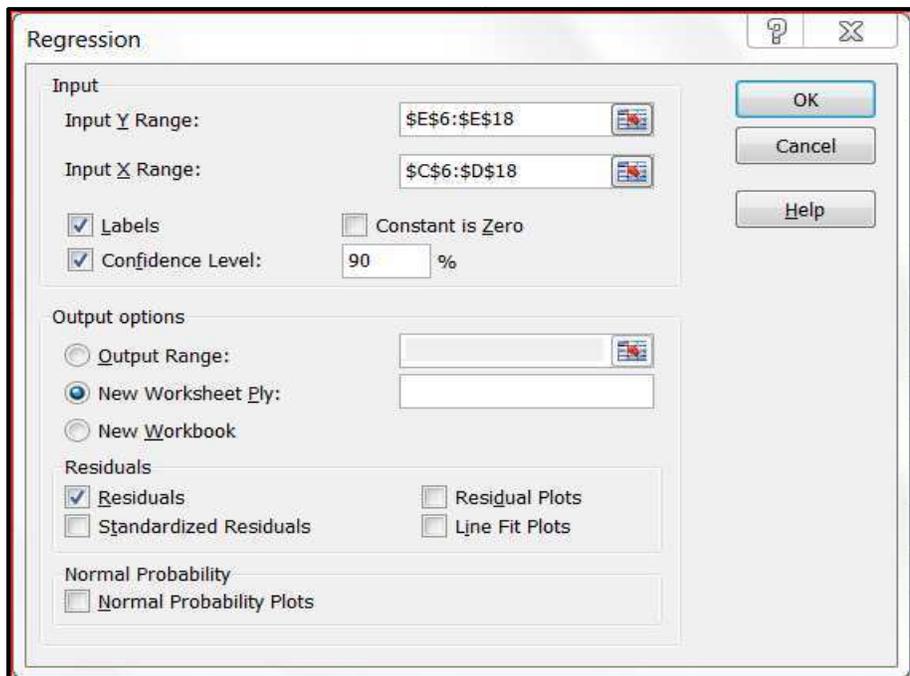
Figure 10
Selected Ranges and Options for Fabuglass Multiple Regression Example

After performing the regression analysis, Excel writes the summary output table, beginning in cell A1 (the upper-left corner of the output range), to column I (nine columns) and rows 1 through 19; and the residual output table to columns A through C and rows 25 through 37. In order to see all of the information in summary output table, you will have to adjust the columns widths to fit the row and column titles. After making those adjustments, your summary output and residual output tables should be similar to those which appear in Figure 11.

Now review the regression output. The coefficient of determination (R^2) is in cell B5; the standard error of the Y estimate is in cell B7; the Y-intercept (or constant) is in cell B17; the t-statistic for the Y-intercept is in cell D17, the coefficient of the first independent variable (i.e., weight of materials) and its related t-statistic are in cells B18 and D18, respectively; and the coefficient of the second independent variable (i.e., dollar value of materials) and its t-statistic are in cells B19 and D19, respectively.

The R^2 value of 0.903132468 indicates that approximately 90.3 percent of the change in the dependent variable is explained by the changes in the two independent variables. Moreover, the values of the t-statistic for the Y-intercept and for both independent variables are greater than 2, indicating that a significant relationship exists between each independent variable and the dependent variable Materials-Handling Costs. Now rename the Sheet 2 worksheet as Multiple Regression Output.

Figure 11 Summary Output and Residual Output Tables for Two Independent Variables

The standard formula for multiple regression with two independent variables is:

$$Y(X_1, X_2) = a + b_1 X_1 + b_2 X_2$$

where:

$Y(X_1, X_2)$ = Predicted value of the dependent variable

a = Y-intercept

b_1 = Coefficient of Weight of Materials (first independent variable)

X_1 = Any value of Weight of Materials (first independent variable)

b_2 = Coefficient of the second independent variable

X_2 = Any value of Dollar Value of Materials (second independent variable)

As noted previously, the Y-intercept can be found in cell B17, the coefficient b_1 of X_1 can be found in cell B18, and the coefficient b_2 of X_2 can be found in cell B19. Thus, using Excel terminology, the regression formula in this example could be written as:

=B17+B18*X_1 +B19*X_2

or as

=554684.79+100.09*Weight of Materials +0.165*Dollar Value of Materials

Because you selected the Residuals box, Excel wrote the residual output, including a predicted value of Y for each observation of Weight of Materials and Dollar Value of Materials, to the range A25 through C37. There are two possible reasons for having Excel compute the predicted values of Y and the residuals: One, you may want to compare the predicted values of Y with the observed values of Y (or review the residuals, which represent the differences between the predicted and the observed values of Y); and, two, at some later point in time, you may want to chart the observed and predicted values of Y relative to one of the independent variables. Moreover, if you have access to a graphics program that can create a three-dimensional (i.e., 3-D) XY chart, you may want to create a multiple-regression chart.

SUMMARY

The regression feature in Excel allows you to perform both linear regression with one independent variable and multiple-regression analysis using the "least squares" method. After you specify the range of the independent and dependent variables and choose options in the Regression dialog box, Excel performs the regression analysis and writes the regression results to a summary output table. Included in the summary output table are R^2, the standard error of the Y estimate, the Y-intercept, the standard error and t-statistic of the Y intercept, the coefficient of each independent variable, and the standard error and t-statistic of each coefficient. Other regression options create a line fit plots chart and write the predicted values of Y and related residuals to a residual output table.

APPENDIX – INSTALLING THE ANALYSIS TOOLPAK

With Microsoft Excel 2010 (similar but different from Excel 2007), you add in the Analysis ToolPak by clicking on the Filecommand on the left-side of the screen and then clicking on Options on the left side and near the bottom of the dialog box. Excel displays the Excel Options dialog box. Click on the Add-ins tab at the left side of the dialog box. Next, click on Analysis Toolpak to highlight it, as shown in Figure 12. Next, make sure that the Manage list box near the bottom of the screen containsExcel Add-Ins, and then click on Go, which appears to the right side of the Manage list box. Excel then displays the Add-Ins dialog box, as shown in Figure 13. Make sure that the Analysis Toolpak has a check in its box and then click OK. You should then see the Data Analysis button in the Analysis group of the Data tab's Ribbon.
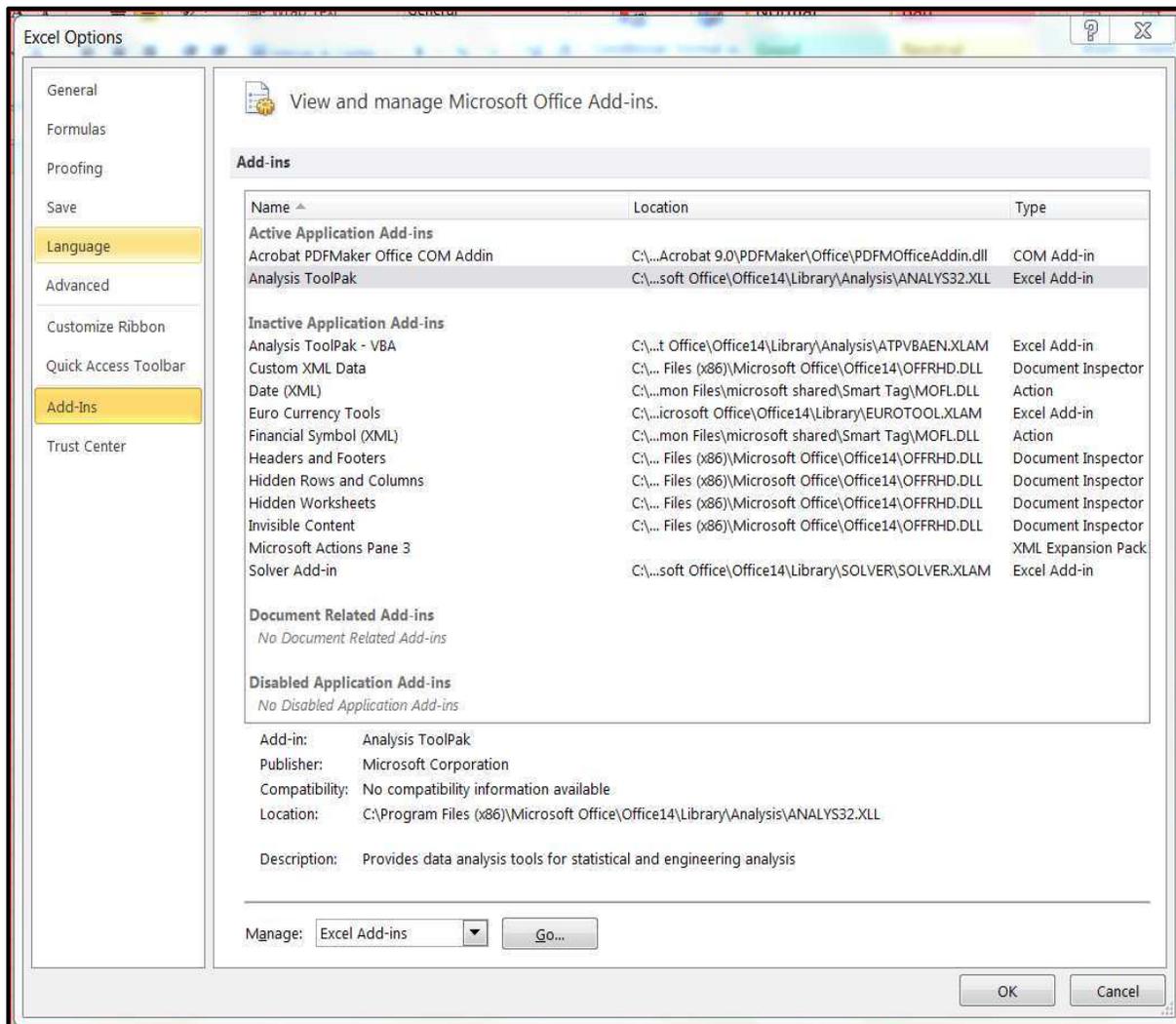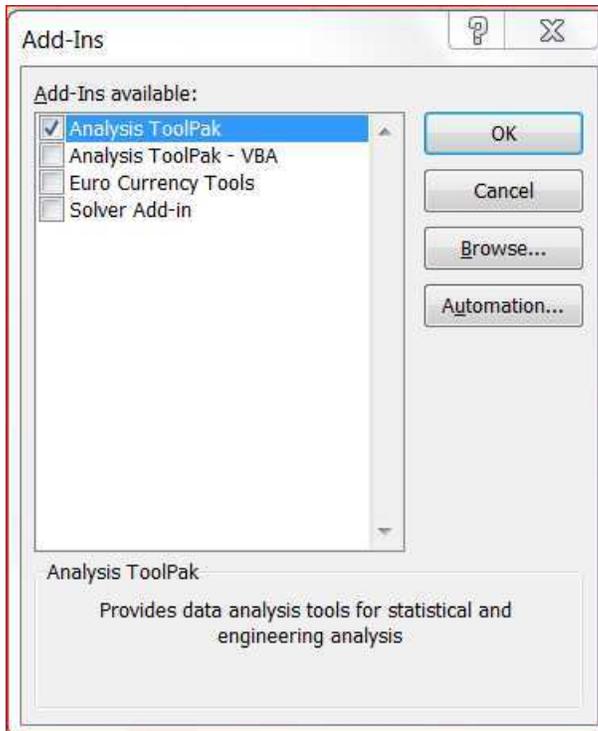

Figure 12
Excel Options Dialog Box

Figure 13
Add-Ins Dialog Box