

Data Visualization: Some basics

Graphical Options

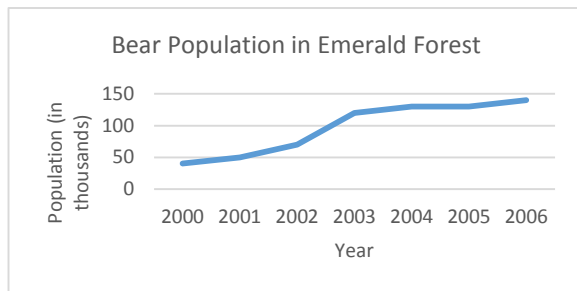
You have many graphical options but certain graphs might be more appropriate depending on the story you want to tell. Here are the most common types of graphs:

Table 1: Types of graphs and their uses

Type of Graph	Uses
Line graph	demonstrating change over time or comparing change over time
Bar chart	comparing differences / similarities between groups
Table	providing exact values
Scatter plot	showing correlation (positive, negative, none) between two variables
Pie chart	illustrating large differences in proportions for simple data

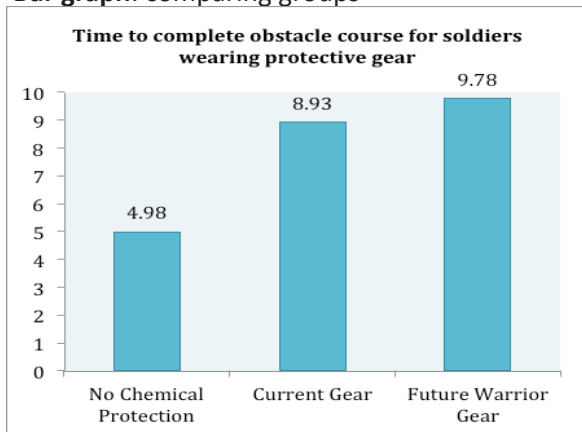
Here are the common types of data visualization with examples and described in more detail:

Line graph: changes over time



Since the unit for the x-axis is Years, we can easily see the growth of the bear population (the y-axis unit) over time by following the line connecting the data points.

Bar graph: comparing groups



A quick glance at this graph shows a large difference between the groups wearing no chemical protection and the groups wearing either the current gear or the future warrior protective gear.

Notice how the gridlines have been removed to eliminate “non data ink” and create a clean look.

Table: presenting exact values

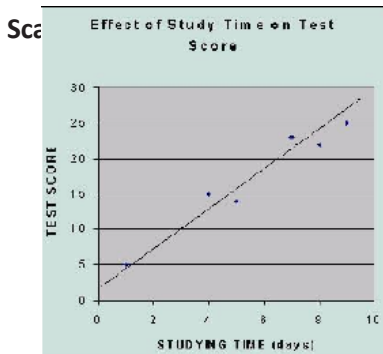
Table 1: Average time to perform task (in seconds)

Task	Firefox 2.0	IE 7.0
Startup speed		
Time to start up from cold	11.6	7.8
Download speeds		
Download benchmark HTML page w/ multiple images	2.0	2.5
Download benchmark Javascript	22.0	36.4

We can easily compare the startup speed between Firefox and IE by glancing at the actual values and knowing that 11.6 is higher than 7.8. It is also useful that we can see not only startup speed in this table, but also download speeds.

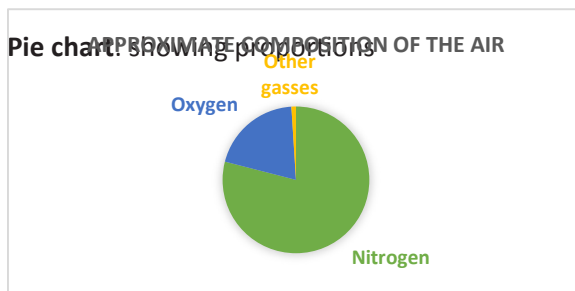
Again, notice how an absence of gridlines simplifies this graph.

Data Visualization: Some basics



In a scatter plot, we can see the spread (i.e., a visualization for how correlated the variables on our axes are); since the points are tight around the line, the correlation is strong. We can also see the general trend that we expect higher test scores for longer studying times.

*It is up to you whether you want to add the trend line to your visualization. Some sources recommend avoiding extra clutter; some prefer the ease of access to the information.



Pie charts can illustrate noticeable differences between a few groups. **USE THIS TYPE OF CHART WITH CAUTION.** It is much easier to distinguish differences in data with a bar graph.

Choosing your story

The graphical representation you choose for your data will depend on the story you want to tell (2).

For example, consider the following data tables. Depending on the way you compile the data, emphasizing total medal count or gold medal count, you could make an argument for whether the US or China “won” the Olympics. Neither interpretation is “wrong” -- they simply create different arguments. First

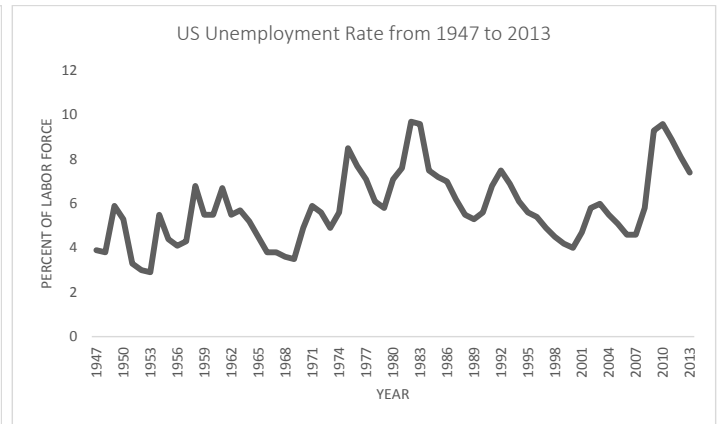
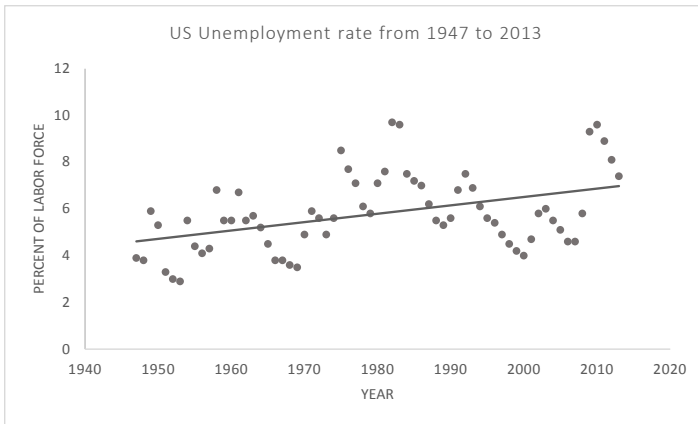
Table 2: 2008 Summer Olympic medals by country. These two tables differ by how the data is sorted but tell two very different stories about who lead the 2008 Olympics.

a) Sorted by total medals					b) Sorted by gold medals				
Country	Gold	Silver	Bronze	Total	Country	Gold	Silver	Bronze	Total
USA	36	38	36	110	China	51	21	28	100
China	51	21	28	100	USA	36	38	36	110
Russia	23	21	29	73	Russia	23	21	29	73
Britain	19	13	15	47	Britain	19	13	15	47
Australia	14	15	17	46	Germany	16	10	15	41
Germany	16	10	15	41	Australia	14	15	17	46
France	7	16	18	41	S. Korea	13	10	8	31

Data Visualization: Some basics

consider your narrative -- then decide how your data visualization can help support it.

Again, consider the below graphs that tell very different stories about the US employment rate over the



Explaining the Data Visualization

Finally, while your data visualize should be clear enough that someone could just look at it and understand the story, you still need to explain the bottom line. You should do this 1)in a caption/title and 2)in the text of your paper.

In a caption OR title, you explain your the basic elements of your visualization:

1. Name the figure or table (Table 1, Figure 1, etc.)
2. Describe the visualization
 - What are the axes
 - What does the data represent

Within the text, you should interpret your visualization:

1. Refer to your table or figure and state the main trend
2. Support the trend with data
3. (if needed) Note any additional trends and support with data
4. (if needed) Note any exceptions to the main trends or unexpected outcomes

For example, consider the interpretation for the table below on internet browser performance.

Average browser performance on speed, memory use, and compliance				
Task	Speed (in seconds)	Memory use (in megabytes)	Compliance (in percent)	JavaScript score
Chrome 31	4.5	261	75%	213.9
Firefox 25	5.3	106	68%	183.1
IE 11	2.3	162	64%	165.3
Opera 17	5.3	200	73%	200.0
Safari 5.1	4.5	130	51%	137.6

Speed was calculated by taking average cold start up, non-cold start up, non-cached page load, and cached page reload speeds. Memory use was calculated by taking average memory for blank tab and 10 open tabs. Compliance was calculated by averaging HTML 5 and CSS 3 compliance rates. See the full report for an explanation of how JavaScript score was calculated

Table 1 suggests that Chrome is the best internet browser as it scored the best on two of the four measures. However, we should note that Chrome had the worst score on memory use. Therefore, we conclude that Chrome is the best internet browser for users whose computers have enough memory to handle it. Users whose computers have lower memory might prefer Firefox 25, while users who want the fastest browser could use IE 11.