How to Create Graphs Using Excel

Getting Started: Installing Data Analysis Toolpak
1. Check to see if it is already installed. Go to Tools and if Data Analysis is listed you, are all set. Otherwise, go to Step 2.

2. Go to Tools/Add-Ins and select Analysis ToolPak and then hit OK. You may be directed to insert a Microsoft Office or Excel disk for personal computers. If so, follow the directions on screen. You only need to do this once on a given computer.

Displaying Toolbars in Excel
1. Go to View/Toolbars/Standard (To display icons for commonly used commands)
2. Go to View/Formula Bar (Allows you to see cell entries and formulas)

Creating Various Graphs From Raw Data in Excel

Histograms (Graphically summarizing a numerical variable):
1. In file containing data highlight the cells containing the label and the numerical data. Copy it.

2. Open the file histogram.xls and click on the tab that says Data.

3. Click the cell A1 (labeled with Arch) and paste your data. If your data set has fewer than 13 observations you will need to delete the extra observations from the default data set.

4. Click on the tab that says Histogram and note the Minimum and Range in the Descriptive Statistics Summary at the top of the page.

5. Click the box that says Choose my own Bin Size.

6. Enter the minimum value. You can round off if needed.

7. Enter the bin width. For the bin width, you can try bin width = range/4 or bin width = range /5.

8. Adjust the bin width so that there are not a lot of empty bins or bins with only 1 observation in it. You also don't want too few bars. You should try to have at least 3 or 4 bars. Be sure that you can get a sense of the distribution.

9. Also make sure that your histogram is not missing any of the bins. If this happens, there will be a warning in red to this effect. If your get this warning, increase your bin width.

10. Edit the title and axes labels by clicking on the axes labels and chart title and typing in informative titles with units.

11. You can copy and paste the histogram into Word (or other documents) and resize as needed.

Note. If you delete a data value within the data in column A, you will mess up the underlying code and your histogram will not be correct. Instead of deleting, just start over with a new histogram.xls sheet.
Bar Chart/Pie Charts (Graphically summarizing a categorical variable):
1. First you must construct the frequency table that lists the different outcomes and the number of observations for each outcome. (If data is in summarized in a frequency table you can skip to step 4.)

2. First sort the data. This will make it easier to count the number of observations for each outcome since they will be grouped together. Highlight column and then click the icon that has an A…Z. This will sort one column from low to high.

3. Make a frequency table that has 2 columns, one that lists the subcategories and one that lists the corresponding frequencies.

4. Once you have your frequency table, highlight both columns (column names then summarized data in a box). If your classes are numerical, you should rename them to be more informative and so that Excel will create an appropriate graph.

5. Select the Chart Wizard (Choose Bar, Column or Pie Chart – depending on view you like)

6. Hit Next and then enter titles. Click Next and Finish when done.

7. Delete legend if it is not needed. You can remove the gap between the bars by double clicking on a bar and choosing the Options tab. Change the Gap Width to 0. You may also need to resize your graph.

Scatterplot (Graphically summarizing 2 dependent, numerical variables or a Time Series Plot):
1. In Excel the column containing the explanatory variable (X) needs to be listed directly to the left of the column containing the response variable (Y). Highlight both the X and Y data values. Click on the Chart Wizard on the tool bar (the icon that looks like a bar chart) or select Insert/Chart on the menu bar and hit enter. Choose the XY(Scatter) chart type, hit Next.

2. Only if you forgot to highlight the data (otherwise skip down to the next step): You must enter the data here. For the data range enter the cells containing your data (ex: A1: B23) or highlight the data in the two columns containing your X and Y data. Make sure the series in selection is for columns. Hit Next.

3. For chart title type an appropriate title. Label your X and Y axes with a short description of the data and units (ex: Time (minutes)). Hit Next when you are done with the titles and axis labels and then choose Finish.

4. After the Scatterplot appears, place your cursor over the legend and delete it. You should also change the background from gray to white. To do this click on the chart in the grey area (choose an empty spot), then right click and choose Format Plot Area. Under Area, choose None, then OK.

5. To change the scale of the y-axis (if there is a lot of empty space vertically), place the cursor over any one of the y-axis numbers and double click. Choose the Scale option and change the minimum value and/or maximum value to fit the range of values for Y. Then hit OK. Be sure to select these numbers appropriately so that you don’t accidentally ignore some points on the graph. The x-axis can be adjusted similarly.
6. To get the least-squares line click on the graph. Then go to Chart/Add Trendline and click on the tab that says options. Then check the boxes next to Display equation. Then click OK.

**IMPORTANT GUIDELINES TO REVIEW WHEN CREATING FIGURES**

- Label both axes and provide a heading to make clear what the graph is representing.
- Vertical axes should usually start at 0 to help our eyes compare relative sizes.
- Remove any clutter that isn’t needed or is distracting.
- The axes may need to be resized to remove extra white space.
- Be careful in using unusual bars since it can be easy to get the relative percentages that the figures represent incorrect.
- Sometimes displaying information for more than one group on the same graph can be difficult especially when the values differ greatly. Consider using relative frequencies or separate graphs instead.
- Y axis should be \( \frac{3}{4} \) as tall as X axis
- When the number of score values on X axis is large, scores should be collapsed so there are at least 5 intervals but no more than 12.
- The width of each interval on the X axis should be equal.
- Frequency on the Y axis must be continuous and regular.
- Range on the Y axis and X axis must neither unduly compress nor unduly stretch the data.