1

Introduction to Excel

In this chapter, students will learn to:

- Undertake basic operations in an Excel worksheet
- Perform mathematical calculations on worksheet data using formulas and functions
- Understand and apply relative and absolute cell referencing
- Visualize and interpret data sets in the form of charts

Excel is a Microsoft spreadsheet application widely used to store, organize, process and analyze many forms of data, including experimental data. It offers great flexibility and is, in many respects, unrivalled in terms of its functions as applicable to scientific experimental data. Researchers use spreadsheet applications such as Excel to work with experimental data. For example, they will transfer data to a spreadsheet such as Excel to:

- Store and organize experimental data
- Manipulate data using mathematical functions
- Visualize data, for example, through charts and tables
- Perform statistical analysis of data
- Apply curve fitting with linear and non-linear regression

Besides Excel, other examples of spreadsheet applications exist including free, open source software packages such as LibreOffice Calc and Google Spreadsheets. They operate in a similar manner to Excel in general, but differ in some features and hence functionality. Microsoft® Excel® has the most features and is currently more widely used than these open source alternatives. That said, the landscape is rapidly changing and these open source

Spreadsheet Applications in Chemistry Using Microsoft® Excel®: Data Processing and Visualization, Second Edition. Aoife Morrin and Dermot Diamond.

Companion Website: www.wiley.com/go/morrin/spreadsheetchemistry2

software packages are increasing in maturity and popularity. If you have access to Excel, it is the spreadsheet software program of choice. As such, the tutorials in this book are designed specifically around Excel. However, if Excel is not accessible, open source alternatives are a good option to work through the tutorials to learn approaches to processing experimental data.

This chapter introduces basic standard worksheet operations in Excel that will be needed for the later chapters. The tutorial exercises have been designed around Excel for PC. If you are using Excel for Mac, you can expect minor deviations from the tutorial instructions, as formats and styles, and locations of commands and options can differ between the two versions. Likewise, accessing tools and commands may differ if you are using an early version of Excel. However, most functionality is equivalent between versions and so all tutorials here can be undertaken using any version of Excel. Of course, it is advisable to upgrade Office if you are using a particularly archaic version. Once you are up and running with Excel, it is worth spending time working through the tutorials in this chapter to ensure that the more basic spreadsheet functions of Excel are understood before moving to the more advanced topics and tutorials in later chapters.

1.1 Navigating the Workbook

1.1.1 The Worksheet

Launching Excel brings you into a workbook containing a set of spreadsheets. Excel refers to each spreadsheet within a workbook as a 'worksheet'. Some basic aspects of the worksheet are labelled in Figure 1.1.

The **Ribbon menu** gives access to all tools and commands. Within the Ribbon tab, you can see several tabs – Home, Insert, Page Layout, Formula, Data, Review, and View. Each of these has their own **Ribbon display**, which

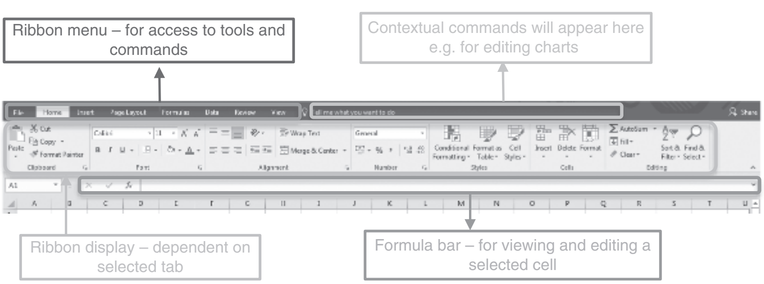


Figure 1.1 Highlighted aspects of Excel worksheet for navigation.

comprises groups of buttons representing a variety of commands that are displayed when each tab is selected.

Contextual tabs are special types of tabs that appear only when an object is selected, such as a chart or a shape. These contextual tabs contain commands specific to whatever object you are currently working on. For example, after you add a shape to a worksheet, a new **Format** tab appears as a **Contextual tab**. These tabs only activate when you work with particular objects. You will use these tabs regularly in the tutorials in this book.

The **Formula bar** is the toolbar at the top of the worksheet window that can be used to enter or copy an existing formula into cells. It is labelled with the function symbol *fx*. By clicking the **Formula bar**, or when you type the equal (=) symbol in a cell, the **Formula bar** will activate.

1.1.2 Worksheet Tools

You can navigate the Excel worksheet fairly intuitively using standard Office365 operations. The first tutorial here will use an already populated worksheet to show you some of the tools available.

Tutorial 1.1 Using Basic Formatting and Analysis Commands

In this tutorial, you will work with a data set relating to the Periodic Table to learn some basic formatting and analysis commands in Excel.

- Open the workbook *1.1_Periodic Table.xls*.
- In the worksheet, you will see columns of data related to the periodic table. Expand the width of the columns so that all text in each of the columns can be seen. To do this, bring the mouse cursor to where the row and column headers meet – see Figure 1.2. By clicking here you will select the whole worksheet. Then double-click any one of the column partition lines. This will readjust all column widths so that you can visualize the data clearly.
- Now take a look at column D – *Atomic Mass*. The values in the cells have 7 decimal places reported which is unnecessary for our purposes. To reduce the number of significant figures, first highlight the data by clicking at the top of column D. Right click and select **Format Cells**. In the pop-up dialogue box, select **Number** and enter *3* in the **Decimal Places** box. Press **OK**.
- Next, format the columns of data into a table so that you can sort the data. Highlight columns A to I and under the **Home** tab, click **Format as Table**. Choose a style you like in the dialogue box that pops up. Ensure

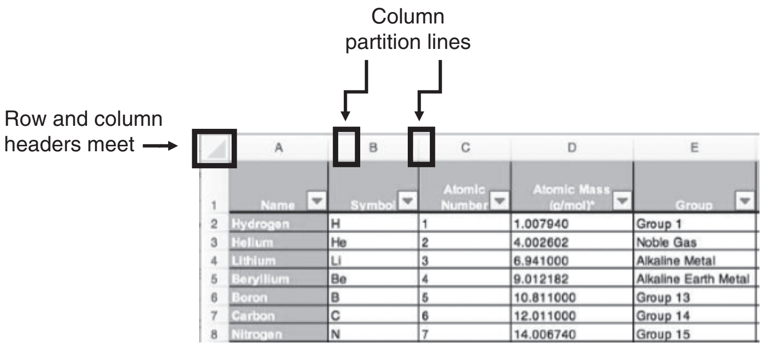


Figure 1.2 Periodic table worksheet highlighting row and column formatting navigation.

Header Row is ticked in the **Table Style Options** under the contextual **Design** tab.

- Next, sort the data in increasing order of atomic radius. Select the greyed icon in G1 to the right of text *Atomic Radius*. Click on **Smallest to Largest** and exit out of the box.
  - Also try sorting the data indifferent ways according to the different properties listed.
- You can visualize the data by creating charts to represent the data. Try graphing *Atomic Number* against *Atomic Mass*. To do this, highlight columns C and D. Click the **Insert** tab and then click **Scatter** chart type as shown in Figure 1.3. This type of chart is very common when working with experimental data.
- To format the chart, select the chart and double click into each **axis title** and **chart title** to edit the text.
- Click on the *x*-axis, and right click and select the **Format Axis** option. Select **Tick Marks** and in the **Major Type** box, and select **Inside** to add tick marks to the *x*-axis. Repeat this for the *y*-axis.
- **Gridlines** are the light grey horizontal and perpendicular lines that divide the chart area into squares to form a grid. To delete these, click on one of the horizontal gridlines, and then right click and select **Delete**. Repeat this step for the vertical gridlines.
- Click through the previews in the **chart styles** to change the layout or style to one you like. Depending on your chosen style, your chart might look something like in Figure 1.4.
- Using the same approach, create charts to visualize the dependence of electronegativity and atomic radius on atomic number. Decide yourself on the chart type and format and design that you use.
- Save and close the *1.1_Periodic Table.xls* workbook.

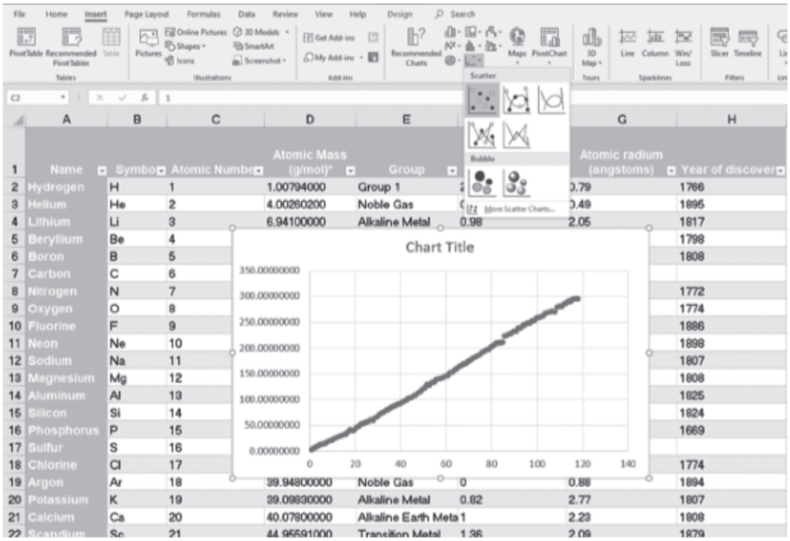


Figure 1.3 Generating a scatterplot in an Excel worksheet.

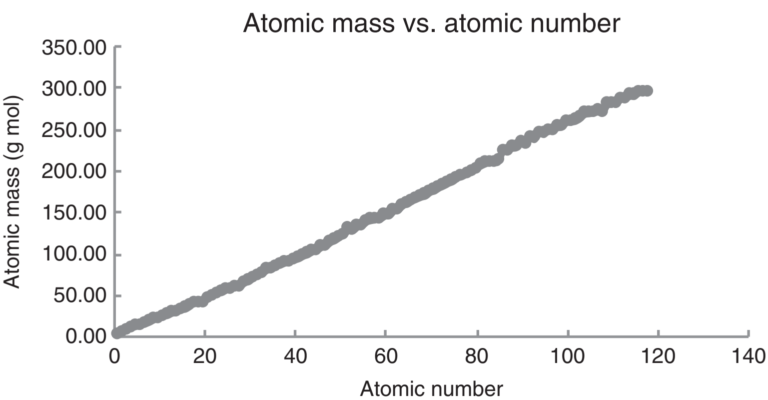


Figure 1.4 Formatted chart showing the linear relationship between atomic mass and atomic number.

1.2 Mathematical Operations on Cells

1.2.1 Formulas and Functions

Once data is entered into a worksheet, operations can be performed to process the data. Excel performs mathematical operations using formulas and functions. Formulas can be written into the formula bar and always begin with an equals sign (=).

These formulas and functions act on specified cells in a worksheet, where variables can be defined in other cells that are referenced. There are two types of cell references used by Excel: relative and absolute. Relative and absolute behave differently when copied and filled from other cells. Using a letter-number combination, e.g. A2, to describe a cell is known as relative referencing. By default, all cell references are relative references. These references change based on position relative to the original cell when the formula is copied and pasted into another cell. The effect is to keep the relative addresses between cells referenced in a formula, in effect making these variables.

In contrast, absolute referencing uses the format $letter$number, e.g. A2, and remains constant when copied and filled from other cells. If the absolute reference A2 had been used as the address, then this address is maintained in the formula across all cells, effectively rendering it a constant (the value of the number in cell A2).

The following tutorials have examples of using both relative and absolute referencing.

Tutorial 1.2 Entering a Simple Formula into a Worksheet

In this tutorial, you will generate model temperature data and convert it from Celsius to both Fahrenheit and Kelvin using relative referencing.

- Open a new workbook and name as *1.2_Temperature.xls*.
- To set up the worksheet, enter the titles *Celsius*, *Fahrenheit*, and *Kelvin* in A1, B1, and C1, respectively.
- Adjust column widths A-C so all titles are visible.
- Bold the titles by highlighting and click on the **Bold** icon under the **Home** tab.
- Enter the centigrade temperature range from 0 to +100 in increments of 5 into column A according to the **Fill→Series…** technique as described in the sub-bullets here:
  - Enter *0* into A2.
  - On the *Home* tab, in the **Editing group**, click **Fill→Series…** to open up the Series dialogue box (Figure 1.5).
  - Select **Series in** as Columns and **Type** as Linear.
  - Use *5* as **Step value** and *100* as **Stop value**.
  - Press OK.

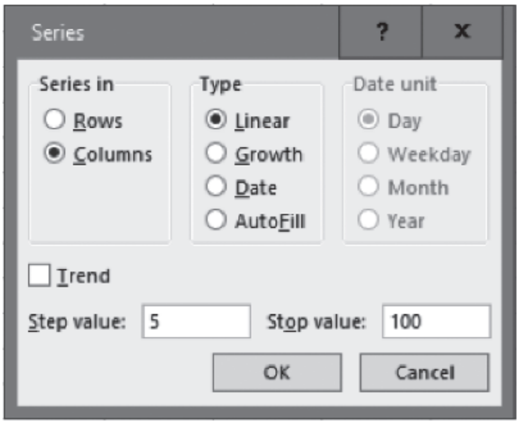
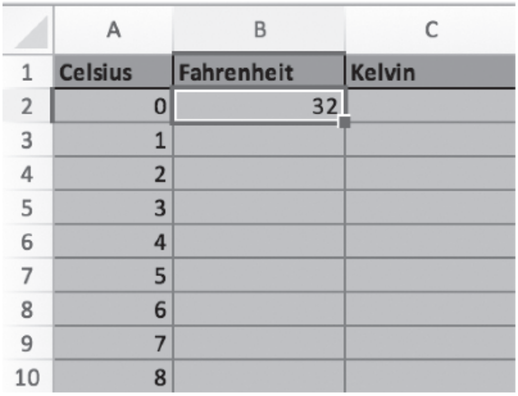


Figure 1.5 Series dialogue box for inputting detail for generating data series.

- In B2, enter the conversion formula from Celsius to Fahrenheit, starting with an equals sign, and using A2 as the centigrade variable = *(A2*9/5)+32*.
- Fill all corresponding values for data in column A into column B by hovering the cursor over the small square on the bottom right corner of B2 (known as the fill handle) until it becomes a black cross and double click on your mouse (Figure 1.6).
- In C2, enter the conversion formula from Celsius to Kelvin, again starting with an equals sign, and using A2 as the centigrade variable = *A2+273.15*.
- Fill down the column as before to report all Kelvin values.
- Save and close the *1.2_Temperature.xls* workbook.



	A	B	C
1	Celsius	Fahrenheit	Kelvin
2	0	32	
3	1		
4	2		
5	3		
6	4		
7	5		
8	6		
9	7		
10	8		

Figure 1.6 Conversion of cursor to black cross symbol for auto-filling cells.

It is important to note that all Excel formulas follow the same rules of algebra, regarding the order of operations. If there is more than one set of brackets or parentheses, the inner-most set will be computed first. Exponent operations will then be calculated. Multiplication and division calculations will be performed next. Finally Excel will then complete any addition and subtraction in the formula.

It can be a good practise to use brackets whenever you can in Excel formulas to structure the equation, even if the use of brackets is superfluous. The use of brackets can help you not only avoid calculation errors but also better understand the formula you are applying.

1.2.2 Entering Functions

Functions in Excel are accessed through the **Formulas** tab under **Insert Function.** A comprehensive range of mathematical, statistical, and scientific functions are available. These all have the general syntax:

= FunctionName(arguments)

For example, = SIN(number) calculates the sine of a number (where the number is an angle in radians) and = SUM(number 1, number 2,…) calculates the summation of the numbers in the cells defined by the argument. Functions can be entered into cells using the **Insert Function** button, or by typing the function directly into the Formula Bar or cell. Here, we will use the **Insert Function** dialogue box to write some formulas.

Tutorial 1.3 Entering a Function into a Worksheet

In this tutorial, you will use the functions AVERAGE and STDEV to describe a set of replicate experimental data.

- Open the workbook *1.3_Sensor Repeatability.xls*. You will see a set of data relating to the anodic current responses of a platinum electrode to 10 repeated measurements of a standard solution of hydrogen peroxide (5 mM).
- Calculate the mean and standard deviation of the data using the AVERAGE and STDEV functions:
  - Click on B13 and then click **Insert Function** under the **Formulas** tab.
  - In the search box that pops up, search for the function *AVERAGE*, highlight it, and double click.

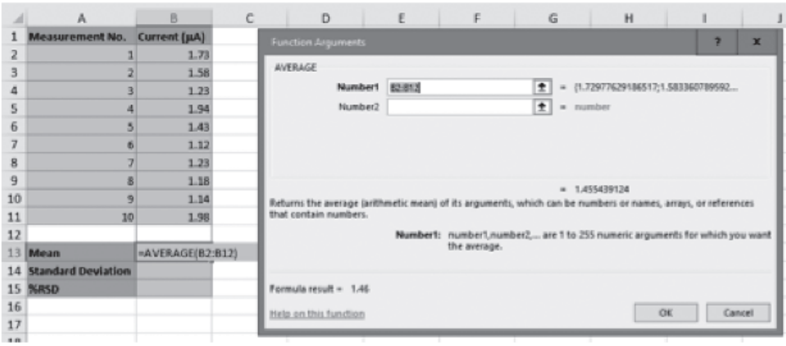


Figure 1.7 Entering a function in a specific cell worksheet.

- With the cursor in the box *Number1*, select B2:B11. The result will be returned in the dialogue box. Look to the bottom of the dialogue box pop-up for specific information on the selected function and arguments (Figure 1.7).
- Press **OK** to enter the calculated value into B13.
- To calculate the standard deviation of the set of data, select cell B14 and following the same steps as above but this time search for and select the STDEV function. You should return a value of 0.33236205. In order to reduce down the number of decimal places used, select the cell and right click. From the drop-down menu, select **Format Cells**. Then select the **Category:** Number. Enter the number of decimal places you require (three in this case) and press **OK** to give just three significant figures in B14.
- Enter the relative standard deviation (RSD) in B15 by typing the formula = *B14/B13*100*. Reduce down the number of decimal places to 1 according to the instructions given above.

• Save and close the *1.3_Sensor Repeatability.xls* workbook.

Note: As you become more accustomed to particular formulas, you can enter them directly, without the need for the **Insert Function** dialogue box. For example, the AVERAGE function can also be entered directly into a cell by typing = *AVERAGE(B2:B11)* in B13 and pressing **Enter**.

Functions in Excel can also be nested, which means placing one function within another. Generally, this is more of a requirement in logic decisions than mathematical calculations. However, they do have their use in scientific data processing. For instance, calculating a formula based on a function inside another function, e.g.

= STDEV(AVERAGE(A1:A5),(AVERAGE(B1:B5),(AVERAGE(C1:C5)) calculates the standard deviation of the set of average values taken from the 3 columns of data.

The following tutorials introduce more functions as well as nesting calculations within functions.

Tutorial 1.4 Using Nested Functions

In this tutorial, you will transform angle data from degrees to radians units in order to compute the sine function for the angles.

- Open a new workbook and save the file as *1.4_Degrees.xls*.
- In the first worksheet, in cell A1, enter the title *Degrees* and below this enter the x range (0–360°) in column A, incrementing every 10°.
- As the sine function, $y = \sin(x)$, stipulates that x must be in radians, the data must be converted to radian units. Enter the title *Radians* in B1. Convert the degree values in column A to radian values in column B using the RADIANS() function available via **Insert Function**. Once the formula is entered in B2, fill down in column B to convert the entire data set.
- In C1, enter the title *Sin(x)* and calculate $\sin(x)$ in C2 with the **Insert Function** dialogue box, using B2 as the value for x in radians. Fill this formula down column C over the full data range using the default relative referencing.
- The alternative here is that these arguments can be nested to merge the steps of calculating radians and sine function together. To demonstrate this, in D1, enter the title *Sin(x)_Nested*. In D2 select the SIN function using **Insert Function** and type the function *RADIANS(A2)* into the arguments box. Alternatively, you can directly type = *SIN(RADIANS(A2))* into the cell. Exit out of the dialogue box and fill down the column. By nesting the arguments in this way, the worksheet needs only contain two columns of data.
- Save and close the *1.4_Degrees.xls* workbook.

Using nested arguments is a matter of preference in Excel. Performing calculations in a stepwise fashion in columns, rather than combining several transformations in a single step, can have advantages when it comes to troubleshooting calculations.

Tutorial 1.5 Entering Functions for Template Design

In this tutorial, you will get more practise entering formulas using relative and absolute referencing by designing a template for assigning elemental composition in organic compounds.

- Open the workbook *1.5_Organic Compounds.xls* where you will see a template setup with three tables. The chemical formulas for methanol, ethanol, and acetic acid are entered in the first table. Populate the rest of this table by entering the number of carbons in each of the corresponding compounds in column C, the number of hydrogens in each of the corresponding compounds in column D, and the number of oxygens in each of the corresponding compounds in column D.
- Now in F1, enter a formula to calculate the molecular weight of methanol. You should use relative referencing to address the number of C, H, and O's and then absolute referencing to address the cells with the relevant atomic mass values. Therefore, the formula entered in F1 should read = *C2*N2+D2*N3+E2*N4*.
- Fill this formula down column F. Notice that the cell references for the number of atoms will change down the column (relative referencing is used as these are variables) but the cell references to the atomic mass values remain fixed (absolute referencing is used as these are constants).
- Next populate the second table with relative composition information. Enter the formulas to calculate the percent composition of C, H, and O for methanol in row B. Enter = *C2*N2/F2*100* in H2, again taking note of the use of relative and absolute referencing. Enter corresponding formulas for %H and %O also. In K2, sum up the percentages across the row (H2:J2) and the value returned should be 100. Now highlight cells H2:K2 and fill down the table to populate the % compositions for the rest of the compounds.
- Report just two decimal places in the cells in the second table by setting this number to two.
- Your final worksheet should look something like in Figure 1.8, depending on how you decide to format it.
- Save and close the *1.5_Organic Compounds.xls* workbook.

	Formula	C	H	O	Total Mass
Methanol	CH3OH	1	4	1	32.042
Ethanol	C2H5OH	2	6	1	46.069
Acetic Acid	CH3COOH	2	4	2	60.052

%C	%H	%O	Total
37.49	12.58	49.93	100.00
52.14	13.13	34.73	100.00
40.00	6.71	53.28	100.00

Atomic Mass	
C	12.011
H	1.008
O	15.999

Figure 1.8 Populated template tables for assigning elemental composition in organic compounds.

1.3 Charts

Graphing experimental data on charts is routinely used as a way of understanding and visualizing data. Originally, Excel was not the platform of choice for performing this function, as the graphing options were clearly designed for financial analysis and many basic operations required by the scientific community were not offered. However, recent versions have redressed this situation and Excel can now easily cope with scientific requirements and offers some advanced features such as optimization modelling using the add-in ***Solver***, which, as we shall see in Chapter 6, can be used for advanced curve fitting. While Excel does not challenge the features offered in specialized statistical and mathematical packages, it has broad applicability and requires the user to enter the mathematical functions to be performed. From a teaching point of view, it provides a powerful tool for teaching undergraduate science students and for helping them explore graphically the dependence of various parameters in equations.

1.3.1 Creating Charts

Charts are used to represent data visually and typically take the form of a graph, a diagram, or a table. There are several chart types in Excel including pie, column, bar, area, and scatter charts. The scatter chart is often used to construct a graph. Excel charts are created using commands under the **Insert** tab and can be edited in the **Chart Design** and **Format** contextual tabs, allowing you good flexibility to tailor every aspect of your chart. Charts can be embedded in a worksheet or placed in a separate chart sheet under its own Sheet tab.

Tutorial 1.6 Constructing a Scatter Chart
In this tutorial, you will generate some model data and use a scatter chart to visualize the data.

- Open up a new workbook and name the file *1.6_E* vs. *Time.xls*.
- To set up the worksheet, enter the titles *Time (s)* and *E (V)* in cells A1 and B1, respectively.
- Enter *0* into A2.
- To generate your model data for the *Time (s)* column, select the range A2:A35 and use the **Fill→Series…** technique (see Tutorial 1.2) to bring up the **Series** dialogue box. Tick the **column** and **linear** choices and enter a **Step Value** of 30 to increment every 30 s.

- To generate corresponding values for $E(V)$, enter *0.1* in B2. Select the range B2:B35, select **Fill→Series...** and tick **Column** and **Linear** options. Enter a **Step Value** of 0.005.
- The data set is now ready to be plotted. Select the range A2:B35 or alternatively, select the two column headings A and B. Selecting the data range before defining the chart type is just one route to constructing a graph and it simplifies the process. The column selection technique should only be used when no other data except that to be graphed are contained in the columns; otherwise editing the data series will be necessary.
- Click the **Insert** tab, and then click the **Scatter** symbol and select the chart option **Scatter**. This type of chart is very common when working with experimental data.
- The chart will appear as an embedded chart in the worksheet. Using the **Chart Design** and **Format** contextual tabs, the chart can be customized and formatted (Figure 1.9).
- Add tick marks to the axes using the **Add Chart Element** command under the **Chart Design** tab. **Select Axis→More axis options**. In the **Format Axis** dialogue box that pops up to the right of the screen, click on **Tick Marks**. Select **Inside** for **Major Type**.
- Add axis titles in a similar manner using the **Add Chart Element** command. Select **Axis Titles→Primary Horizontal**. A default title 'Axis Title' will appear under the *x*-axis. Click into this textbox, clear the text and enter *Time (s)*. Similarly, add the appropriate axis title to the *y*-axis.
- Remove the gridlines using **Add Chart Element→Gridlines**. Unclick **Major Primary Horizontal** and **Major Primary Vertical**.
- Change the default title text to '*Model Potential Data*' by clicking into the title text and editing it.
- Your scatter chart should look similar to Figure 1.10, depending on your selected formatting options.
- Many other aspects of the chart can be customized using the **Format** contextual tab. Explore the formatting options here. Under the **Chart Design** tab, there are pre-set **Quick Layouts** and **Styles** you can use also. Experiment with formatting of your chart using these functions to

Figure 1.9 Chart design contextual tab showing chart style options.

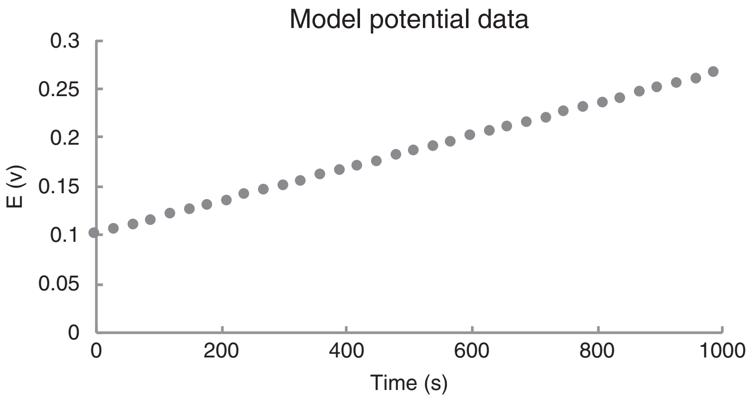


Figure 1.10 Scatter chart style for plotting data.

understand the full design capability of Excel. Settle on a style you are happy with.

- Save and close the *1.6_E* vs. *Time.xls* workbook.

1.3.2 Charting Mathematical Functions

This section reinforces techniques introduced earlier through graphing common mathematical functions that are routine for Excel users from scientific backgrounds.

Tutorial 1.7 Graphing a Simple Function

In this tutorial, you will generate data and chart the function $y = Ax^2$. Relative and absolute referencing will be used in the generation of the data.

- Open a new workbook and name the file as *1.7_y = Ax^2.xls*.
- Enter the titles *x* and *y* in cells A1 and B1, respectively.
- Enter the title *Constant A* in C1, and enter *2* in C2.
- Enter the value *0* in A2 as the first value of *x*.
- To enter a data series in column A, first highlight A2. In the **Editing** group under the **Home** tab, click **Fill→Series**· · · Ensure **Columns** is selected and enter a **Step value** of *5* and a **Stop value** of *500*. Press OK and data is generated for x with an increment of 5 over the required range.
- Split the screen by going to the **View** tab and selecting the **Split** command. This can be a useful way to view different areas of large data sets simultaneously.

- Drag the vertical divider bar over to the far right until it disappears, as we will not be using it in this exercise.
- Scroll down in the lower screen portion until the final row in the series is in view (i.e. row 102).

- Enter the formula for y in cell B2 (= *C2*A2^2*). The caret, ^, represents raising the preceding number to the power of the following number. In this case, the preceding number is given by the relative reference A2 (i.e. $x = 0$). Press **ENTER** to execute the calculation and the result 0 should be returned in B2. As an alternative approach to entering a formula, instead of typing a cell reference into a formula as before, click on the cell to enter the cell reference into the formula. In this case the process is as follows:
  - In B2, enter =.
  - With the cursor still flashing in B2, select C2 and press **F4** to change this from a relative to an absolute reference. (Continued pressing of **F4** toggles through four variations of a cell reference, from absolute column and absolute row to relative column and relative row.)
  - Type the multiply operator *.
  - Select A2.
  - Type ^*2*. Press **ENTER**.
- Fill all corresponding values for data in *x* (Column A) into *y* (Column B) by hovering the cross-hair over the bottom right corner of B2 until it becomes black and then double click.
- To chart the data as a graph, highlight the full range A1 to B102 by clicking on A1 in the upper split pane, hold down the SHIFT key, and click on B102 in the lower split pane. Then, under the **Insert** tab, select the scatter chart option with a sub-type of line chart (without markers).
- To perform further formatting, double click into the element of the chart that you want to format and work through the dialogue box that pops up to the right of the worksheet. For example, change the *x*-axis scale to 0–500 by double clicking on the *x*-axis to bring up the **Format Axis** dialogue box. Use *500* as your maximum value under **Axis Options→Bounds**. Alternatively, you can edit or format a chart by selecting the chart, and under the contextual tab **Chart Design**, all formatting options are available, e.g. chart elements can be edited under **Add Chart Element**.
- When you are finished formatting your chart, save and close the workbook *1.7_y = Ax^2.xls*.

Tutorial 1.8 Adding Additional Plots to an Existing Chart

In this tutorial, you will learn how to add additional data series to an existing graph.

- Make a copy of file *1.7_y = Ax^2.xls* and rename as *1.8_y = Ax^2.xls*. In this new workbook, highlight columns A, B, and C and copy and paste them into columns E, F, and G, respectively.
- Click on F2 and modify the formula so the reference for the constant A is now G2 and fill down the column.
- Change the value of A in G2 from 2 to 6. Check the data in column F changes as a result of changing this value.
- In the chart, assign a legend to the existing series. You can add a legend using **Add Chart Element** under the **Chart Design** tab. The default name for that data series will be *Series 1*. In order to edit the legend text, right click on the chart and click **Select Data**. A dialogue box called **Select Data Source** should appear. This contains the source data of the chart. Under Legend Entries (Series), select 'Series 1' and select **Edit.** An Edit Series dialogue box will appear (Figure 1.11), and under Series name, type $A = 2$, and press **OK.**
- Now, add the additional series $A = 6$ to the chart. Select **Add** in the Legend Entries (Series) box. Enter $A = 6$ as the Series Name. In the Series X values box, click on the data source button and highlight the data in column E. Press the data source button again to return to the dialogue box. In the Series Y values box, include all the data in column F. Press **OK.**
- Press **OK** again in the **Select Data Source** box.
- Repeat the steps earlier to add a third series to the chart where $A = 10$.
- To edit the default fonts used in the chart or axes titles (e.g. use superscript for the 2 in the chart title), highlight the text you would like to format, and right click. Select **Font…** and a dialogue box will pop up where you can select the appropriate format command. After formatting, your chart should look something like that in Figure 1.12.
- Save the file as *1.8_y = Ax^2.xls* and close the workbook.

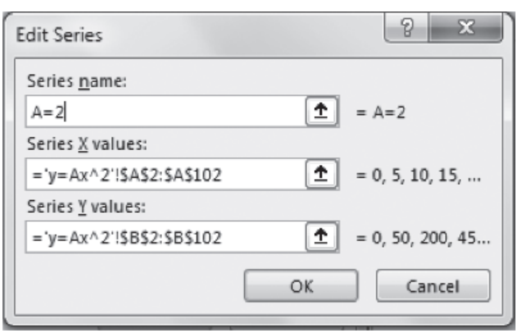


Figure 1.11 Edit Series dialogue box.

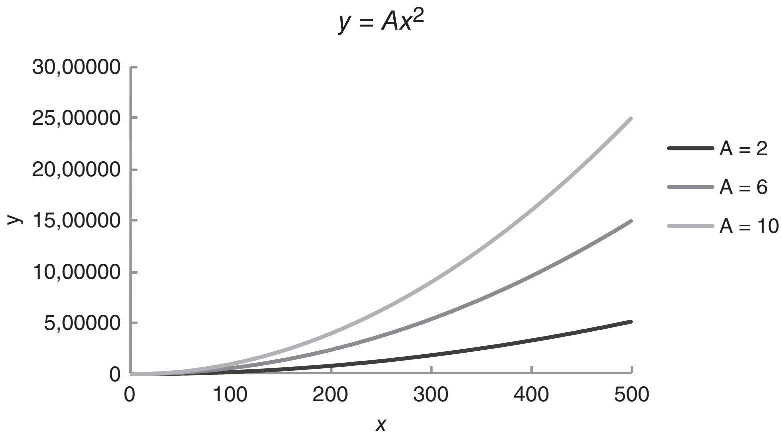


Figure 1.12 Chart plotted with additional data series.

1.3.3 Linear Regression

Trendline in Excel is the tool that is used to add a best-fit regression line to your data. The type of trendline that you choose depends on your data. Trendlines that you can choose from in Excel include linear, exponential, logarithmic, and polynomial. Linear trendlines are often applicable, but much data can be described more effectively with other types.

Adding a trendline to data in Excel enables you to visually see if an experimental data set has a linear (or other) fit. You can label the trendline, edit its properties and forecast the trend beyond the data range if required. You can set the intercept value and output the equation of the line with a corresponding Pearson's correlation coefficient (R^2).

Tutorial 1.9 Performing Linear Regression on a Set of Data by Insertion of a Trendline

In this tutorial, you will apply a linear regression model o experimental data.

- Open the workbook *1.9_E* vs. *Time.xls* to see a set of data for the time dependency of the measured voltage of a galvanic cell. The data is already charted and you should see a clear linear trend in the data points. Activate the chart by clicking on it.

- To add a trendline, in the **Design** tab, click **Add Chart Element→ Trendline→Linear**.
- Double click on the trendline to bring up the **Format Trendline** dialogue box to the right of the screen.
- Click on **Trendline Options** and select **Linear**. Also, tick **Display Equation on chart** and **Display R-squared value on chart**.
- A linear regression equation and the R^2 value should appear on your chart. The trendline can be further customized and formatted from within the **Format Trendline** dialogue box. Spend some time working through the options to find a style (e.g. colour, thickness, etc.) of trendline that can be seen clearly when overlaid on the data.
- Save and close the *1.9_E* vs. *Time.xls* workbook.

1.4 Summary

This chapter demonstrates many of the basic aspects of functionality in Excel that should be of interest when processing experimental data. It is not intended to be exhaustive, but rather highlights the more common data processing basics used by the scientific community. In the proceeding chapters, you will employ these basic operations when learning more specialized features in Excel as they relate to processing and analyzing scientific data.

1.5 Further Exercises

1.5.1 Stoichiometry

Excel worksheets can be used as templates to perform quantitative chemistry calculations. For example, based on a balanced chemical equation, you can use Excel to calculate the amount of a product substance that will form if beginning with a specific amount of one or more reactants.

For example, if 5 g of Fe_2O_3(s) is mixed with an excess of CO(g), how many grams and molecules of Fe(s) and CO_2(g) will form according to the following equation?

$$Fe_2O_3(s) + 3CO(g) \rightarrow 2Fe(s) + 3CO_2(g) \tag{1.1}$$

- Start by generating the table shown in Figure 1.13 in a worksheet.
- Using simple formulas, solve for the number of moles of Fe_2O_3(s).
- Convert all molar values into scientific notation and present all relevant values within the table to have three decimal places.

	Stoichiometry	Grams	Molecular Weight	Moles	Molecules
Fe_2O_3(s)	1	5.000	159.690		
CO (g)	3	Excess	28.010		
Fe (s)	2		55.845		
CO_2 (g)	3		44.010		

Figure 1.13 Tabulated data for computing stoichiometric quantities of products in equation (1.1).

- Use this template design to investigate how changing your starting value of Fe_2O_3 from 5.00 to 6.00 g effects the products.

1.5.2 Sine Wave

Many applications involving trigonometry in chemistry require use of the sine wave function. It can be used to model sound, light, and electromagnetic waves. In this exercise you will chart a sine wave to visualize some aspects of it.

- Open the workbook *1.6.2_Sine Function.xls* and generate a scatter chart plotting the sine function (Sin θ) against θ^{o}. (In order to select the relevant data, you can use the CTRL button to highlight columns A and C only.) Note that the title in column C is used as the default title and legend in the chart – you can change this if you like.
- The function is periodic, repeating itself every 360° (or 2π radians) and is defined for any angle, positive or negative. It has an amplitude of 1, i.e. it ranges from −1 to 1.
- Add a second data series to the chart with a phase shift of 90°. Hint: You will need to generate a new set of θ values in another column ($= \theta + 90°$).
- Now create a new chart where the sine function plotted against θ^{o} has twice the amplitude of the original function (2Sin θ). What is the effect seen in the plot?
- Create another chart where the sine function plotted against θ^{o} has half the period of the original function (Sin 2θ). What is the effect seen in the plot?
- Enter axes titles and format your charts before saving and closing the workbook.

1.5.3 Bragg's Law

Bragg's law is one of the most frequently encountered relationships in chemistry that involves a trigonometric function. It describes the angle at which a beam of X-rays of a particular wavelength, λ, diffract from a crystalline surface in which the lattice planes are separated by a distance d. Reflected X-ray beams constructively interfere and so only appear at certain angles, θ. The first of these angles is described by an integer representing the order of the diffraction, n. Thus for first order reflection, $n = 1$; second order reflection, $n = 2$; etc. The Bragg equation, which governs this behaviour, is:

$$n\lambda = 2d \sin \theta \tag{1.2}$$

In this exercise you will set up a template to investigate the Bragg equation and then use this template to calculate lattice plane distances in a crystal.

- Open up a new workbook where you will set up a template that will allow you investigate the relationships governed by Bragg's equation. Save this file as *1.6.3_Bragg.xls*.
- Enter the parameters n, λ, d, and θ in cells A1:A4. In order to enter Greek symbols, under the **Insert** tab, click on **Symbol** and a dialogue box will pop up. Select Font: Symbol and all Greek character notations will be displayed. Click on the desired character and press **Insert** to enter the symbol into the cell.
- Assign a corresponding set of values for n, λ, d, and θ, in B1:B4 in the first instance where d is the unknown. For n, λ, and θ, assign values of 1 initially. In the cell assigned to d, enter a version of equation (1.2) that allows you to solve for d. Be very careful about the use of your brackets!! Use absolute referencing when referring to the cells containing values for n, λ, and θ. Give this set of cells a title such as 'Calculation of d given θ'
- Similarly, assign another set of cells elsewhere in the worksheet for values for n, λ, d, and θ, where θ is the unknown. Here, you will need to calculate the inverse sine function. To do this, use the function *ASIN(number)*, where number refers to the sine of the angle you want and must have a value between −1 and +1.
- Use this template to calculate the lattice plane distance (d) in a lithium fluoride (LiF) crystal where the first order reflection from X-rays of wavelength 0.707 Å occurs at 34.68°. Based on these same conditions, calculate the angle at which X-rays will be diffracted for the second order reflection.

1.5.4 Nernst Equation

A number of analytical techniques require measurement data to be transformed in some manner before a calibration graph can be constructed. One common example is potentiometry, in which the measured response (electrode potential, E) is related to the logarithm of the corresponding ion activity (a) or concentration (C) via the Nernst equation [equation (1.3)]. For a simple electrochemical cell involving a single metal species being reduced at an electrode, the equation for the reaction for $M^{n+} + ne^-$ can be written as:

$$E = E^o + \frac{RT}{nF} \ln C_{M^{n+}} \quad (1.3)$$

where

E Measured potential, V
E° Formal potential, V
R Universal gas constant, 8.314 JK^{-1} mol^{-1}
T Temperature, K
F Faraday constant, 96485.3 c mol^{-1}
${C_M}^{n+}$ Concentration of metal species, M^{n+}, M

- Open a new workbook and name as *1.6.4_LOG and LN.xls*.
- Set up a table of constants somewhere in the worksheet as below (Figure 1.14):
- Enter the column heading ${C_M}^{n+}(M)$ in the top row of a column. Create a series of concentration values from 0.05 to 1.00 M in increments of 0.05 going down the column.
- Enter $ln({C_M}^{n+})$ in the top row of the column to the right and calculate the natural logarithm of each of the concentration values in the previous column.
- Enter *E(V)* in the top row of next column to the right and write the formula for equation (1.3) for E in the cell below. Fill the cells down to complete the column of calculated electrode potentials. Make sure to use absolute referencing when entering addresses for each of your constants.
- Create a chart by plotting this electrode potential, E, against ln ${C_M}^{n+}$
- By performing a regression analysis on the charted data, show that the value for E° is 0.337 and the oxidation state, n of the metal species is 2.
- Visualize the dependence of E on ${C_M}^{n+}$ by plotting E against log ${C_M}^{n+}$ according to the logarithmic form of the Nernst equation:

$$E = E^o + \frac{2.3026RT}{nF} \log C_{M^{n+}} \quad (1.4)$$

- Demonstrate equation (1.4) is equivalent to 1.3 by again calculating E° and n and observing that the same values are calculated.

Figure 1.14 Tabulated values for constants in the Nernst equation.

E°(Cu²⁺)	0.337
R	8.31451
T	298.15
n	2
F	96485.3