Data representation and analysis in Excel

Let’s Get Started!
This course will teach you how to analyze data and make charts in Excel so that the data may be represented in a visual way that reflects the nature of data. The nature of Excel program makes cross-platform importing and styling charts much easier, and boasts a fluid and visual interface.

To make a chart, first ensure that you have data.

What is the nature of your data?
Think of the data that you have and how you want to showcase it.

Dimensions
How many groups of data are associated at every single data point? This determines the dimension of your data. If you are trying to represent the speed of a car with respect to time, this would be a 2-D data group. If you are trying to represent the mass and velocity of a launching rocket over a period of time, your data are 3-D. An energy landscape of (X,Y) coordinates and respective energy values is also a 3-D data group. Anything dealing with parts of a whole, or adding up to a certain percentage, is a 1-D data group. You can also create charts that presents to up to 7-D (or more) of data, such as a scalar field in 3-D directions, but most people find hard to register with more than 7 or 8 dimensions.

Chart Types

One Dimensional Chart

Pie
Use a pie graph to show any amount of data
adding up to 100% -- for example, use a pie chart for counting votes. 40% are blue, 24% are red, 26% are green, 10% are purple. Before you do anything, make sure all your data are of the same unit and can add up. You might not want to add the votes from one voting to those of another.

Figure legends are of utmost importance in a pie chart, without which you cannot know which one group of data is which. You can set up a legend by specifying which group is which on the data spread sheet. It would look even better of you can add figures on each slice and a title on the top.

Two Dimensional Charts

The most significant 2 types of 2-D charts are column/bar charts and line/scatter charts. To choose between them, you might want to count the total number of your data points: usually anything above 5 data points requires a line/scatter charter, whereas with a lower number of data points a column/bar chart is better. This is because bar charts are more visual in nature and should convey all their information within one glimpse. Line and Scatter point charts are more analytical. Bar charts are excellent for presentations.

Column / Bar

Column chart and Bar chart are the same kind of charts only with different directions. It is good for data that uses dates or conveys the passage of time.

When you set up your original data table, it is a good habit to have column titles and units. Notice 2-D data have 2 columns. The number of data points would equal to the number of rows.

Always include sufficient data on your charts! For presentation purposes, always include an appropriate explanatory title. Include the units on either the legend or the axis titles. Utilize the axis titles to present what the natures of the data are.
Line/Scatter

A line graph is just a scatter point graph with dots connected and evenly distributing a value along the x-axis. For this reason it is much less used than a scatter point. Use a line/scatter graph to see a trend in one set of data, or to compare the trends of several sets of data. They are also good for graphs using dates or time. To produce a trend, we usually require a decent number of data points. Scatter point charts are widely used in research and industries because of its excellent potential to be analyzed.

A scatter point chart starts with a data table of more than 5 data points. As an example, a free-fall process of 5 seconds with the distance travels measured every 0.5 seconds would have 11 data points. The resulting plot should have two variables: time and distance. Of the 2 variables, the independent variable should be on the x-axis and the dependent variable should be on the y-axis.

When creating a chart for research purposes, several elements are essential to your chart representation: the axis labels with units and figure legend explaining the nature of the graph below the graph. Often a title is not necessary.

Multi-Dimensional Charts

Line/ scatter

Often times our data might have more than one dependents variable associated with the independent variable. For example, in the experiment above, the velocity can be measured at every point in addition to the distance traveled. Such data would have more than 2 columns and would require more than 1 point group/ line to represent.
Heat map/Wind rose

Sometimes our data would have more than one independent variable.

Heat maps are used to present range at given locations, it would have two independent variables: latitude and longitude, and one dependent variable, the temperature. The temperature would presented by a range of colors, usually warmer colors mean higher temperatures and cooler colors mean lower.

Another kind of map is the wind rose. It’s two independent variables are direction and frequency, and the dependent variable is the percentage.

Note the color scheme on the charts. When presenting the figures, it’s almost never a good idea to present a heat map with black and white, as this would dull the distinction between warm and cool colors. On the other hand, if you think your graph might be presented in black and white, consider choosing a monochromatic color scheme.
Inserting a Chart

If you don’t select any data, Excel will create a chart that includes all of the data on your sheet. You can also choose to select data by clicking and dragging, or by holding down the Ctrl key and selecting rows and columns. Make sure your column and row headers are included if you want your data in the chart labeled.

Once you’ve decided whether or not to select your data, go ahead and either select the appropriate data or make sure that no cells are selected.
Click the **Insert** tab. In the **Insert Charts** box, click **Column** and select a type of column chart to insert. Your chart will appear an object in your sheet. You may find it easier to work with your chart if it is another sheet, and you can move it by selecting it (it will be selected by default after you insert it) and typing **Ctrl + X** to cut the chart. Select another sheet by clicking **Sheet 2**, for example, at the bottom left-hand corner of your screen, and type **Ctrl + V** to paste the chart. This way, you can work with your chart and your data without having them overlap. Any changes you make with the data will still be reflected in your chart.

**Chart Tools**

**Modifying Chart Design**

**Add Chart Element**

The Add Chart Element button in the leftmost of the ribbon helps to edit the various elements on the site. Among the elements, Trendlines and Error Bars are very useful in data analysis and could be used in scatter point graphs.

**Axis Labels**

In the **Design** toolbar, click **Add Chart Element > Axis Label** to label the axis on which your data is displayed. Always remember to add the units.

**Legends**

In the **Design** toolbar, click **Add Chart Element > Legends** to label the data sets with corresponding colors. This is only necessary when the is more than one point group and the colors need to be differentiated.

**Trend lines, equations and R-squared**

In the **Design** toolbar, click **Add Chart Element > Trendline** to add a trendline to a given data set. Once the trendline is added, the **Format Trendline** box should appear on the right. You can also bring out the box by double clicking your trendline on the graph.

The Format Trendline box controls which algorithm would be used to define the data set, as well as the different elements to be display on the graph.

To adjust the elements, go to the **Trendline options** tab.
The **Linear** option should be used if your dependent variable and the independent variable are directly proportional with each other. It would display a straight line on the graph.

The **Polynomial** option can be used if the dependent and the independent variable are related in the certain order. I.e, the time and distance traveled example on page three. The order depends on the number of peaks in the data set, and always equal to a maxim/minimum +1.

The **Moving Average** option is most often used in statistical samples when each data point have no correlation with the independent variable.

Sometimes the graphs call for Equations and R-squared values. These can be displayed by checking the **Display Equation on chart** and **Display R-squared value** on chart options.

**Error Bars**

In the **Design** toolbar, click **Add Chart Element** > **Error Bars** to label the data sets with corresponding Error bars. With a given data set, Excel can calculate the standard deviation of the set and label it on the graph. Longer error bar would indicate larger more deviation to an average value. Similar to **Format Trendline** box, the **Format Error Bar** box can help you adjust the kinds of error bars or even help define a custom amount.

**Select Data**

If you change your mind about the data you selected and would like to select different data, you can accomplish this by clicking on **Select Data** in the **Design** ribbon. Here you can decide what needs to go in the chart. In the **Legend Entries (Series)** column, you can decide how many sets of data you want to have in the chart. Most will automatically have 2 series set, even if you only select one series of data. Click on a Series, and then the Edit button. You can change the name of what each line, bar, or plot point will be. You can also add more by clicking the Add button. In the **Series name** box, type = “Name of Series”. In **Series values**
box, click and drag on the data you want from your dataset, and it should automatically add to your chart. Any series can also be deleted with the Remove button. In the Horizontal (Category) Axis Labels column, you can select what labels will go along the x-axis: just click and drag the on the cells that you want. Below both columns is the Hidden and Empty Cells button, which will let you decide how you want to deal with cells that have no data in them. You can either choose to have it show as empty space (Gaps) or counted as a number (Zero).

**Move Chart Location**
In the Design toolbar, click Move Chart Location if you would like to move your chart to a new sheet or make it an object in an existing sheet. This can also be accomplished by selecting the chart and cutting and pasting it to a new location.

**Change Colors**
In the Design toolbar, click Change Colors if you would like to adjust the default color scheme of your graph. All graphs in one manuscript should have a consistent theme, ie, different shades of blue. Excel also offers a variety of monochromatic scheme that’s compatible with black and white printing.

**Clean up unnecessary elements**
You can always delete unwanted elements by double clicking on the element until a rectangle appears around the element, then click backspace on the keyboard.

**Charts in Other Applications**
It is possible to insert tables charts in other Microsoft Office applications directly from those applications by using the Insert menu. This is an effective way to insert charts into Word documents or into PowerPoint presentations, but it has some complications, and we recommend that you simply insert your chart into your Excel spreadsheet and customize it appropriately, then simply copy and paste the chart into your document or presentation. It will be possible to modify the chart only by changing its size. Any other modifications will need to be made in Excel, and then you will need to copy and paste the new version of the chart.

**Help**
There are several ways to get help in Microsoft Excel should you find that you have a problem.

PowerPoint Help menu: Click the F1 key on your keyboard or the Help button available at the right-hand side of each ribbon to browse through a collection of help topics or to search for a topic.

Google is an excellent tool so long as you are sufficiently specific in your search terms: “how to change chart layout in Excel 2013” is an example.

Technology Assistants and Consultants and Peer2Peer teachers: We can be found at the Technology HelpDesk in the Info Commons, 4th floor Library. If we don’t know the answer to your question, we know someone else who does!

Prepared by Wanshu Zhu, LITS/CTS/ Peer2Peer Team, September 2014