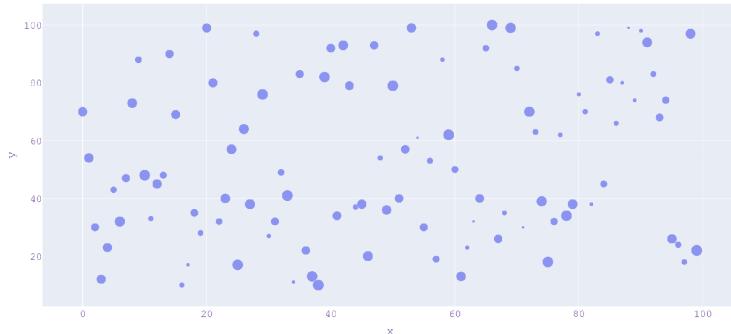


Data Visualization Continued

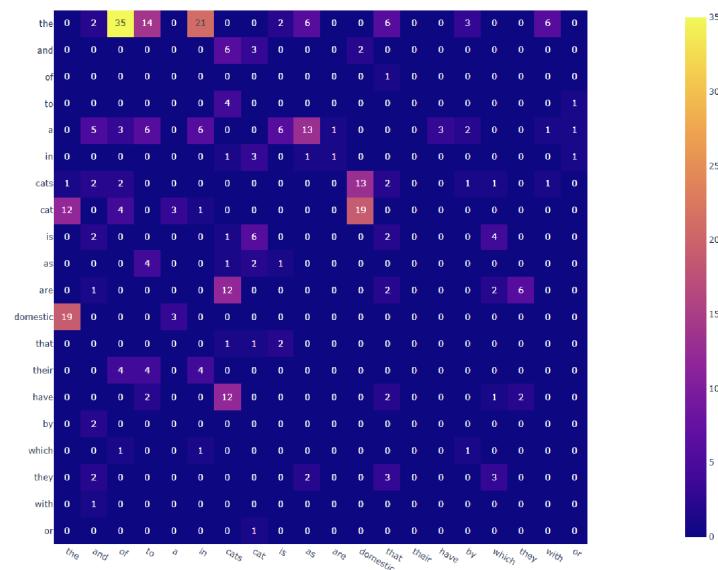
Visualization with Three or More Attributes – Bubble Chart

- Displays three quantitative variables
- Scatterplot with points of various sizes, called bubbles
- Two variables are represented by the x -axis and y -axis, and the third is represented by the size of the bubble



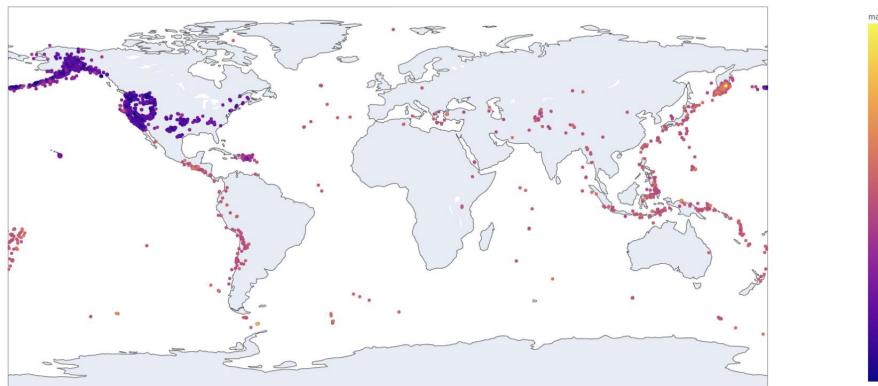
Visualization with Three or More Attributes – Heat Map

- Uses color to represent quantitative data
- Colors are used to quickly identify higher and lower values
- Often cool blue to red hot



Visualization with Three or More Attributes – Geographic Map

- Combines a quantitative variable associated with a geographic region and its geographic shapes



General Notes

- There are many different types of visualization techniques
- Which one you select is based on what data you are working with and what you want to show
- When possible, all visualizations should:
 - Be easily readable
 - * Proper Font
 - * Good distance between labels
 - * Not too cluttered
 - * ...
 - Quickly understood to some degree
 - Have self-explanatory labels

Why we Write Code instead of using Software

- Applications like Excel are great and you can do a lot with them
- Code can be very flexible and diverse
- Code can be easily automated
 - By changing one simple value, I can recreate multiple graphs extremely quickly
- Visualization code can be easily integrated with code for data extraction and pre-processing

Visual Misrepresentation

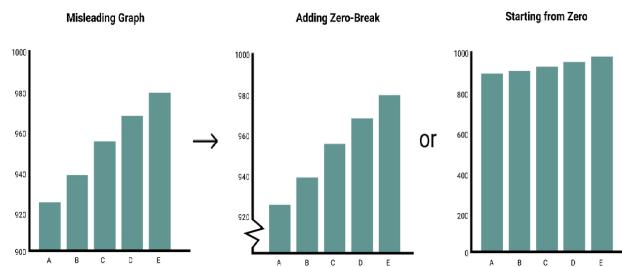
Dangers of Visual Misrepresentation

- Misrepresenting data happens when data and visuals are altered in some way to communicate a message that is misleading or inaccurate
- This can be done using true values

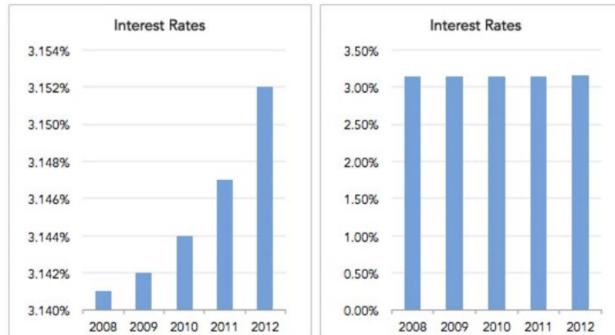
- Observers of presented data / visuals could be using that material to make important long-lasting decisions
- When presented with information, it is a good idea to be critical on who is presenting the data and try to be aware of any personal agendas or benefits

Truncated Axes

- Truncated Axes are the result of manipulating the scale of axes to exaggerate differences or make changes seem more significant than they are
- Similar to wearing very tall shoes to get on a ride at the amusement park
- An example with a graph is selected a range that only shows part of the data

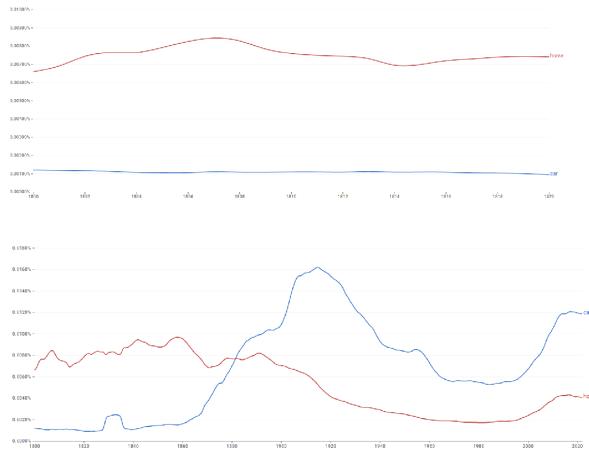


- In the plot below, we have the same data displayed on two different y -axes



Cherry Picking

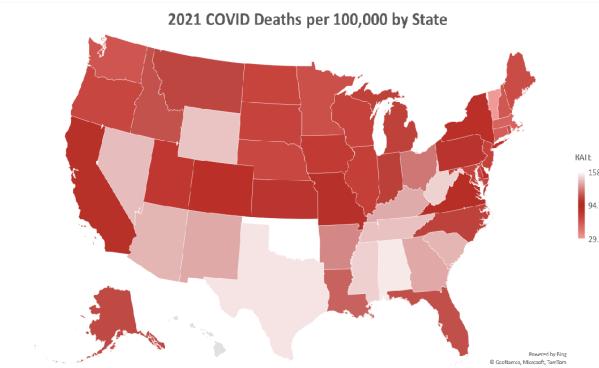
- Cherry Picking refers to the selection of specific data points or time periods to support a particular narrative while ignoring the broader context / story and relevant data
- Similar to saying you are good at basketball because you won against 5 people that have never played before, but you lost to 50 people that have at least one month of practice
- An example with a graph is highlighting and only selected a portion of it that tells a specific narrative



- The first graph above makes the word horse appear more often than the word car, whereas the opposite is told in the graph below it

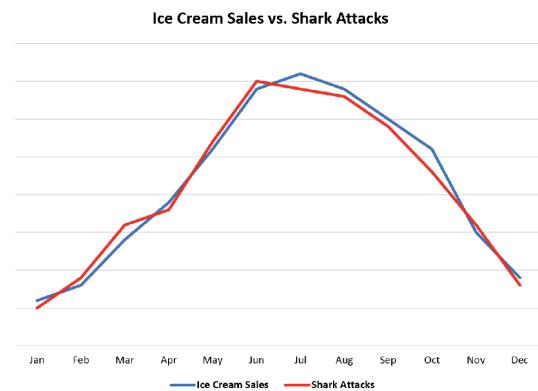
Misleading Use of Color

- Misleading Use of Color refers to using color in a manner that confuses or misleads the viewer
 - Inconsistent color scheme
 - Applying color to highlight certain data points without a clear rational
- Similar to selling a 25 pound bumper plate and making it blue
- An example with a graph is making bad financial situations black and good financial situations red



Misrepresentation of Association and Causation

- Misrepresentation of Association and Causation is incorrectly assuming that an association between two variables implies a causation
- Similar to telling people that you got 100% on tests in school because you ate pizza the week before those tests, and the few times you did not eat pizza, you did not get 100%
- An example with a graph is showing similar trends between two variables, but one variable is not causing the trend in the other



Mislabeling

- Mislabeling is when variables on the chart do not match their data
 - Mixing up x -axes and y -axes
 - Actually incorrect information
 - Can be done by accident
- Similar to selling someone a cat and calling it a hamster
- An example with a graph is flipping the axes so that it decreases instead of increases, but keeping the data points where they were

