Data Visualization Continued

- Interactive

  - Interactive visualizations are engaging
  - Plotly graphs allow you to hover over, select, and zoom
  - Gives more functionality to the observer

- Shared Socially

  - The ability to share data visualizations on social media or other publicly facing platforms can increase the impact of the data

- Data Democratization

  - Making data visualization accessible to the average person means more people can understand and benefit from the information
  - Know your audience

- Storytelling

  - Data visualization uses data and evidence to assist a story that the data is telling
  - Storytelling is an effective way of delivering information
  - What are you trying to say with your data?
  - Infographic is another visualization that can be used which combines text, art, and charts to tell a story

- Data Journalism

  - Reporting of information to society using large datasets
  - Visualization is an effective tool to help deliver information to readers of selected media

- Use of Artificial Intelligence (AI)

  - AI applications / models can help us improve our visualizations
  - Example: Caption generation (image to text generation) to give more context to the image
  - Natural language processing (NLP) can assist with analyzing textual data for better visualizations

- Mobile Friendly

  - Many people consume their information through smartphones
  - Making visualizations mobile friendly allows the visualizations to be more impactful and understood
  - What does it mean to be mobile friendly?

* Try to find a website that looks good on a laptop but not great on a smartphone and you can see the differences

- Staying Current

  - New data is always being produced
  - To keep our visualization current, we can connect them to a dataset that is being updated

Data Visualization for Impaired Individuals

- People can have impairments that makes reading and understanding data visualizations more difficult

- There are many things to consider to assist those with impairments

  - Don't rely solely on color to express information
    * Add shapes or labels to express similar information
    * Color vision deficiency safe color schemes
  - Use large and friendly fonts
  - Sonification – Usage of non language sounds to deliver information in a graph / chart
  - Use alt text (alternative text) so that people can use screen readers

Data Exploration

Data Exploration

- Exploratory Data Analysis (Descriptive Analytics) – Process of using statistical techniques to describe and summarize data features

- Descriptive Analysis – Systematically examines current and historical data, utilizing various benchmarks, to provide a detailed view of past and present outcomes

- Descriptive Statistics – Quantities like the mean, median, mode and standard deviation used to describe and quantify data features

  - There are 4 categories
    * Central Tendency
    * Variability
    * Shape (Already talked about previously when discussing skewness and symmetry)
    * Association

Case Study – Shopify and Descriptive Analysis

- Shopify is an e-commerce platform

- Shopify Analytics helps retailers

    – Know when to offer the right products

    – Understand trends

    – Target new and existing customers with deals

    – View information about their store's performance

- Metrics

    – Current number of visitors

    – Top products

    – First-time versus returning customers

    – Growth of company

Central Tendency

- Central tendency (typicality) aims to describe a quantitative attribute in a dataset via a single number representing the middle, center, or "typical" value

- Measurements of central tendency

    – Mean (Average)

    – Median

    – Mode

Averages

Trimmed Average

- An $X\%$ trimmed average is the average of the middle $X\%$ of the values

- If we want a 90% trimmed average of a variable:

    1. Sort values

    2. Remove the lowest 5% and upper 5% (round to the nearest integer if it does not divide evenly

    3. Calculate the mean of the remaining 90%

- Consider the table below:

Cat ID	Weight (lbs)
0	6
1	6
2	7
3	7
4	8
5	8
6	8
7	10
8	12
9	39

- The mean of cat weights is 11.1 pounds, the 80% trimmed average is 8.25 pounds (this trimmed average excludes cat 0 and cat 9)

Simple Moving / Rolling Average (SMA)

- The X-day simple moving average only looks at the most recent X days when calculating the mean

Day	Cost	3-day simple moving average
1	10	
2	9	
3	11	10
4	12	10.7
5	6	9.7
6	13	10.3