

Project Part 1: Improve Data Visualization Skills

1. In your own words, explain three principles of effective data communication from the lecture material.
 - a. Impactful, meaningful charts and graphs
 - b. Easy to Understand
 - c. Levels of detail
 - d. Clear purpose
2. In your own words, summarize the following concepts as they relate to visual perception:
 - Order - the order in which viewers view your graph and which can impact the message
 - Hierarchy - the way in which certain elements call attention to a visualization: bright colors vs muted
 - Relationships - highlighting data connections you want to convey to your viewers
 - Convention - placement (composition) should be standardized for ease of understanding, correct graph choice for the information shared.

Answer the following:

How will an understanding of these concepts help you create better data visualizations?

Use of these concepts will create visualizations that are impactful, easy to understand, and able to convey an intended narrative or point with ease.

3. Given the scenarios below, write which graph would be best to use for the data and what makes it an effective choice:
 - Comparison between values – bar or column
 - Comparison to the whole – columns or stacked columns
 - Change over time – column, line, slope, or area
 - Ranking data – bar, slope, gauge
 - Correlation – plots or bubbles
 - Geographical charts – maps using countries or states
 - Measuring a target – gauge, kpi, numbers

- Showing Outliers - table

How will an understanding of these concepts help you create better data visualizations?

Selecting the best graph for the data ensures that the visualization makes sense, doesn't confuse the reader or detract from the truth of the information.

4. Consider the following quote:

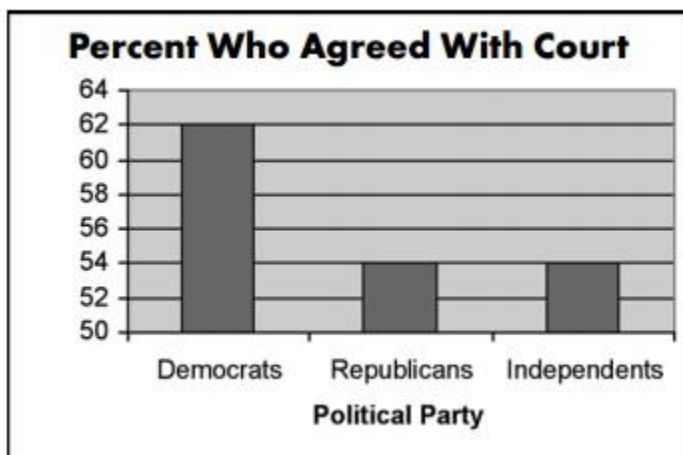
"It is easy to lie with statistics. It is hard to tell the truth without it." - Andrejs Dunkels

What do you think is meant by this statement?

It is easy to misrepresent data by obscuring it with bad or inappropriate visualizations, summaries, and narratives. However, without visualizations it can be hard to express large amounts of data and the information we can interpret from it.

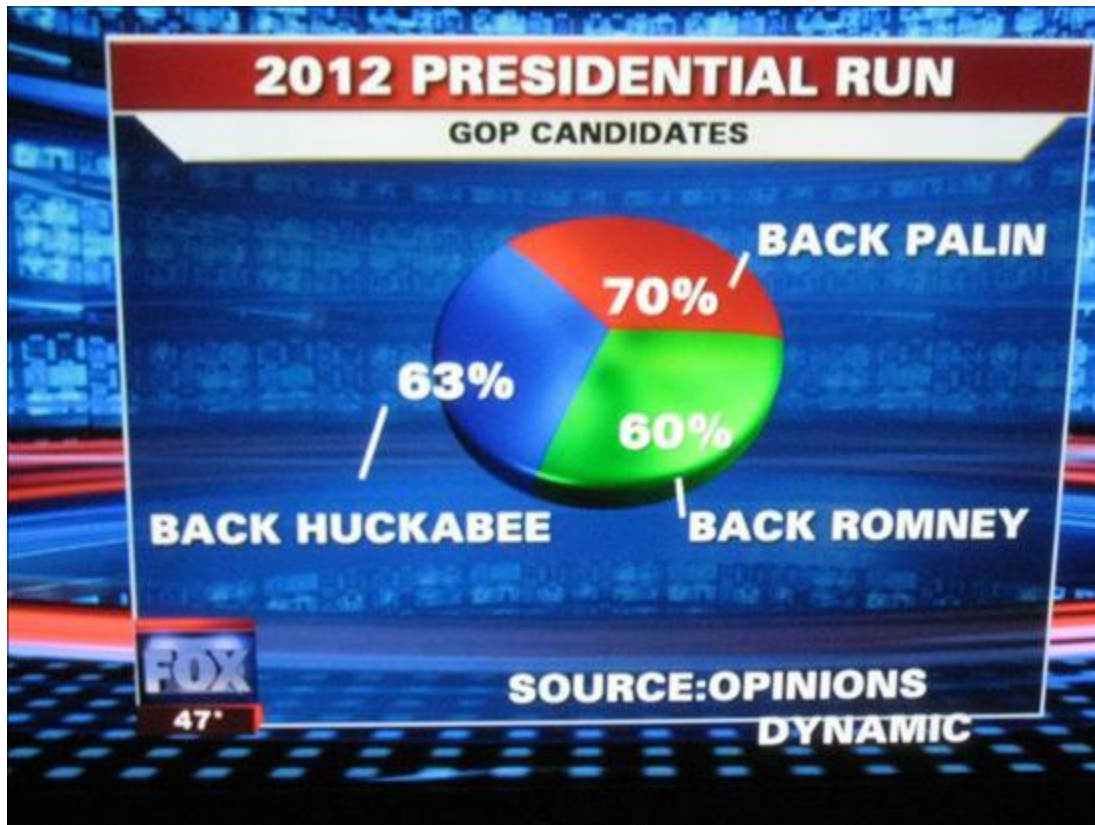
Provide three examples of misleading graphs. Explain what is misleading about the graph in your example and what should be changed to make the graph objective and accurate. Feel free to consult Google for real-life examples of misleading graphs to use as your examples.

Ex 1



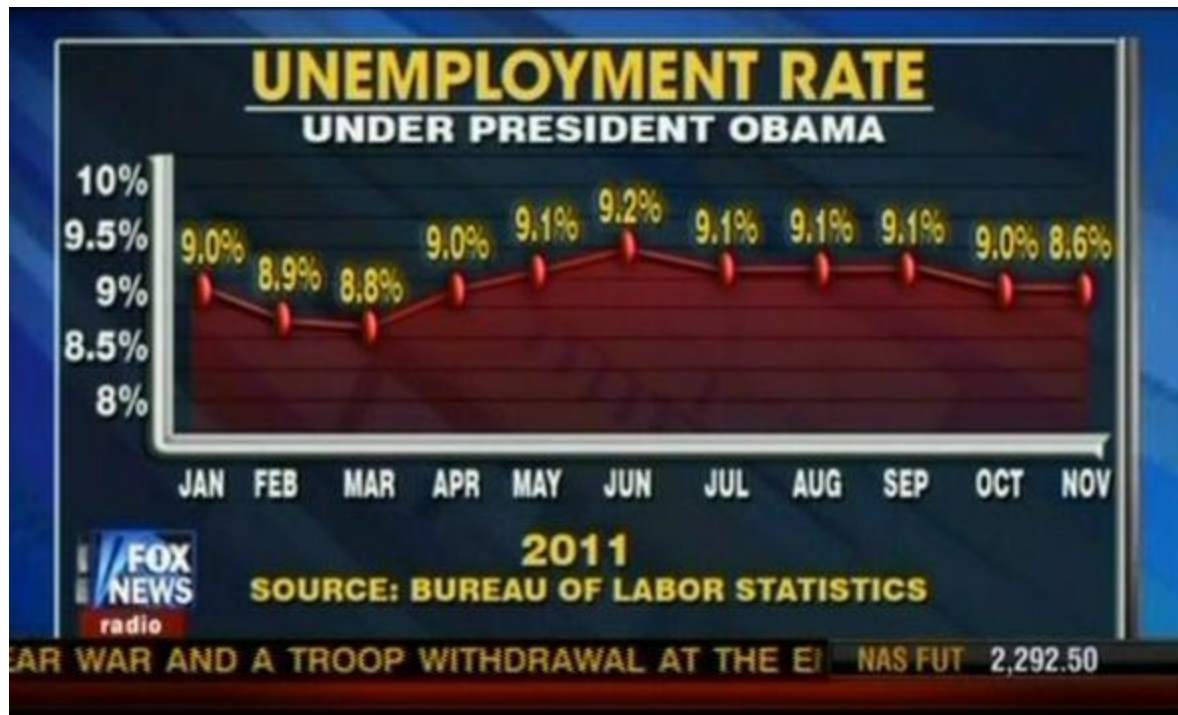
This chart lacks a baseline (0) and is truncated which creates a narrative that far more Democrats agreed with the court than Republicans or Independents. This is because the baseline starts at 50; in truth the different is only 8 percent.

Ex 2



A pie chart should add up to 100, whether that is percentage or absolute. This is plain bad math. Potentially a different chart type or a different narrative (surely those percentages represent **something** in the data) would be a better visual.

Ex 3



This example is particularly egregious. Lacking a true baseline, the graph also randomly omits December which is still included in the year 2011 as far as I'm aware. Additionally, the y-axis and the graph points don't align, showing 8.6 equal to 9.0 and higher than 8.8 (March). I'd say fix it by throwing the whole thing out, but adding a baseline, using true measures and labels to align the data points with the y-axis, and including all representation points along the x-axis would be a good start.

5. In your own words, answer the following questions:

- What is "visualization clutter?"
 - Too many elements that are not supportive of the vital information communicated; can distract viewers
- What are the main components of a graph?
 - Title
 - Subtitle
 - Source
 - Data Label
 - X&Y Axis
 - Legend

- What are three techniques you learned to make data visualizations more clear?
 - Grouping similar objects together
 - Closure
 - Proximity - we create relationships based on object proximity; spacing those objects close together implies relationships and farther apart implies no relationship
- How can the use of color affect the way your visualizations are understood?
 - Too many colors, colors representing unexpected data (red for good, green for bad), or colors that are too similar can confuse or mislead viewers.
-