

UN 0701 Fundamentals of Risk and Reliability



Microsoft Excel 2007 / 2010 Tutorial

Mahesh Pandey

NSERC-UNENE Industrial Research Chair

Risk-Based Life Cycle Management of Engineering Systems

Mikko Jyrkama

Research Associate

January 2012

Dept. of Civil and Environmental Engineering
University of Waterloo



© Mahesh D. Pandey 2012
University of Waterloo

GENERAL INFORMATION

The University Network of Excellence in Nuclear Engineering (UNENE) in partnership with the Natural Sciences and Engineering Research Council (NSERC) established the Industrial Research Chair in Risk-Based Life Cycle Management of Engineering Systems at the University of Waterloo in September 2004.

The research program, chaired by Dr. Mahesh Pandey, is in its second five-year phase and will run from 2009 to 2014.

CONTACT INFORMATION

Dr. Mahesh Pandey
Professor / NSERC-UNENE Industrial Research Chair
Risk-Based Life Cycle Management of Engineering Systems
Department of Civil and Environmental Engineering
University of Waterloo, Waterloo, ON, N2L 3G1, Canada
Phone: (519) 888-4567 ext. 35858
Fax: (519) 888-4349
E-mail: mdpandey@uwaterloo.ca
Web: <http://www.watrisk.ca>

Contents

1. BASIC TERMINOLOGY	1
1.1 The Microsoft Excel Window	1
1.2 Navigation	3
1.3 Selecting Cells	3
2. ENTERING AND EDITING DATA	5
2.1 Types of information	5
2.2 Entering Data	5
2.3 Entering an Equation	6
2.4 Entering a Function	7
2.5 Error Messages	9
2.6 Copying and Pasting Information	9
2.7 Naming or Declaring Constants	11
3. FORMATTING	12
3.1 Cell Formatting	12
3.2 Sorting Data	12
4. CREATING CHARTS	13
4.1 Basics	13
4.2 Creating a Chart	13
4.3 Column Charts	14
4.4 Chart Formatting	17
4.5 Scatter (X-Y) Plots	18
Appendix A – Common Excel Functions	23
Appendix B – Common Statistical Functions	24

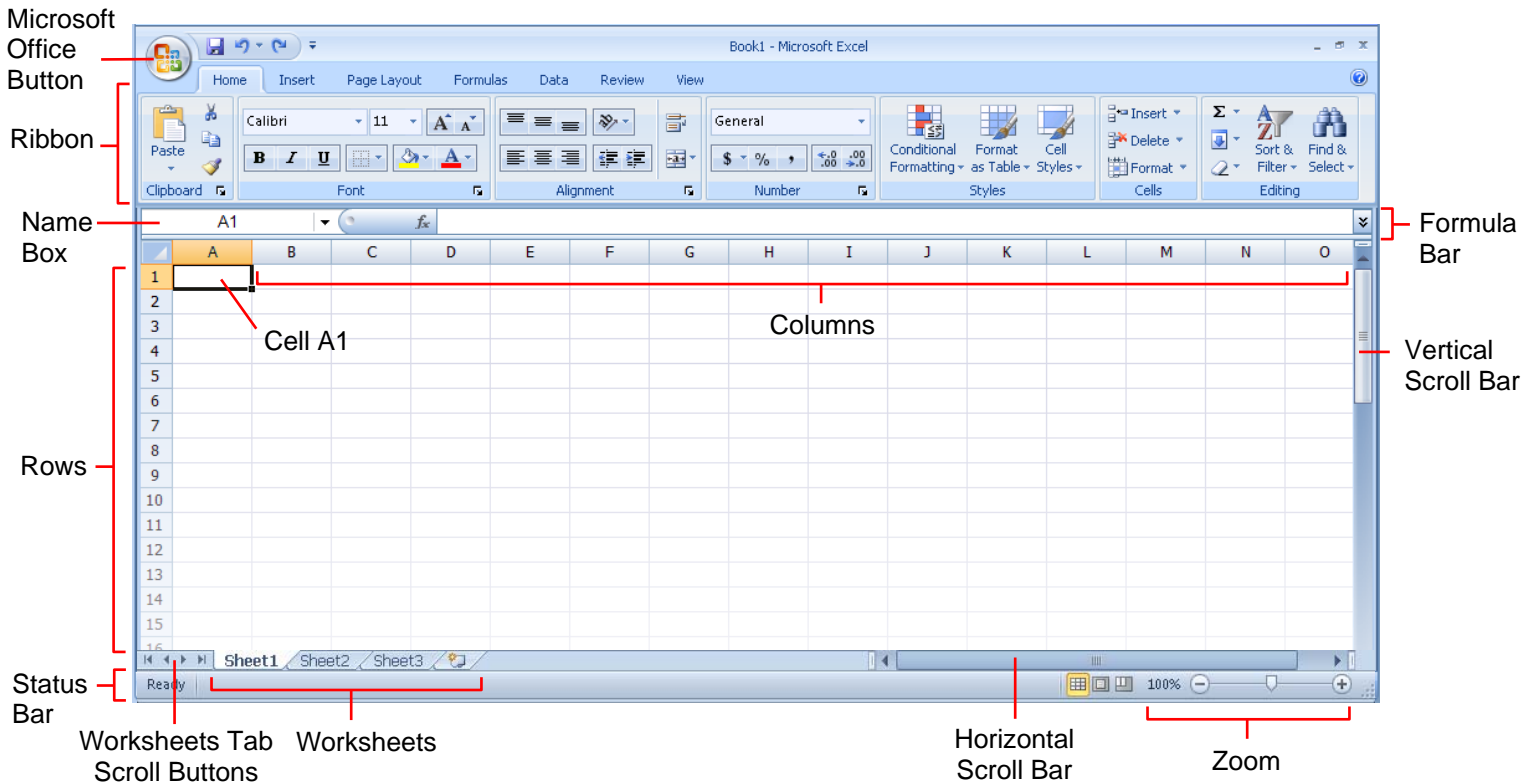
(this page left blank intentionally)

1. BASIC TERMINOLOGY

The objective of this tutorial is to teach the basics of Microsoft Excel 2007. Basic knowledge of how to navigate in the Windows environment using the mouse and keyboard is expected.

Excel 2007 is similar but also very different from the previous versions of the program. Some of the major new features include the Microsoft Office Button and the Ribbon toolbar.

1.1 The Microsoft Excel Window



1.1.1 Microsoft Office Button

The **Microsoft Office Button** contains many of the functions that were found under the **File menu** in older versions of Excel, such as Open, Save, Close, Print, etc.

Excel **Options**, which used to be found under the **Tools menu** in previous versions of Excel, are also found here, as well as the **Add-ins** feature.

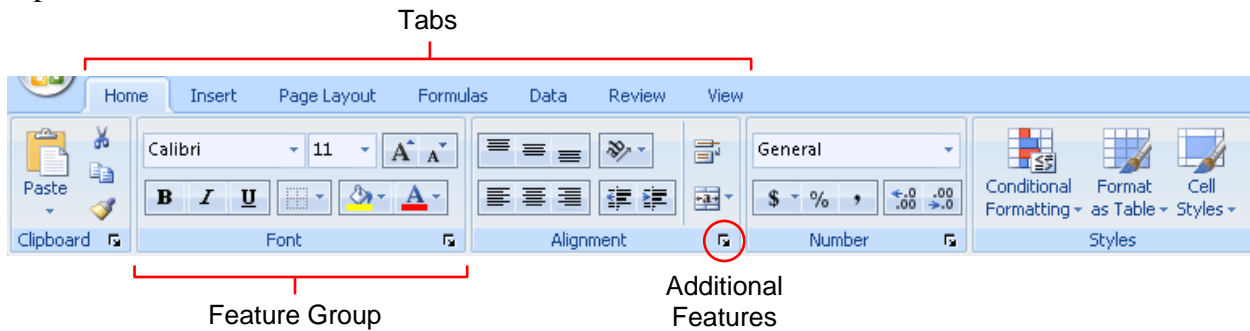


1.1.2 Ribbon

Perhaps the most noticeable change from the previous versions of Excel is the **Ribbon panel** or toolbar. This panel has essentially replaced the old drop-down menu structure of the previous Microsoft Office programs.

1. BASIC TERMINOLOGY

The Ribbon panel contains the most common features and functions for Excel under a number of separate **Tabs**.



Commonly used features are displayed and **Grouped** within each **Tab** in the **Ribbon**. Additional features for each **Group** are found by clicking the arrow at the bottom right hand corner of each **Group**.

The seven Tabs contain the following feature groupings:

Home:	Clipboard, Fonts, Alignment, Number, Styles, Cells, Editing
Insert:	Tables, Illustrations, Charts, Links, Text
Page Layouts:	Themes, Page Setup, Scale to Fit, Sheet Options, Arrange
Formulas:	Function Library, Defined Names, Formula Auditing, Calculation
Data:	Get External Data, Connections, Sort & Filter, Data Tools, Outline
Review:	Proofing, Comments, Changes
View:	Workbook Views, Show/Hide, Zoom, Window, Macros

The **Home Tab** essentially corresponds to the **Formatting toolbar** in the previous versions of Excel, while the **Chart** options are found under the **Insert Tab**.

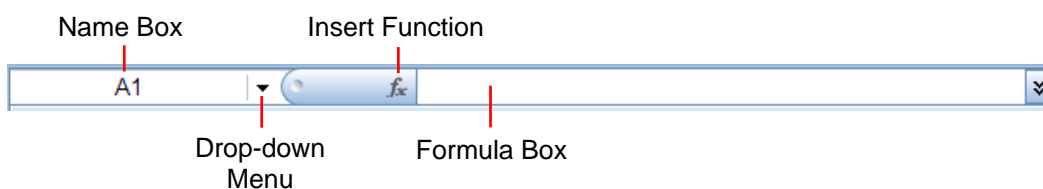
Three additional **Tabs** also appear on the **Ribbon** when editing/selecting an Excel **Chart**:

Design:	Type, Data, Chart Layouts, Chart Styles, Location
Layout:	Current Selection, Insert, Labels, Axes, Background, Analysis, Properties
Format:	Current Selection, Shape Styles, WordArt Styles, Arrange, Size

1.1.3 Basic Objects

Microsoft Excel consists of **worksheets**. Each worksheet consists of **rows** (up to 1,048,576) and **columns** (A to XFD, for a total of 16,384 columns). The combination of a column and row coordinate defines a **cell** address. For example, A1 refers to the cell in column A, row 1.

1.1.4 Formula Bar



The **Formula bar** consists of the **Name Box** and the **Formula Box**. The Name Box identifies the name of the currently selected cell. It can also be used to navigate to a specific cell location in the worksheet simply by entering the desired cell address. All **Named cells** (more on this later...) will also be listed in the *drop-down menu* of the Name Box.

The Formula Box can be used to enter data (i.e., text, functions, etc.) into the cell reference identified in the Name Box.

1.1.5 Status Bar



The **Status bar** appears at the bottom of the screen. It displays information with respect to Excel's status (i.e., opening, calculating, etc.) and contains the **Zoom** options for the worksheet. The status bar also identifies whether control selection keys, such as Num Lock or Caps Lock, are in use. Various options can be accessed by *right-clicking* on the Status bar.

1.2 Navigation

Basic movement around Excel is accomplished using either the mouse or keyboard (or both). Use the mouse to select a cell(s) you wish to manipulate (i.e., enter data, copy, clear, etc.). In many cases, the keyboard also offers an efficient way for moving around Excel. The most useful keyboard strokes are listed in the table below.

Movement	Key Stroke
One cell up	Up arrow key
One cell down	Down arrow key or Enter
One cell left	Left arrow key
One cell right	Right arrow key or Tab
Top of row	Ctrl + up arrow key
Bottom of row	Ctrl + down arrow key
Any cell	<F5> or Ctrl+g

There are also many special mouse actions that allow the efficient manipulation and navigation around the spreadsheet. For example, *double-clicking* on the **edge** of a selected cell will move the selection to the top-most (or bottom, left, or right, depending on the edge that is double-clicked) cell in a *group of cells* containing data.

1.2.1 Name Box

Similar to function key <F5> and/or **Ctrl+g**, the **Name Box** in the Formula bar can also be used to navigate to a specific cell in the worksheet simply by entering the cell reference into the box.

1.3 Selecting Cells

Before a cell can be modified (copied, etc.) or formatted, it must first be selected. Selecting a single cell should be straightforward (see above), while the selection of multiple cells follows standard Windows convention. That is, the **Ctrl** key can be used to select multiple cells individually, while the **Shift** key can be used to select a **range of cells** (between two selections). Naturally, both keys can also be used in combination.

1. BASIC TERMINOLOGY

Cell selection is generally easy to accomplish with the mouse in conjunction with the *Ctrl* and *Shift* keys. Selecting a group of cells can be done with the mouse by simply holding down the left mouse button while selecting the cells (i.e., essentially **dragging** the mouse over the desired cells with the left mouse button depressed). While this method is easy, it may become cumbersome with a large data set.

For large data sets, it may be beneficial to combine the *Ctrl* and *Shift* keys with the navigational instructions provided in the previous section. For example, selecting many rows of data arranged in a single column could be accomplished by

1. Selecting the top most cell.
2. Holding down the *Shift* key.
3. Double-clicking on the bottom edge of the selected cell (while holding down *Shift*) (or pressing *Ctrl* + *down arrow* keys while also holding down the *Shift* key).

2. ENTERING AND EDITING DATA

This section describes the methods and types of data that can be entered into the cells of an Excel worksheet.

2.1 *Types of information*

There are three basic types of information that can be entered into the cells in Excel. These are:

- Text
- Numbers
- Functions, equations, or formulas

Text or labels are simply text with no numerical value. Numbers represent values that are defined as *constants*. Formulas and functions are mathematical equations used to describe the relationship between constant values and/or various cells.

2.2 *Entering Data*

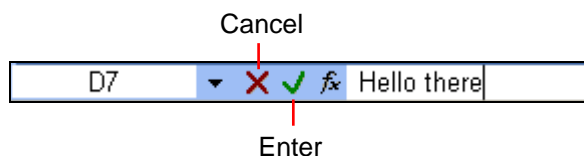
There are various ways to enter information or data into a cell. After *selecting* a particular cell, for example, cell A1, you can enter information by

1. Typing the information using the keyboard directly.
2. *Double-clicking* on the cell with the mouse and then typing in the information.
3. Clicking on the **Formula Box** and then typing in the information.
4. Pressing <F2> and then typing in the information.

Note that the information that is being entered will be shown in two places: within the **active cell** as well as the **Formula Box** at the top of the worksheet.

Once the desired information has been typed into the cell using any of the above methods, for example, the text “Hello there”, the information is entered using the **Enter** (or *Tab*) key. The information can also be entered by clicking the **green checkmark** on the Formula Bar right beside the Formula Box.

To *cancel* the operation, i.e., deciding not to enter the data, press the **Esc** key or the **red x-mark** on the Formula Bar. Note that the *Cancel* and *Enter* symbols are only shown on the Formula bar while the information is being entered into a cell.



2.2.1 Cursor Styles

There are five common cursor or mouse pointer styles in Excel. Understanding the different styles is important for editing purposes. Each of the styles is shown below:

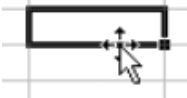
2. ENTERING AND EDITING DATA



Click in a cell to select a particular cell, or *click and drag* to highlight a group of cells.



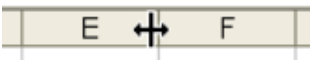
Click and *drag* to perform two operations: *inward* will **clear or erase** the cell contents; *outward* **fills** the cell contents into cells below or to the right.



Click and *drag* to move the cell to another location.



Click to select an entire column (or row).



Click and *drag* to change the column width (or row height).



Denotes data entry mode.

2.2.2 Editing a cell entry

Editing the contents of an existing cell is accomplished the same way as entering data, i.e., by *double-clicking* on the cell, clicking on the **Formula Box**, or by pressing <F2>.

To **clear** or delete the contents of an *active cell*, or multiple selected cells, press the **Delete** key or select **Clear** from the Editing group under the Home tab on the Ribbon. You can also **replace** the contents of an existing cell directly by first selecting it and then typing the new data using the keyboard (i.e., without *double-clicking* or pressing <F2> which initiates the editing mode).

2.3 Entering an Equation

The main feature of a spreadsheet is the ability to create and evaluate mathematical functions and formulas. The formulas and functions entered into Excel describe the relationship between various types of data. The most significant aspect, however, is not only to describe the relationship between text and constants, but to relate the contents of various cells with each other.

All equations and functions entered into Excel must begin with the *equal sign* “=”. The basic operators used in Excel are listed in the following:

Operator	Description
+	Addition
-	Subtraction
*	Multiplication
/	Division
^	Exponentiation
%	Percentage (i.e. divides the value by 100)
&	Connects or concatenates two values as text

For example, to evaluate the sum of two numbers (constants), type into a cell (or Formula Box) $=2+2$. Note that the result of the function is displayed on the spreadsheet, while the underlying formula is displayed in the Formula Box (when the cell is selected). Similarly, the following equation: $=A1-B1$, subtracts the value in cell B1 from the value in cell A1.

2.4 Entering a Function

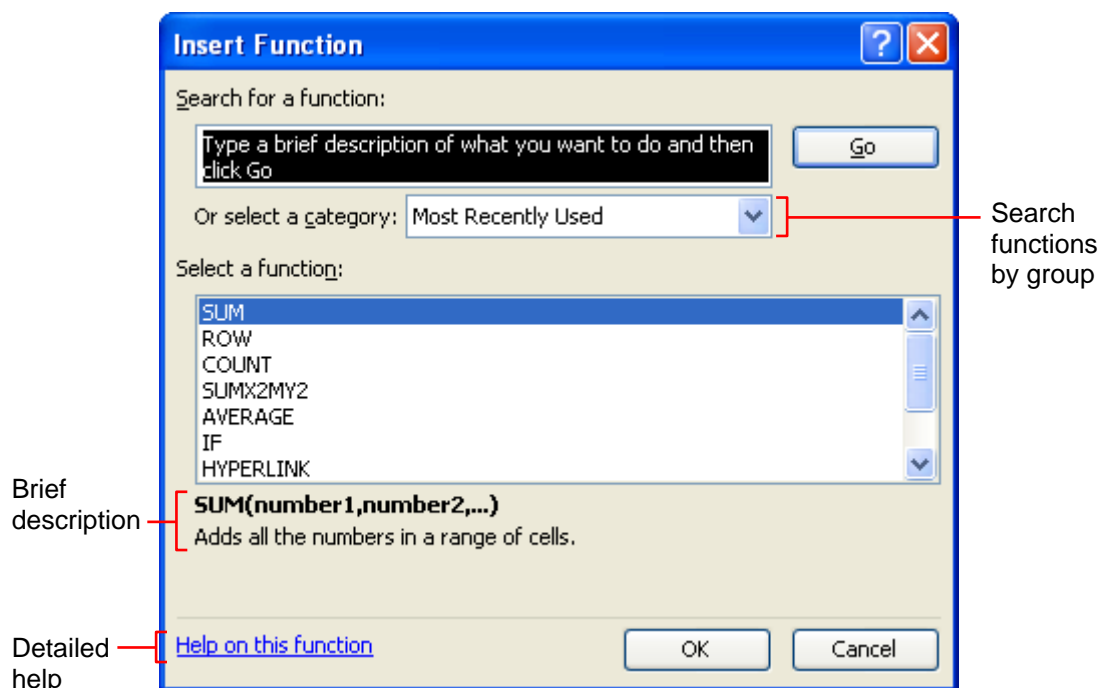
The real power or significance of the spreadsheet is in its ability to evaluate various functions, such as SUM, AVERAGE, etc. There are numerous functions that are built-in to Excel, while others can be loaded in using the Add-ins feature, or even custom created using the Visual Basic (VBA) programming environment. Similar to equations, all functions must begin with the *equal sign* “=” . A list of commonly used Excel functions is found in Appendix A at the end of this tutorial.

In Excel, a function is called by its **name**, followed by the **arguments** which are enclosed in brackets. A function can be inserted into a cell by

1. Typing it directly into a cell (by knowing the name and the arguments).
2. Selecting the **Formulas** tab from the Ribbon panel.
3. Clicking the *Insert Function* button in the Formula Bar.
4. Pressing *Shift* + <F3>.



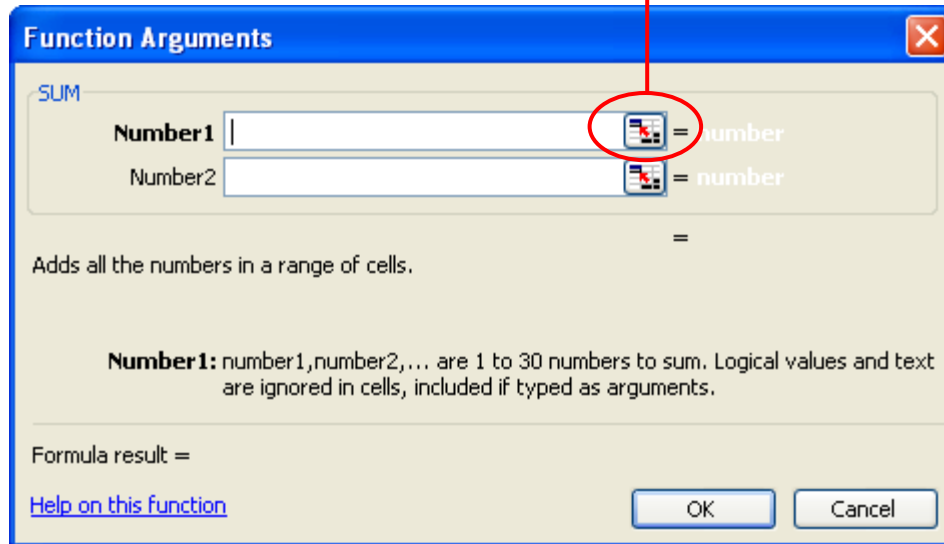
The **Insert Function** dialog box is shown below. This dialog box allows you to **search** all of Excel’s built-in functions by *name*, or by *group* or *category*. For example, all **statistical functions** are listed under the *Statistical* group. A brief description of each function is provided at the bottom of the dialog box. More detailed help can also be found by pressing the “[Help on this function](#)” link. Accessing additional help is highly recommended, especially when dealing with a function for the first time.



2. ENTERING AND EDITING DATA

The function is selected by clicking the [OK] button. This brings up another dialog box for the **Function Arguments**. Below is the example for the SUM function. The SUM function adds a list of numbers (constants) or all the numbers in a **range** (i.e. group) of cells.

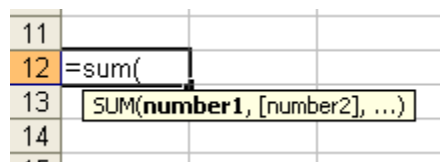
Click here to select a range of cells in the spreadsheet



All function *arguments* are separated by **commas** in Excel. For example, entering 1,2,3 above as the arguments for the SUM function would return the sum of these numbers. A **range** of cells is denoted by a **colon**. For example, entering A1:A10 (or selecting the cells by mouse or keyboard) as the argument for the SUM function above would return the sum of all the values between cells A1 and A10, inclusive. Similarly, entering A1:A5, B1:B5, C1:C5 (or A1:C5) would return the sum of all numbers in the area bounded by rows 1 to 5 and columns A to C.

Once the function has been entered by clicking the [OK] button, the spreadsheet displays the result of the function, while the Formula Box shows the actual function (or equation). In the case of the above example, the cell would therefore contain **=SUM(A1:A10)**.

Knowing the different built-in functions allows the functions to be written directly into each cell without having to go through the Insert Function dialog box. For example, the above function could be entered into cell A12 by first selecting the cell A12, and then typing in the function name, followed by the *left bracket*



If the function is recognized by Excel, a small dialog will be displayed under the cell indicating the required arguments for the function. The arguments can now be entered manually, or by selecting the desired cells in the spreadsheet using the mouse or keyboard, followed by *right bracket* and *Enter*. Below is the example showing the addition of numbers in cells A1 to A10.

The screenshot shows the Microsoft Excel 2007 interface. The ribbon is set to the 'Formulas' tab. The formula bar displays the formula `=SUM(A1:A10)`. The spreadsheet shows a column of numbers from 1 to 10 in cells A1 through A10. Cell A12 contains the result of the sum, which is 55. Red brackets point to the formula bar and the result cell with labels 'Function is displayed here' and 'Result is displayed here' respectively.

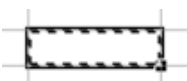
2.5 Error Messages

There are various **error messages** displayed by Excel when the cell entries violate specific requirements or conditions. For example, including text as an argument in the SUM function returns the #NAME? error. The different errors and their interpretation are provided below.

Error	Description
#NULL!	Occurs when you specify an intersection of two areas that do not intersect. The intersection operator is a space between references.
#DIV/0!	Number is divided by zero (0).
#VALUE!	Wrong type of argument or operand is used.
#REF!	Cell reference is not valid.
#NAME?	Microsoft Excel doesn't recognize text in a formula.
#NUM!	Invalid numeric values in a formula or function.
#N/A	Value is not available (e.g. missing) to a function or formula.

2.6 Copying and Pasting Information

Cutting, copying, and pasting cell contents follows the standard Windows convention. The keyboard shortcuts are **Ctrl+x**, **Ctrl+c**, and **Ctrl+v**, for cutting, copying, and pasting, respectively. These functions are also found in the Clipboard group under the Home tab on the Ribbon. The Cut and Copy operations are acknowledged in Excel by special *highlighting* around the selected cells and a message in the Status Bar at the bottom of the spreadsheet



Select destination and press ENTER or choose Paste

2. ENTERING AND EDITING DATA

2.6.1 Cell Referencing

By default, Excel uses **relative referencing** to refer to a particular cell, e.g., cell A1. However, when a formula containing relative referencing is copied from one cell to another, Excel does not create an exact copy of the formula. Rather, it will change the cell addresses relative to the row and column they are moved to. For example, copying the formula `=SUM(A1:A10)` in cell A12 to cell B12 will result in `=SUM(B1:B10)` in cell B12.

In many instances, relative cell referencing is a desirable feature of the spreadsheet. However, to avoid the referencing changes when copying a formula, the relative references can be replaced by **absolute cell references**. This is accomplished by placing the *dollar sign* “\$” in front of the row or column reference for a particular cell. For example, copying the formula `=SUM(A1:A12)` in A12 into B12 would now result in the same formula `=SUM(A1:A12)` in cell B12.

It is also possible to use **mixed cell referencing** by only fixing the row OR column reference, depending on the situation. For example, as \$A1 or A\$1. Rather than manual typing, the function key `<F4>` can be used to quickly insert absolute or mixed referencing for a given cell while editing the formula or function arguments. Pressing `<F4>` repeatedly will cycle through the different combinations of mixed references.

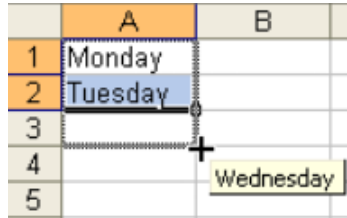
2.6.2 Paste Special

When the cell contents are text or constants, the copying and pasting operations are straightforward. However, when copying functions, it is important to realize that Excel is copying the *contents* of the cell, which in this case is the actual function and NOT the *result* of the function. The result of the function can be copied **as a value** to another cell in Excel using the **Paste Special** feature. Paste Special is found under the Paste function in the Clipboard group in the Home tab on the Ribbon or by *right-clicking* on the destination using the mouse. As shown below, Paste Special allows the pasting of specific features as opposed to all cell contents.



2.6.3 Filling Cells Automatically

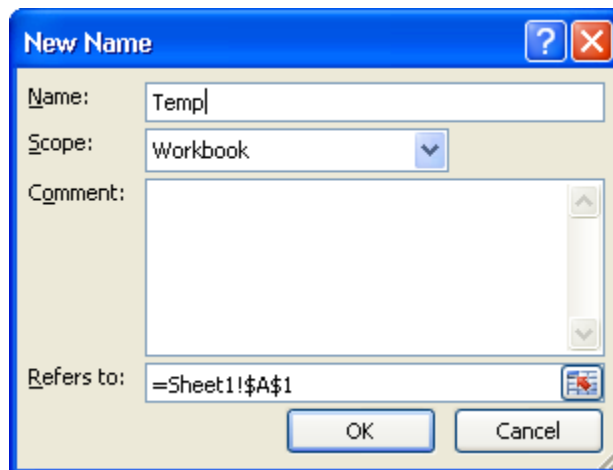
Excel is able to detect patterns in the spreadsheet data, which can then be reproduced automatically. This applies to both text and numbers (constants). For example, typing in Monday to cell A1 and Tuesday to cell A2, then selecting cells A1 and A2 and *clicking and dragging* the **bottom right corner** of the range downward (while holding down the mouse button) will fill in the appropriate weekdays automatically. The same logic applies to numbers, for example, typing in the numbers 1 and 2, and then filling down the rest of the numbers in the sequence by clicking and dragging.



2.7 Naming or Declaring Constants

To avoid the confusion of relative and absolute cell referencing, it is possible to **assign a name** to a specific cell or range of cells. This is equivalent to declaring variables in a programming environment. The named cells can then be used as arguments in formulas and functions, rather than referring to the particular cell address.

A cell, or group of cells can be named simply by selecting them and then typing the desired name directly into the **Name Box** in the Formula bar. The cell name definition is also found in the **Defined Names group** under the **Formulas tab** on the **Ribbon**. Below is an example of naming cell A1 as Temp using the > **Define Name** option dialog box. All named constants can be managed (i.e., modified, deleted, etc.) through the **Name Manager**.



Once the new named cell is added (by clicking [OK]), it can be used as an argument in any formula or function. For example, the equation $=\mathbf{Temp+B2}$ is now equivalent to the equation $=\mathbf{\$A\$1+B2}$. The named cell always refers to the same location and hence acts as an *absolute cell reference*.

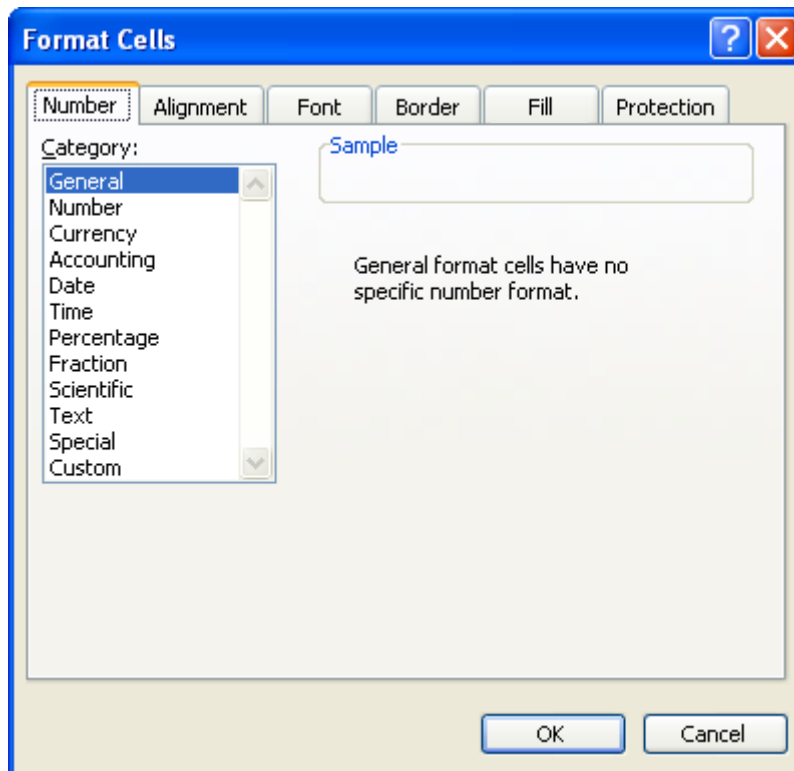
3. FORMATTING

There are many ways to format and customize the way in which information is displayed in Excel. The objective of this section is only to give a very brief overview of the basics.

3.1 Cell Formatting

The cells in Excel can be formatted individually or as a group. Cell formatting can be applied to the selected cells through

1. The various formatting options under the **Home tab** on the Ribbon.
2. *Right-clicking* on the particular cell or group of cells for various options.
3. Pressing **Ctrl+I** to launch the **Format Cells** dialog box.



The cell formatting options allow changes to be made to the number or text formatting (e.g., the number of significant digits to display, date and time formatting, etc.), cell alignment, font style, and cell border and colouring. The cell contents can also be *protected* and *locked* to prevent other users from modifying them.

Besides cell formatting, it is also possible to change the column and row properties (i.e., width and number) as well as modify the worksheets in the workbook (*add, delete, rename, etc.*).

3.2 Sorting Data

Data can be sorted in Excel by clicking on the **Sort & Filter button** in the Editing group under the Home tab on the Ribbon, or by *right-clicking* the selected cells and choosing > **Sort...** from the drop-down menu. Make sure to *select the data* prior to the operation.

4. CREATING CHARTS

4.1 Basics

In addition to powerful data analysis through equations and functions, a key aspect of Excel is the ability to chart or plot the data in graphical form. Excel offers a large library of different types of charts that can be produced, however, the most useful **chart types** are Column charts and Scatter (X-Y) plots.

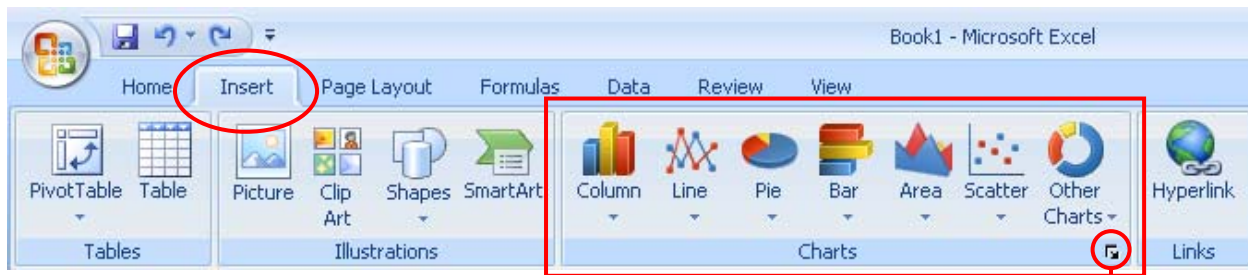
Similar to equations and functions, Excel charts are dynamic (when the chart data is linked to cells in the spreadsheet) and will change (update) automatically upon modification of the data.

4.1.1 Chart Wizard

Previous versions of Excel contained an automated step-by-step process for creating charts called the Chart Wizard. This feature is **no longer available** in Excel 2007.

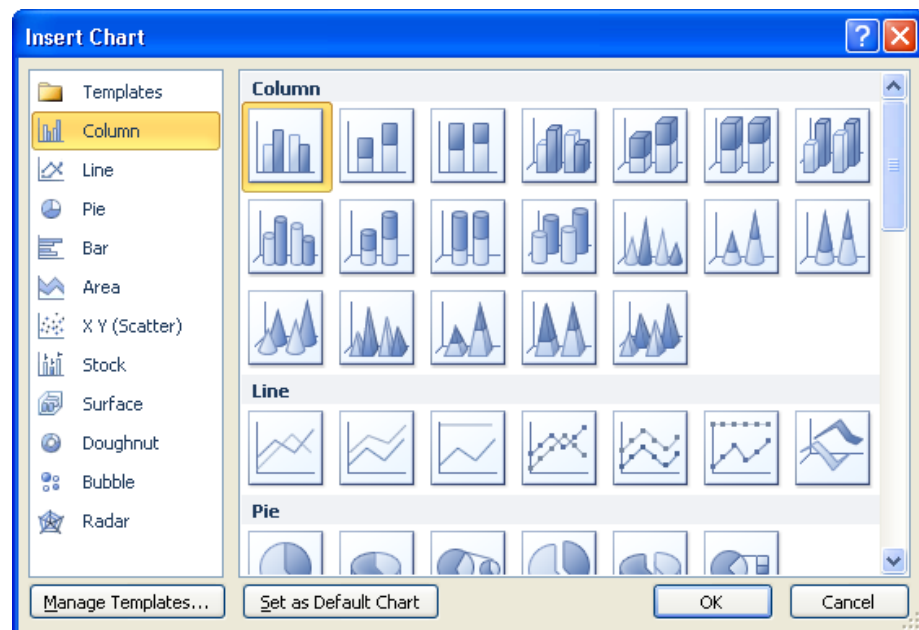
4.2 Creating a Chart

In Excel 2007, charts are created by choosing one of the available Chart types from the **Charts Group** under the **Insert Tab** on the **Ribbon**.



Insert Chart Dialog Box

The **Insert Chart** dialog box can also be activated by selecting the Additional Features from the Chart group or by pressing the keys **Alt-i-h** in sequence.



4. CREATING CHARTS

The **data** for charting can be chosen either **before** or **after** selecting a particular Chart type. Selecting the data beforehand is generally preferred, however, it requires knowledge about the way in which Excel interprets the selected data (i.e., which part of the selected data will be plotted on which axis). If no data is selected, Excel will create an **empty Chart** as an object within the worksheet.

Note: Excel attempts to create the Chart data based on the cell(s) selected and its neighbours at the time a particular Chart type is selected. To avoid confusion, it is therefore important to either select the specific data to be plotted or **select an empty cell with empty neighbours** before creating a Chart.

Whenever a **Chart object** (or worksheet) is selected in Excel, three new tabs labeled **Chart Tools** will appear on the Ribbon. These tabs contain the most commonly used functions and features related to Excel Charts, and are grouped as follows:



- Design:** Type, Data, Chart Layouts, Chart Styles, Location
- Layout:** Current Selection, Insert, Labels, Axes, Background, Analysis, Properties
- Format:** Current Selection, Shape Styles, WordArt Styles, Arrange, Size

4.3 Column Charts

This example shows how to make a simple column chart. The example consists of some simple data shown below.

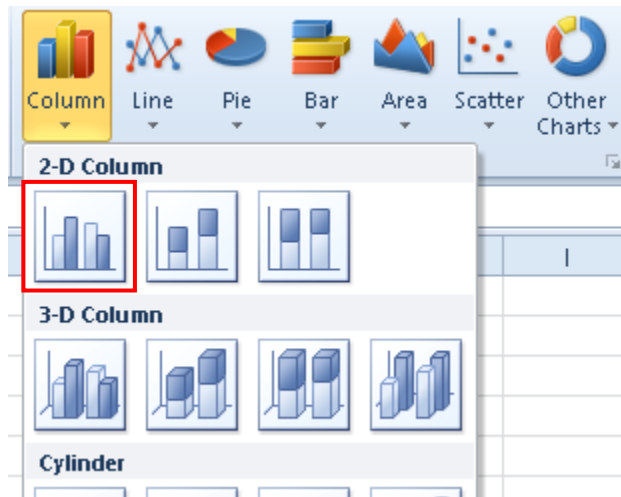
	A	B	C	D	E	F
1	Number of Pump Failures					
2						
3	Year	Failures				
4	1995	5				
5	1996	4				
6	1997	7				
7	1998	2				
8	1999	1				
9						
10						
11						

The data consists of number of pump failures observed in each year. The *label* for the years is located in cell A3, while the actual years are listed in cells A4:A8. Similarly, the *label* for the failures is in cell B3, with the data in cells B4:B8.

For this example, we make an **empty Chart** first and then explicitly define the source data from the Chart Tools Design tab. As noted above for empty Charts, it is important to **select an empty cell with empty neighbours** before creating the Chart. In the above example, we have selected cell C10.

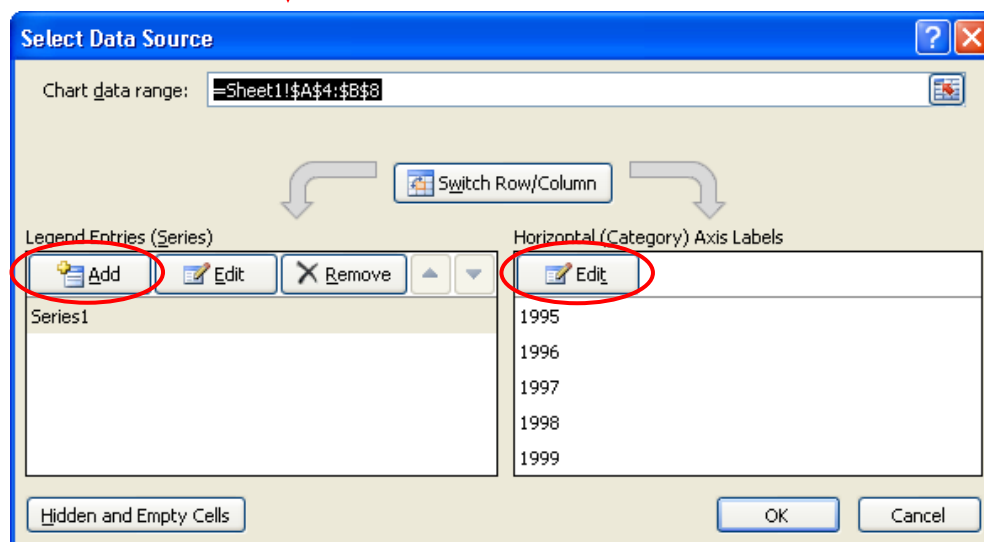
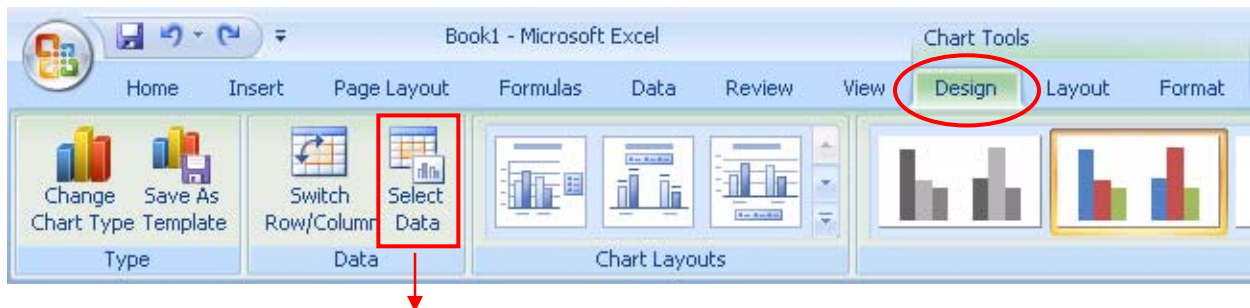
4.3.1 Step 1 - Create an Empty Column Chart

After selecting an empty cell with empty neighbours, we create an empty Chart by selecting the **2-D Clustered Column** chart from the Charts group. This creates an empty Chart object within the worksheet.



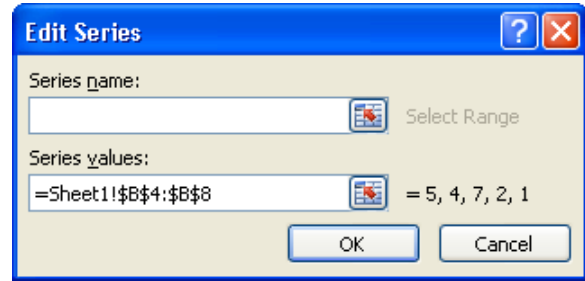
4.3.2 Step 2 – Define the Source Data

The source data for the Chart is defined by the **Select Data** option in the Data group under the Design tab in the Chart Tools. This option can also be accessed by *right-clicking* on the Chart object and then selecting > Select Data... from the drop-down menu.

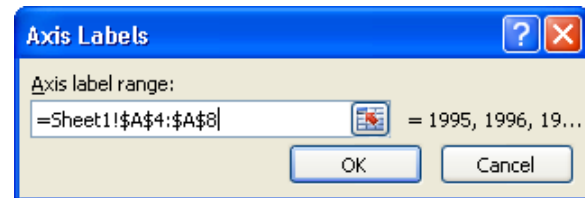


4. CREATING CHARTS

From the Select Data Source dialog box, we press the **[Add] Legend Entries (Series)** button to specify the data we want to plot for each column. In this case, these are the pump failures in range B4:B8 in the worksheet named “Sheet1”. Note that Excel automatically adds the “\$” signs for the selected data to make the references fixed. Similar to equations, each definition begins with the equal sign “=”, while the worksheet name is followed by the exclamation mark “!”.

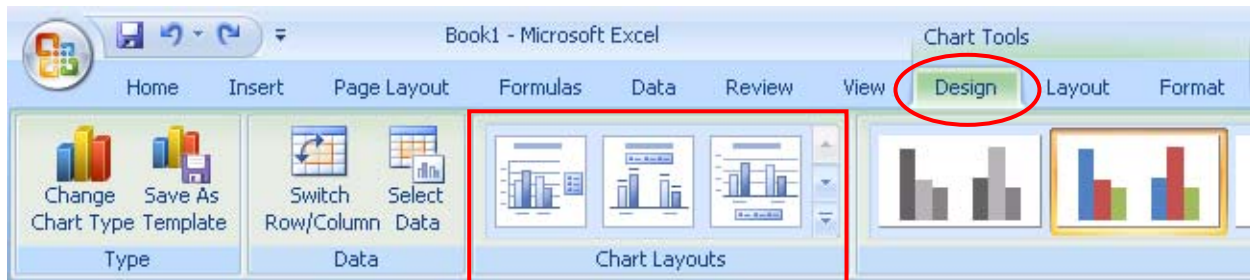


We then press the **[Edit] Horizontal (Category) Axis Labels** button to specify the column labels on the horizontal axis and select the data (years) in the range A4:A8.



4.3.3 Step 3 - Chart Formatting

Once the appropriate data have been selected or defined for the Chart through the Source Data dialog, we need to label and format the Chart appropriately. This can be done using the **Chart Layouts** Group under the Design tab in the Chart Tools.

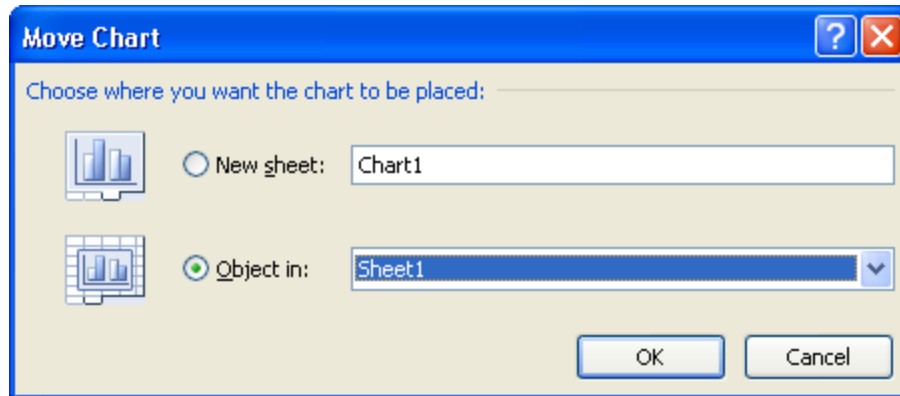


The Chart Layouts group contains various standard formats that specify the presence of Chart and axes titles, as well as the display of **Axes**, **Gridlines**, the **Legend**, **Data Labels**, and the **Data Table**. General Chart formatting is available in the **Styles** group under the Design tab, while the specific formatting of individual Chart features or elements can be done through the various options under the **Layout** tab.

4.3.4 Step 4 - Chart Location

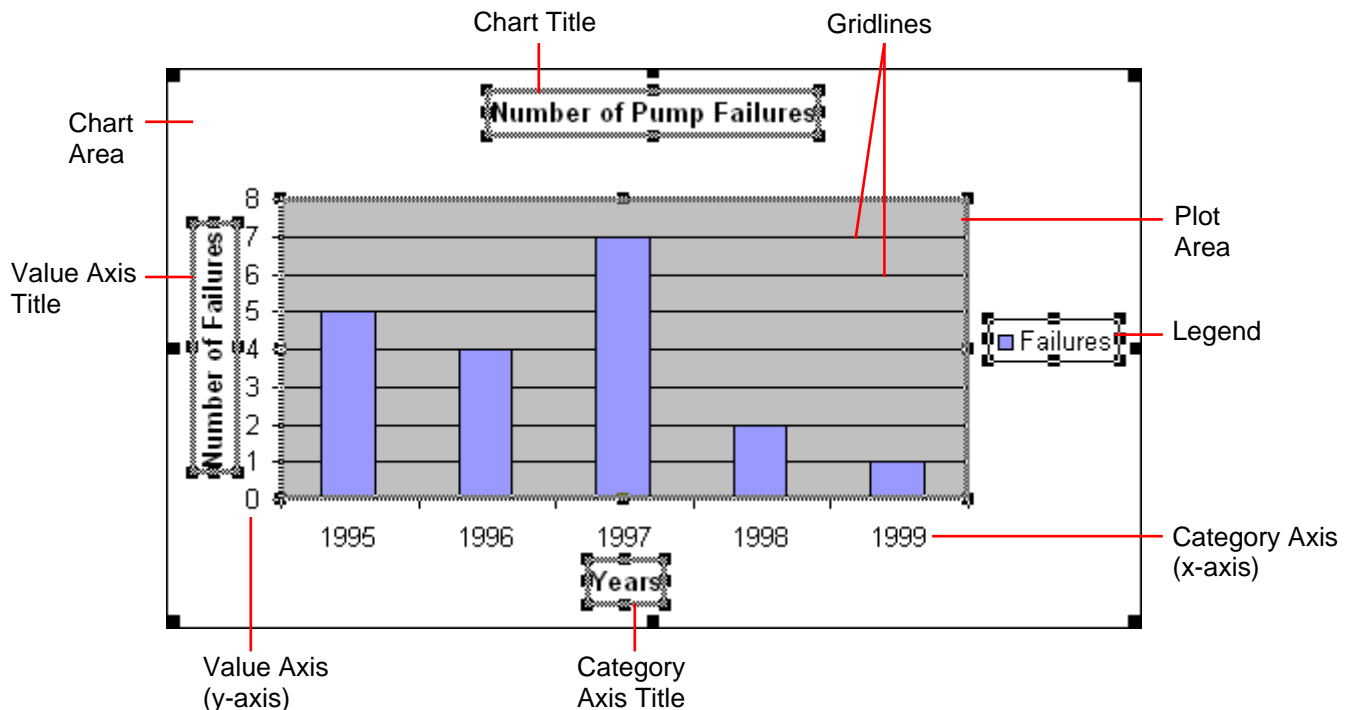
The default status of all Charts in Excel 2007 is as a Chart object within a worksheet. However, the Chart can also be located on a separate **worksheet** by itself. Having the Chart within the worksheet containing the data is often advantageous not only for referencing purposes, but also for directly visualizing the results if/when the data is modified.

The Chart location can be specified from the **Location** group under the Design tab in the Chart Tools, or by *right-clicking* on the Chart object and selecting > Move Chart... from the dropdown menu.



4.3.5 Finished Chart

The finished Chart is shown in the following with the different parts or objects of the chart labeled accordingly.



4.4 Chart Formatting

As shown above, the Chart consists of various objects which can be modified and edited as required. In order to edit the Chart and its objects, however, the Chart must first be **selected** (e.g., single click with the mouse). Text objects and labels, such as the Chart Title and Axis labels, can be **edited** similar to all cell contents by first selecting them, and then entering the appropriate text either directly (*double-clicking* or <F2>), or through the Formula box.

The various objects within a Chart can be **formatted** by

1. **Double-clicking** or **right-clicking** them with the mouse.
2. Using the various options under the **Layout tab** in the Chart Tools.

4. CREATING CHARTS

4.5 Scatter (X-Y) Plots

While column charts are used to plot data in specific categories or groups, the **Scatter plot** is used to plot two (or more) variables against each other. The Scatter plot can also be used to plot a **function** and perform **regression analysis** by fitting a **trendline** through the data points.

4.5.1 Plotting a Function

Suppose we have a simple equation of a parabola $y = ax^2$, where $a = 0.1$. This is a simple equation describing the relationship between the variables y and x ; for any value of x , we can compute the corresponding value of y . To plot this function in Excel, we therefore have to create pairs of values of x and y . The simplest way to do this is to assume the values of x , for example, ranging from -10 to +10, and then computing the corresponding value of y .

Adjacent is the example of the spreadsheet showing this computation. The x -values are specified in cells A6:A26, while the corresponding y -values are computed in B6:B26. The constant a is in cell B3. The equations to compute the y -values are

$$= \$B\$3 * A6^2 \text{ for cell B6}$$

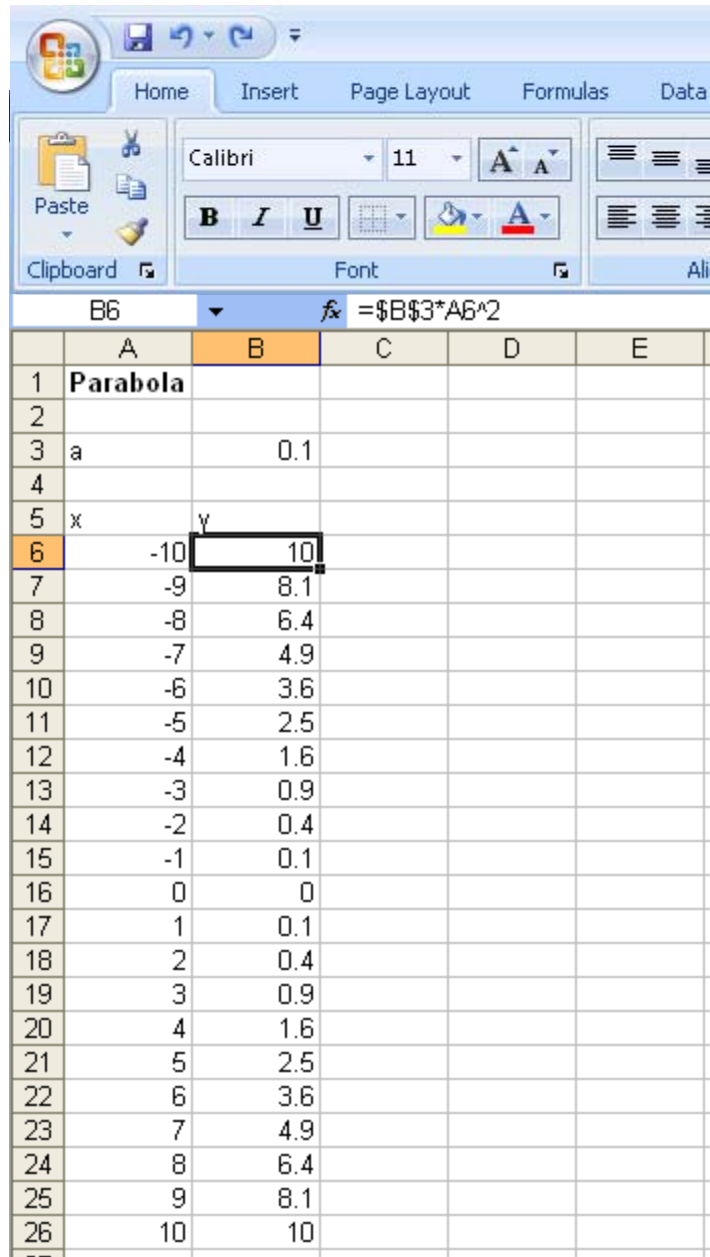
$$= \$B\$3 * A7^2 \text{ for cell B7}$$

etc.

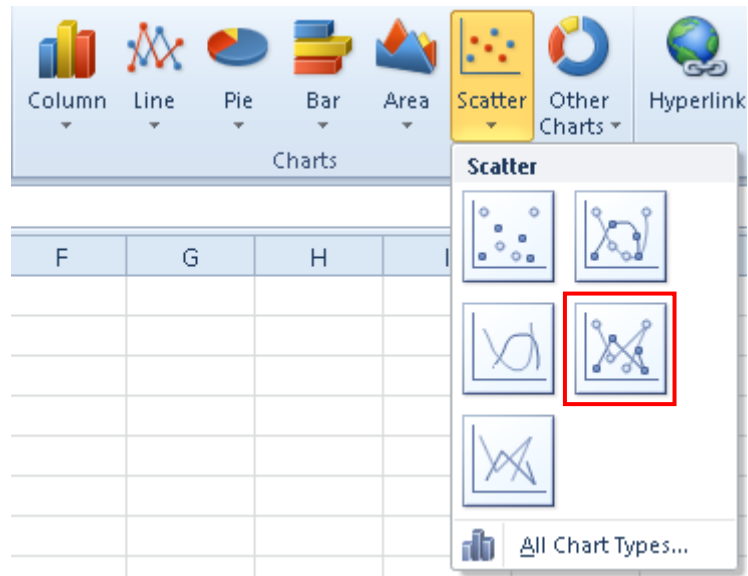
Because we used the absolute cell reference $\$B\3 for the constant a , we can copy the formula down all the way to cell B26. We could have just as easily used the constant value 0.1 in place of $\$B\3 , or *named* the cell $\$B\3 as a (e.g., using the Name Box) and then used a in the resulting equation.

To plot this function in Excel, we first **select all of the data** in the range A6:B26. Excel will automatically use the data in the first column as the x -values and all data in the subsequent column(s) as the y -values.

To make the Chart, we select the **Scatter plot with Straight Lines and Markers** from the Charts group under the Insert tab on the Ribbon.

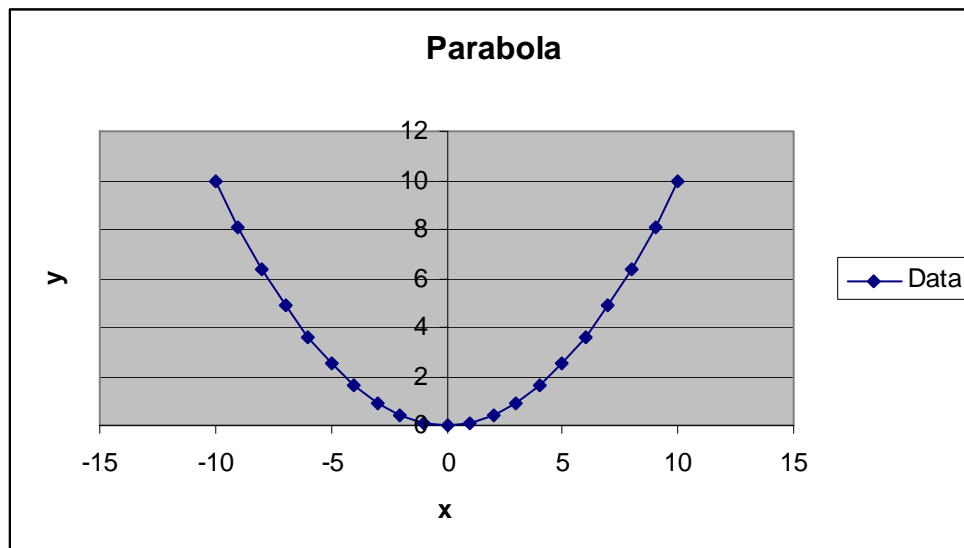


	A	B	C	D	E
1	Parabola				
2					
3	a	0.1			
4					
5	x	y			
6	-10	10			
7	-9	8.1			
8	-8	6.4			
9	-7	4.9			
10	-6	3.6			
11	-5	2.5			
12	-4	1.6			
13	-3	0.9			
14	-2	0.4			
15	-1	0.1			
16	0	0			
17	1	0.1			
18	2	0.4			
19	3	0.9			
20	4	1.6			
21	5	2.5			
22	6	3.6			
23	7	4.9			
24	8	6.4			
25	9	8.1			
26	10	10			



Similar to the column chart, there are many different **sub-types** for the Scatter plot to choose from. In this case, we want a scatter plot with the **data points connected by lines** because we are trying to plot a continuous function. Including the points in the line is optional, however, they serve as reminders of the fact that we are approximating the continuous function at discrete intervals. This may be important in cases when the resulting line is jagged or irregular, which may be the result of a coarse discretization scheme. It is also possible to plot the data using a **smoothed** line through the points, however, this may not always result in a realistic representation of the underlying function. To reduce the jaggedness in these cases, it is preferable to change the discretization scheme by adding more (computed) points to the plot.

After adding labels and formatting, the final plot is shown below. Because the Chart is an object in the worksheet (and because the equation for y contains a reference to a cell, rather than a constant value of 0.1) we can immediately visualize the change in the results by changing the value of the constant a in cell B3.



4. CREATING CHARTS

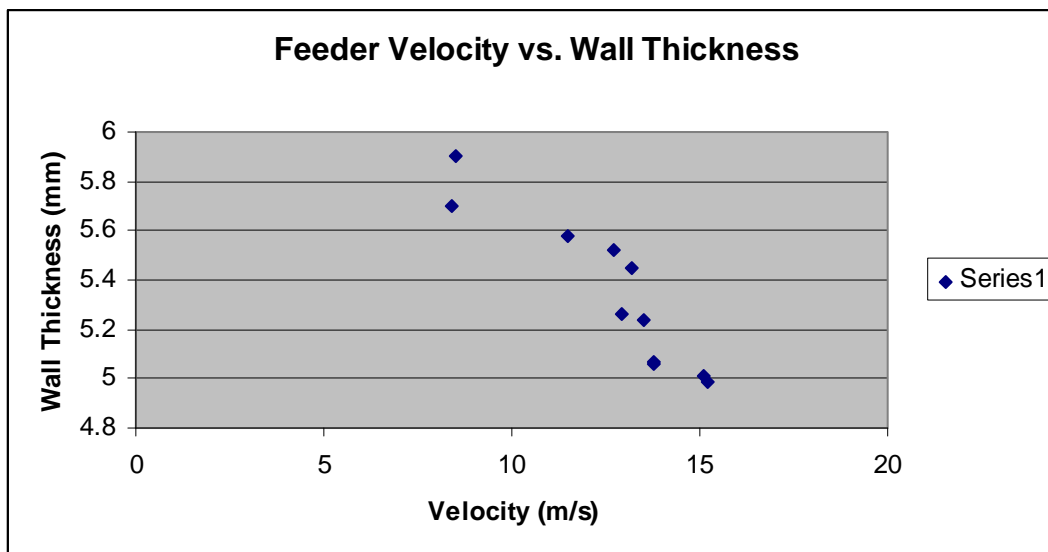
4.5.2 Adding a Trendline

Another important aspect of the Scatter plot is the ability to **fit a function** through the data points. This is equivalent to **regression analysis**, where the objective is to find the “best” fit line through the data points.

Consider the following data consisting of 11 observations of coolant velocity and the corresponding wall thickness of feeder pipes.

	A	B	C	D	E
1	Feeder Data				
2					
3	V	wt			
4	(m/s)	(mm)			
5	8.5	5.9			
6	8.4	5.7			
7	11.5	5.58			
8	12.7	5.52			
9	12.9	5.26			
10	13.2	5.45			
11	13.5	5.24			
12	13.8	5.07			
13	13.8	5.06			
14	15.1	5.01			
15	15.2	4.99			
16					

The objective is to find a relationship between the two variables velocity and wall thickness. To do this, we can use the **Trendline** feature of the Scatter plot. Similar to the previous example, we first select the data in the range A5:B15 and then choose the Scatter plot from the Charts group under the Insert tab. In this case, however, we want to **plot the points (or Markers) only**, since our objective is to fit a function (or line) through the data.

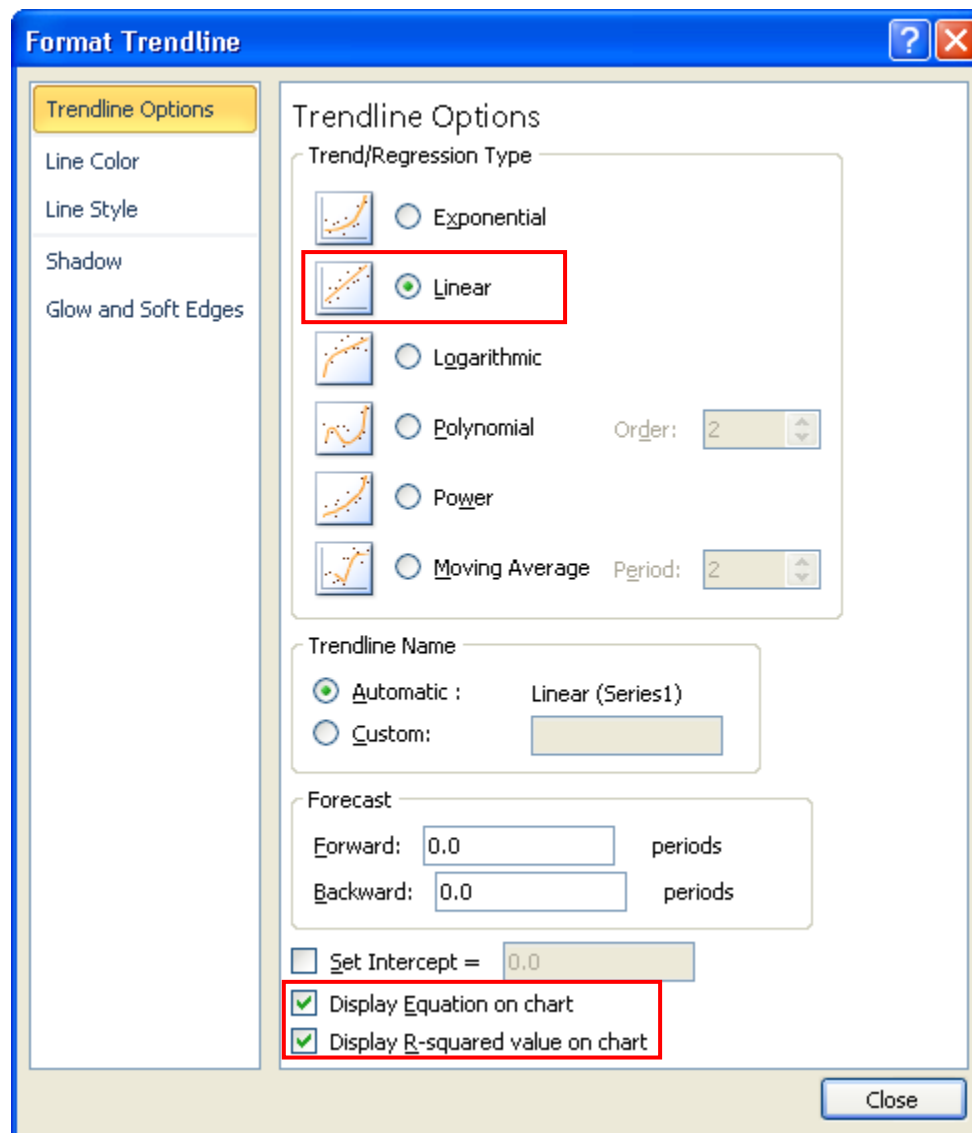


As we can see from the above plot, there appears to be a relationship between velocity and wall thickness. That is, the wall thickness tends to decrease with increasing coolant velocity. We can describe this relationship mathematically, by fitting a function through the data points using the **Trendline** option in Excel.

A **Trendline** can be added to a **data series** by

1. Selecting Trendline in the Analysis group under the Layout tab in the Chart Tools.
2. *Right-clicking* the appropriate data series (the data points in the chart) and selecting the > Add Trendline... option from the drop-down menu.

The **Format Trendline** dialog box is shown below. In this case, we select a **Linear** trend and display the **equation** and **R-squared** value on the Chart.

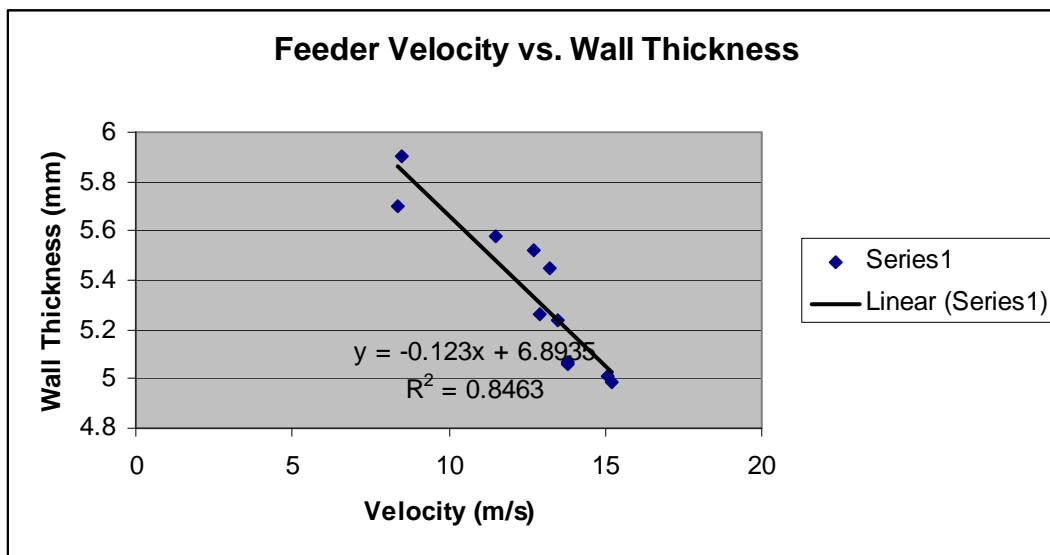


4. CREATING CHARTS

These options allow us to see the resulting mathematical relationship between the two variables and also the quality of the fit (linear in this case) based on the R-squared statistic. R-squared varies between 0 and 1, where 1 indicates a perfect (linear) relationship between the two variables, while 0 indicates that there is no relationship between the two.

The resulting plot is shown below. The linear best-fit trendline has now been added to the plot as **an object** or separate data series (which can then be formatted as required). The equation of the line has also been added to the plot as an object (which can then also be formatted for font size, number of significant digits, etc.).

The resulting fitted equation implies that $w_t = -0.123 V + 6.8935$. The R-squared value of 0.8463 indicates a fairly good fit, which can also be observed visually from the chart.



Appendix A – Common Excel Functions

Function	Description
ABS	Returns the absolute value of a number
AVERAGE	Returns the average (arithmetic mean) of its arguments
COUNT	Counts the number of cells that contain numbers
COUNTIF	Counts the number of cells within a range that meet a given condition
DEGREES	Converts radians to degrees
DEVSQ	Returns the sum of squares of deviations of data points from their sample mean value
EXP	Returns e raised to the power of a given number
FACT	Returns the factorial of a number
FREQUENCY	Calculates how often values occur within a range of values and then returns a vertical array of numbers
GAMMALN	Returns the natural logarithm of the gamma function
IF	Checks whether a condition is met, and returns one value if TRUE, and another value if FALSE
INDEX	Returns a value or reference of the cell at the intersection of a particular row and column, in a given range
INT	Rounds a number down to the nearest integer
INTERCEPT	Calculates the point at which a line will intersect the y-axis by using a best-fit regression line plotted through the known x and y values
LEN	Returns the number of characters in a text string
LN	Returns the natural logarithm of a number
LOG	Returns the logarithm of a number to the base you specify
LOOKUP	Looks up a value either from a one-row or one-column range or from an array
MATCH	Returns the relative position of an item in an array that matches a specified value in a specified order
MAX	Returns the largest value in a set of values
MIN	Returns the smallest value in a set of values
PI()	Returns the value of Pi
RADIANS	Converts degrees to radians
RAND	Returns a random number between 0 and 1
RANK	Returns the rank of a number in a list of numbers
ROUND	Rounds a number to a specified number of digits
RSQ	Returns the coefficient of determination of the best-fit linear regression line through the given data points
SLOPE	Returns the slope of the best-fit linear regression line through the given data points
SQRT	Returns the square root of a number
STDEV	Returns the standard deviation of a sample
SUM	Adds all the numbers in a range of cells
SUMIF	Adds the cells specified by a given condition or criteria
VAR	Returns the variance of a sample

Appendix B – Common Statistical Functions

Function	Description
BETADIST	Returns the cumulative beta probability distribution
BETAINV	Returns the inverse of the cumulative beta probability distribution
BINOMDIST	Returns the binomial probability distribution
CHIDIST	Returns the one-tailed probability of the chi-squared distribution
CHIINV	Returns the inverse of the one-tailed probability of the chi-squared distribution
COMBIN	Returns the number of combinations for a given number of items
CONFIDENCE	Returns the confidence interval for a population mean
CORREL	Returns the correlation coefficient between two data sets
COVAR	Returns the covariance between two data sets
EXPONDIST	Returns the exponential distribution
FDIST	Returns the F probability distribution
FINV	Returns the inverse of the F probability distribution
GAMMADIST	Returns the gamma distribution
GAMMAINV	Returns the inverse of the cumulative gamma distribution
LOGNORMDIST	Returns the cumulative lognormal distribution
LOGINV	Returns the inverse of the lognormal cumulative distribution function
NORMDIST	Returns the normal cumulative distribution for the specified mean and standard deviation
NORMINV	Returns the inverse of the normal cumulative distribution for the specified mean and standard deviation
NORMSDIST	Returns the standard normal cumulative distribution (has a mean of zero and a standard deviation of one)
NORMSINV	Returns the inverse of the standard normal cumulative distribution (has a mean of zero and a standard deviation of one)
PERMUT	Returns the number of permutations for a given number of objects that can be selected from the total objects
POISSON	Returns the Poisson distribution
STANDARDIZE	Returns a normalized value from a distribution characterized by a mean and standard deviation
TDIST	Returns the Student's t-distribution
TINV	Returns the inverse of the Student's t-distribution
WEIBULL	Returns the Weibull distribution