Crash Course:
Data Analysis & Presentation

Week 1

Jordan Krell
jkrell@flintsciencefair.org

Safety

- Your safety is paramount
  - Please read our virtual handbook
  - If at any point you feel unsafe, logout immediately and contact the FRSEF.

- Read the FRSEF Virtual Programming Handbook
  - In your email

- Contact us in an Emergency or for Routine Assistance
  - Jordan: 248-330-4269 jkrell@flintsciencefair.org
  - FRSEF: 810-797-5290 kdutton@flintsciencefair.org

Data Acquisition & Measurement

Poor measurement and acquisition cannot be fixed by excellent analysis

Measurement Setup *If logging data yourself*

- What is your logging setup?
  - Logger
    - If analog measurement, how many bit ADC?
      - Why is this important??
  - Sensor / s
    - What can be measured?
    - Are you using the proper sensor type for the measurement?
      - Range
      - % error
  - Calibration

Sensor Selection - Accelerometer

- Exercise to select an accelerometer for helmet impact testing
  - What's important for our measurement?

A) ST LIS3DH

- 3-axis
- ±2, 4, 8, 16g
  - 10 bit (2^10 = 1024)
  - 16/1023 = 0.01564 g / increment
- 1-5 kHz output
- I2C / SPI output
- Temp Sensitivity: ± 0.01%/C

B) ST H3LIS331

- 3-axis
- ±100g, ± 200g, ± 400g
  - 16 bit (2^16 = 65536)
  - 800/65536 = 0.0122 g / increment
- 0.5 – 1 kHz output
- I2C / SPI output
- Nonlinearity: 2%
- Temp Sensitivity: ± 0.01%/C

3) AD ADXL377

- 3-axis
- ± 200g
- 0.5 – 1 kHz output
- Analog output
- Nonlinearity: ± 0.5%
- Temp Sensitivity: ± 0.02%/C

4) AD ADXL326

- 3-axis
- ± 16g
- 0.5 – 1 kHz output
- Analog output
- Nonlinearity: ± 0.3%
- Temp Sensitivity: ± 0.01%/C

Sensor Selection - Temperature

- Select a temperature sensor
  - Best Insulation to Use
  - Lowering the Freezing Point of Water
  - Thermal Efficiency of Different Bio-Fuels

A) DS18B20 Probe

- Range: -55 to 125 °C Range
- Interface: 12bit ADC (4096)
- Resolution: 0.0625°C
- ±0.5°C accuracy from -10°C to +85°C
- Time Constant: 180 s

B) Type K Thermocouple

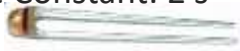
- Range: 0 to 400 °C Range
- Interface: Analog amplifier
- ±2°C accuracy
- Time Constant: 20 s

C) Temp Sticker

- Range: -30 to 70 °C Range
- Interface: NTC Thermistor, use voltage divider
- ±2°C accuracy
- Time Constant: 10 s

D) Thermistor

- Range: -40 to 125 °C Range
- Interface: NTC Thermistor, use voltage divider
- ±2% accuracy
- Time Constant: 2 s



Sensor – Error & Calibration

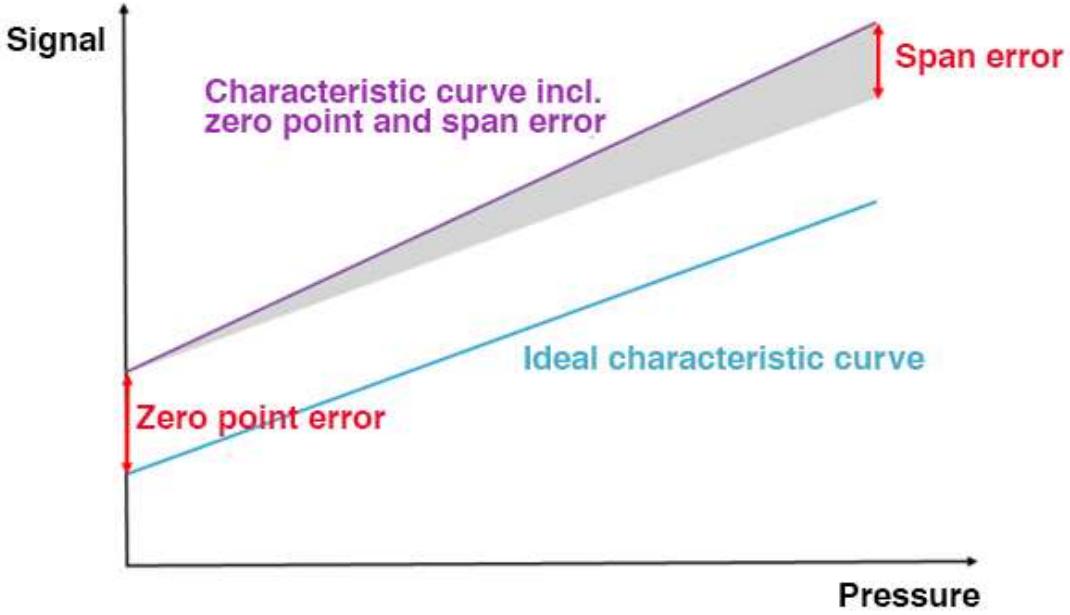
- **Zero-point error –** The start of the measuring range is too high or too low (this is the *zero offset*), thus shifting the entire scale up or down by that zero offset value.

- **Span error** – The distances between the individual divisions (the span) from the zero point to the full-scale value are even but wrong, which has the effect of magnifying errors at the upper end of the scale.

- **Non-linearity** – The distances between the individual divisions from the zero point to the full-scale value are *not* even (non-linear), thus making the ideal straight line into a curve.

https://blog.wika.us/knowhow/pressure-sensor-accuracy-3-errors/

Sensor – Error & Calibration



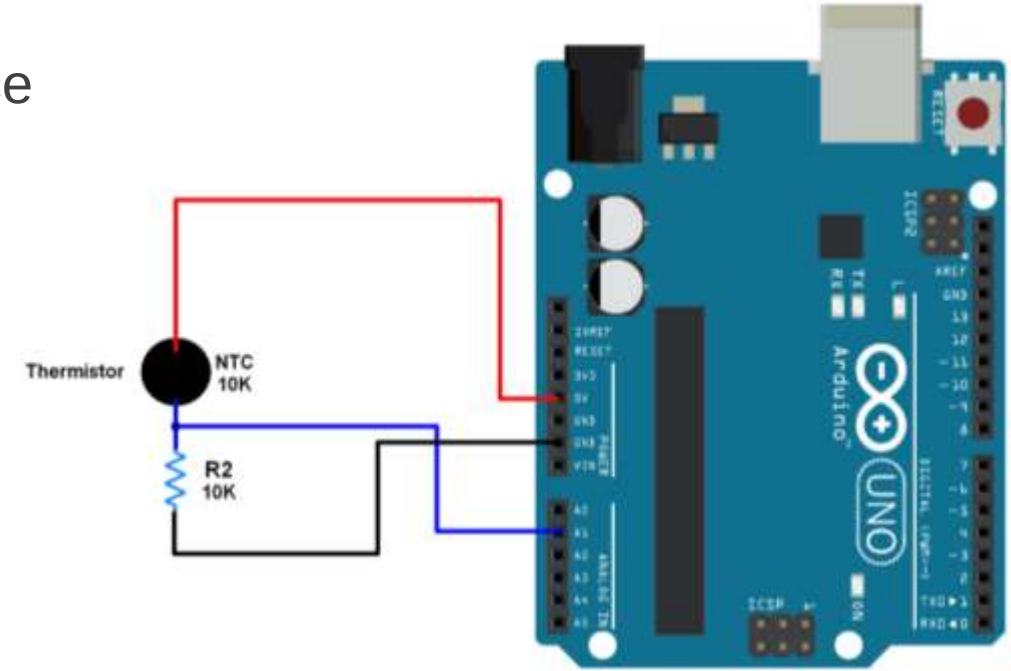
To note: other errors, such as non-linearity, are not considered in the chart

Sensor – Thermistor Measurement

- Thermistor: changes resistance with temperature
- Use a voltage divider to measure the resistance of the thermistor

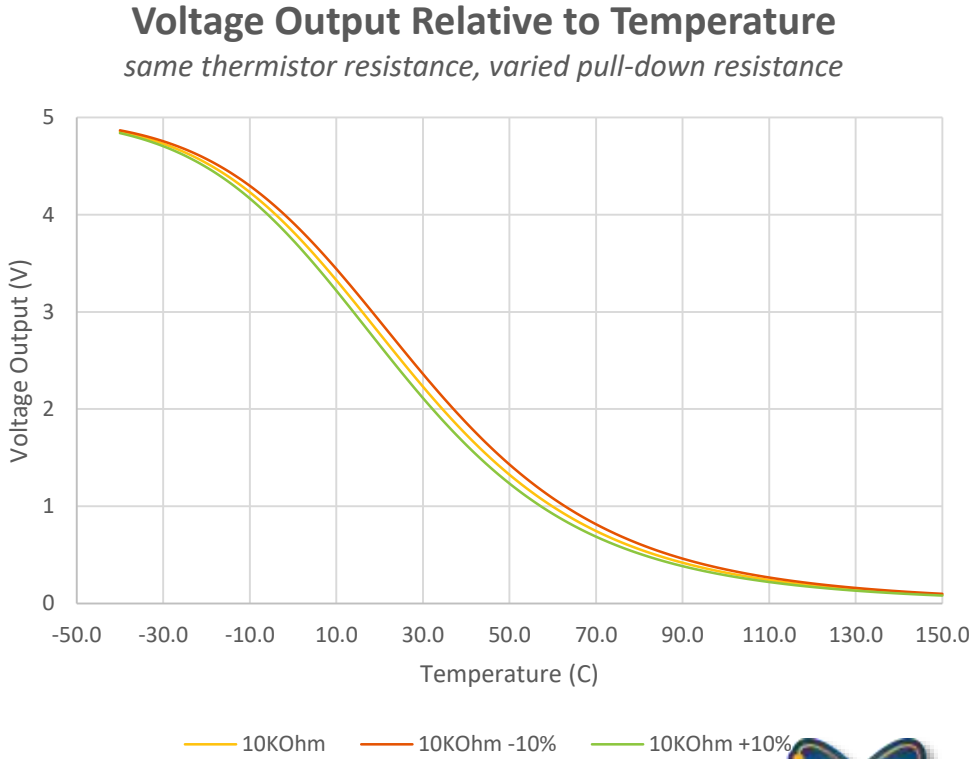
$$V_{out} = V_{in} \times \left(\frac{R2}{R1 + R2} \right)$$

  - R1: Known resistor
  - R2: Thermistor



Sensor – Error & Calibration

- Thermistor with a 10kΩ pull-up resistor (20% error)
  - -30°C
    - +10%: -28°C
    - -10%: -32°C
  - 25°C
    - +10%: 23°C
    - -10%: 27°C
  - 60°C
    - +10%: 57°C
    - -10%: 63°C
  - 100°C
    - +10%: 97°C
    - -10%: 103°C



Know Your Sensor & How It Works

- What does our measurement mean?
  - pH
  - Turbidity

- Critical to know your sensor
  - What is changing on the sensor (resistance, capacitance, etc)
  - Limitations
  - Sources of error
  - Find the data sheet!
- Judges will often ask about sources of error and how this can be reduced.

What can all be measured?

- Petri dish samples

Data Presentation

Share your work so it can be understood.

Communications: What is our goal + how will it be seen / viewed?
Science and Engineering Fair

- What is our goal?
  - Share the results of our research.
    - Hypothesis is supported or not supported
    - Design solves a problem and fits our design criteria
    - How external factors have / do not have an effect

- How will it be viewed?
  - Preliminary Judging (5-6 judges)
    - 10-15 minutes for judges to learn about your project and score.
  - Finalist Judging (5-6 judges)
    - 12 minute Zoom interviews

Communications: What is our goal + how will it be seen / viewed?
Internal Design Reviews

- What is our goal?
  - Approve a product for sale to customers
    - Product is safe for all customer scenarios
    - Product meets all design requirements

- How will it be viewed?
  - In-depth design review meeting
    - 2-3 hour meeting to review with 5-10 colleagues
  - Report review
    - Report emailed and on drive for many to view on their own time

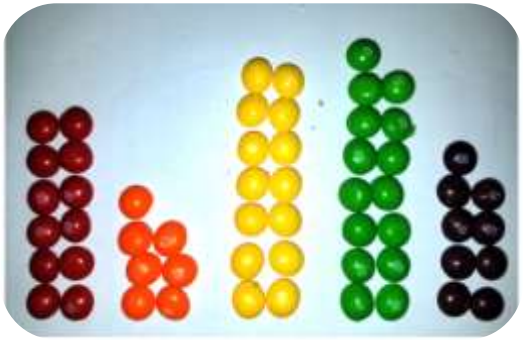
Communicating - Visuals

Process: Sorting Skittles by color

1) Open bag of Skittles and pour onto a table
2) Group Skittles by color
3) Count the number of each color

Step 1: Open bag of Skittles

Step 2: Group Skittles by color.

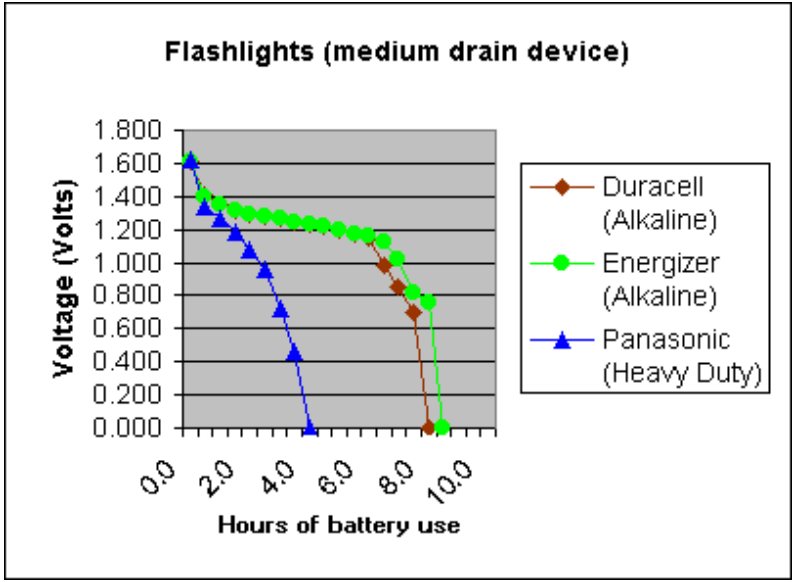


Step 3: Count the number of each color.

What is our goal?

- Communicate our results: what tells us more?

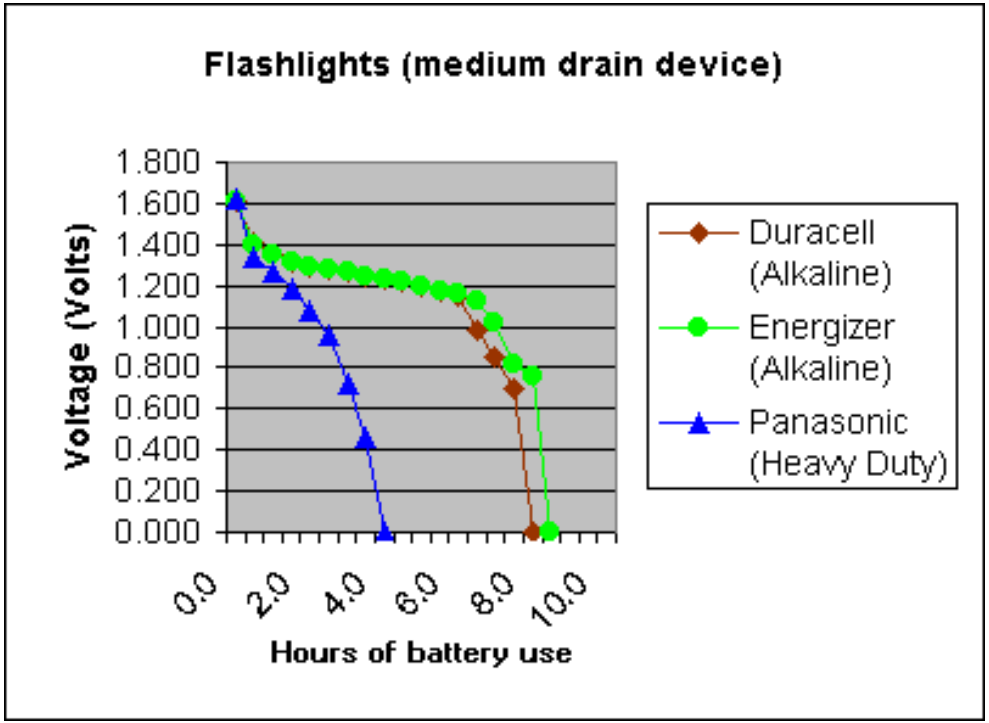
After conducting an analysis of the outcome with 3 trials, the Energizer Alkaline lasted an average of 9.5 hours, the Duracell Alkaline an average of 9.2 hours and the Panasonic Heavy Duty an average of 5.1 hours.



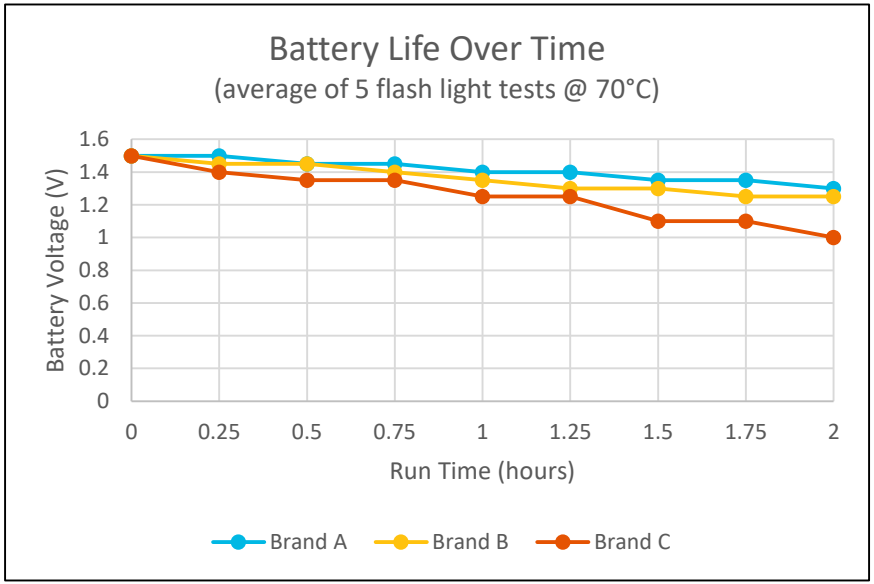
https://www.sciencebuddies.org/science-fair-projects/science-fair/data-analysis-graphs

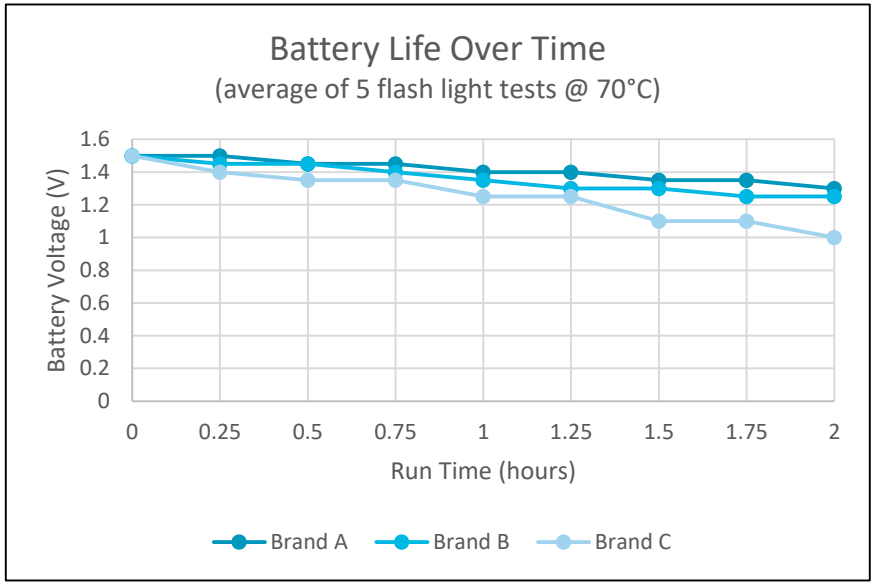
Graph Etiquette

- Title
- Axis Labels
  - Include units
- Legend / Key
- Readability
  - Background
  - Colors
  - Excess Precision

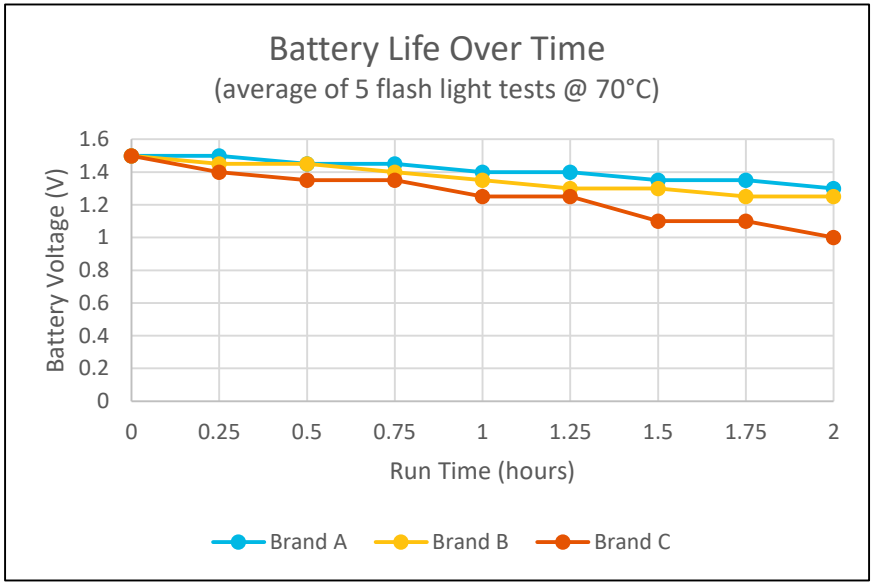


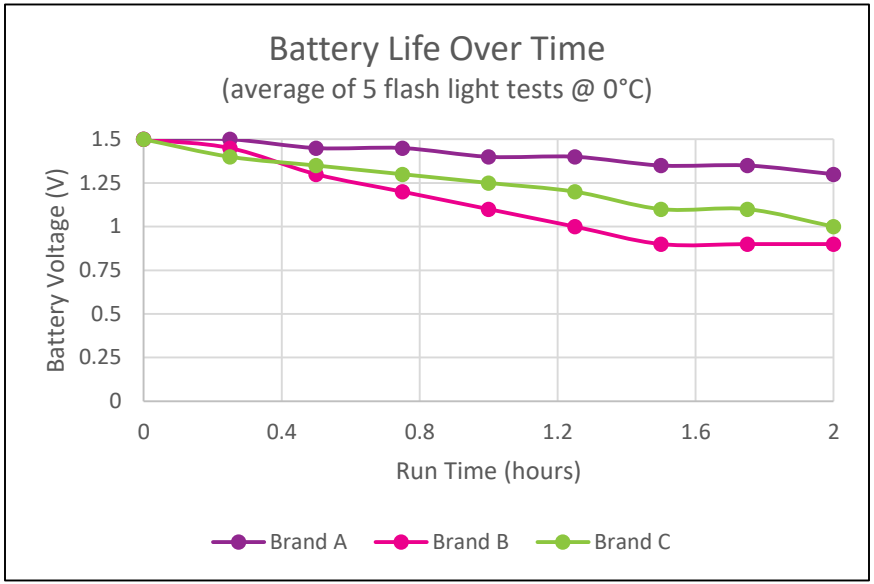
Graph Etiquette - Colors



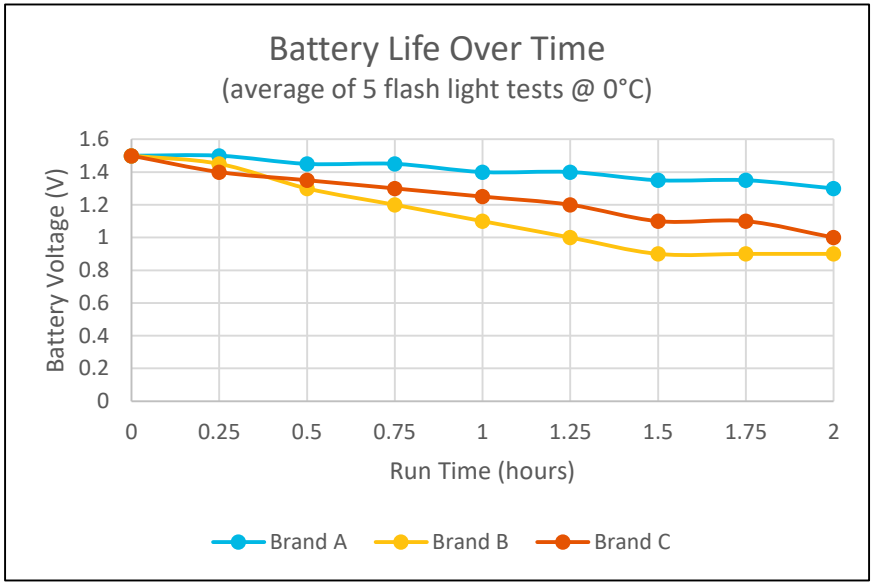


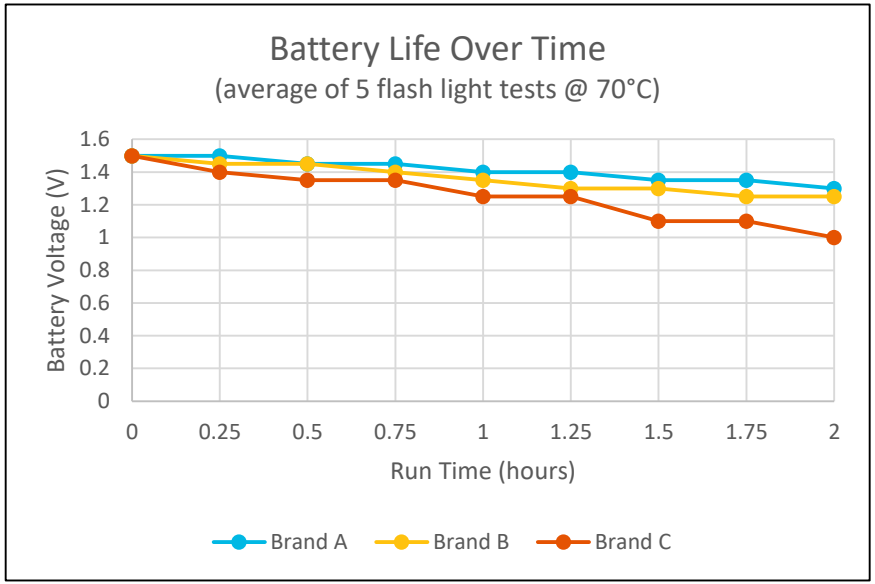
Graph Etiquette - Consistency





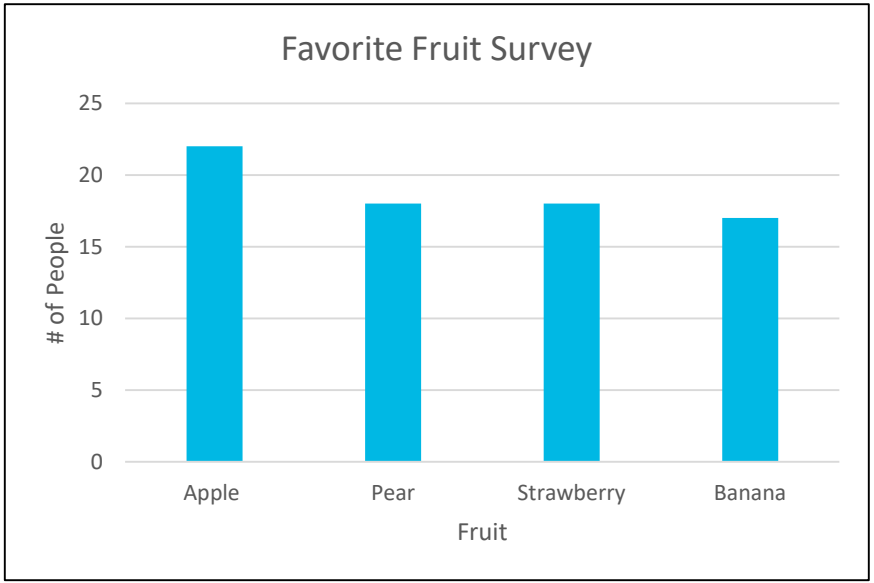
Graph Etiquette - Consistency

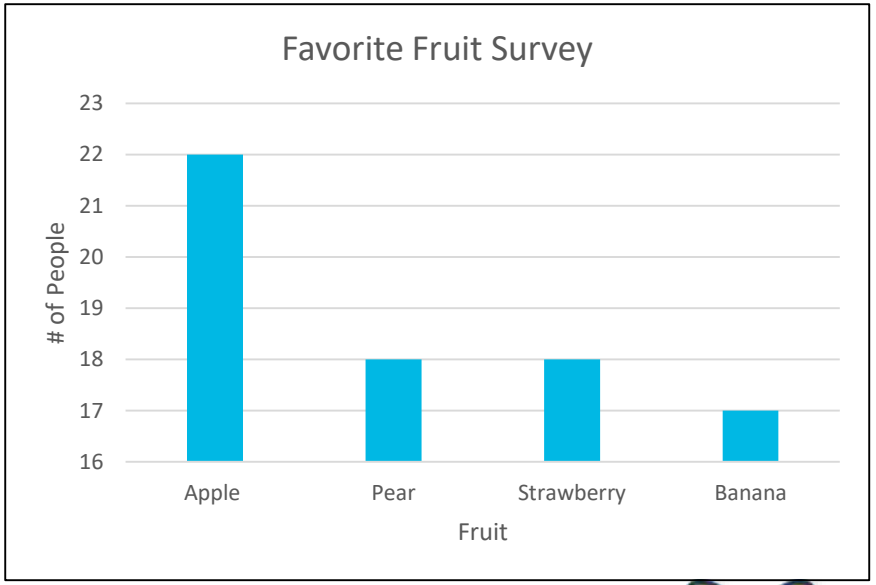




Graph Etiquette – “Cable News” Graphs

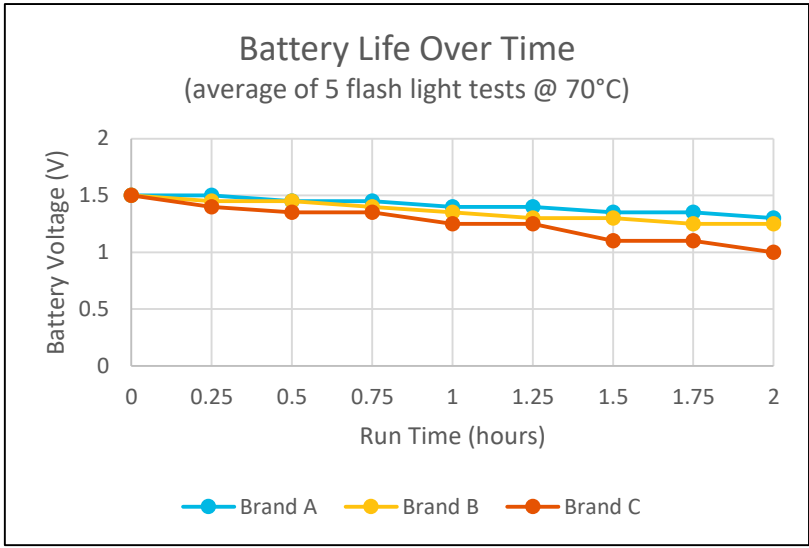
- For a bar chart our baseline always needs to be 0





Graph – Layout Rule of Thumb

- X-Axis: Independent Variable
  - What we are changing
- Y-Axis: Dependent Variable
  - The outcome
- Battery Experiment
  - Independent Variable: Run time of flashlight
  - Dependent Variable: Battery voltage



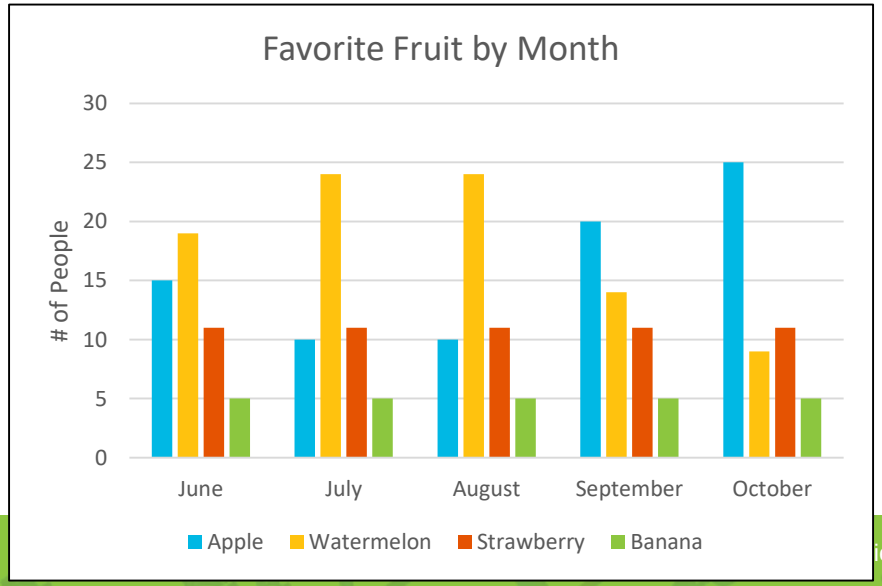
What is our goal?

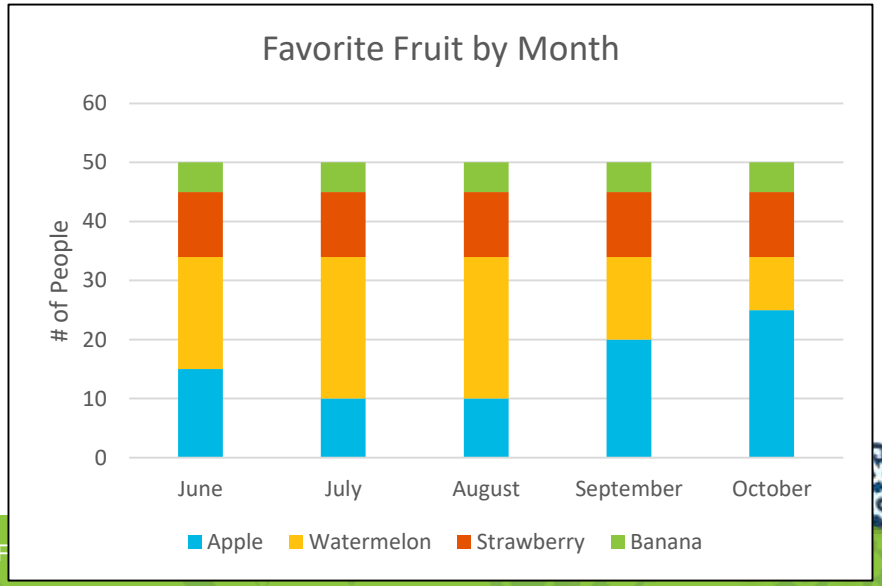
- Communicating: How is our data changing over time
  - Line Chart
  - Scatter Plot (with connecting lines, aka line chart w/ dots)
  - Bar Chart (each time period is a bar)
  - Box Plot (advanced, will show at end)

- Examples are data from a survey of 50 people
  - Survey sent each month
  - Ask participants to choose their favorite fruit
  - Goal: Does someone's favorite fruit change throughout the year

Change Over Time – Bar Chart

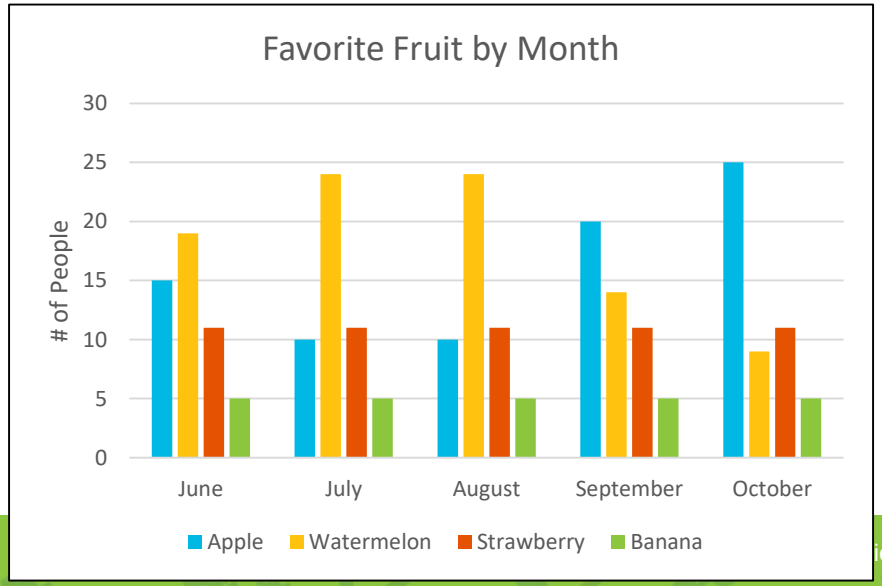
- Bar Chart (each time period is a bar) https://chartio.com/learn/charts/bar-chart-complete-guide/
  - Good choice if your independent variable is not numerical
    - Standard bar chart (left)
    - Stacked bar chart (right)

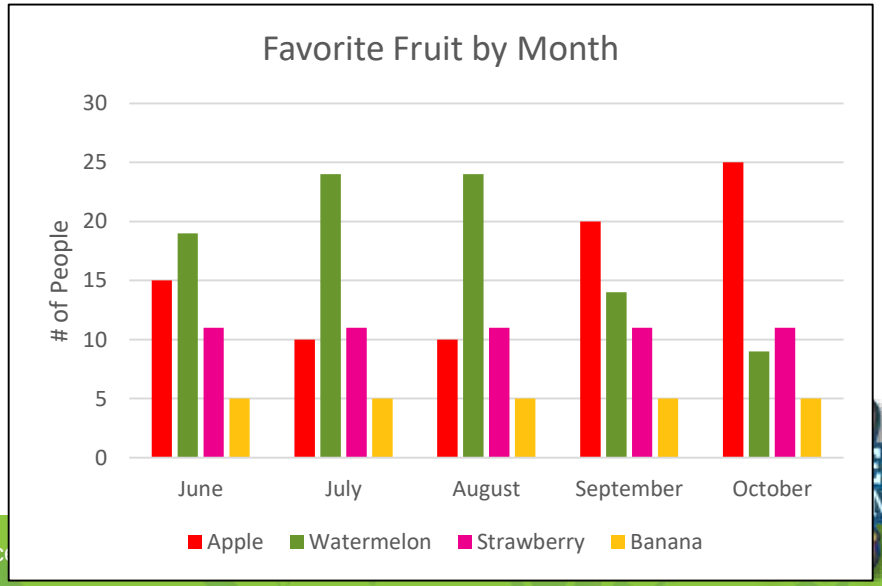




Hints – Color is Your Friend

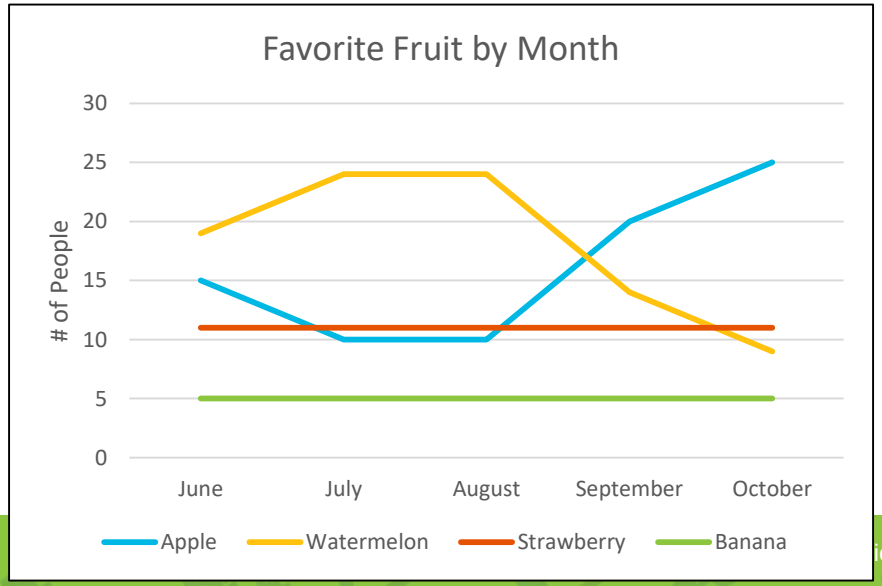
- If applicable: have your colors match the impression of the item
  - Apple: red
  - Banana: yellow

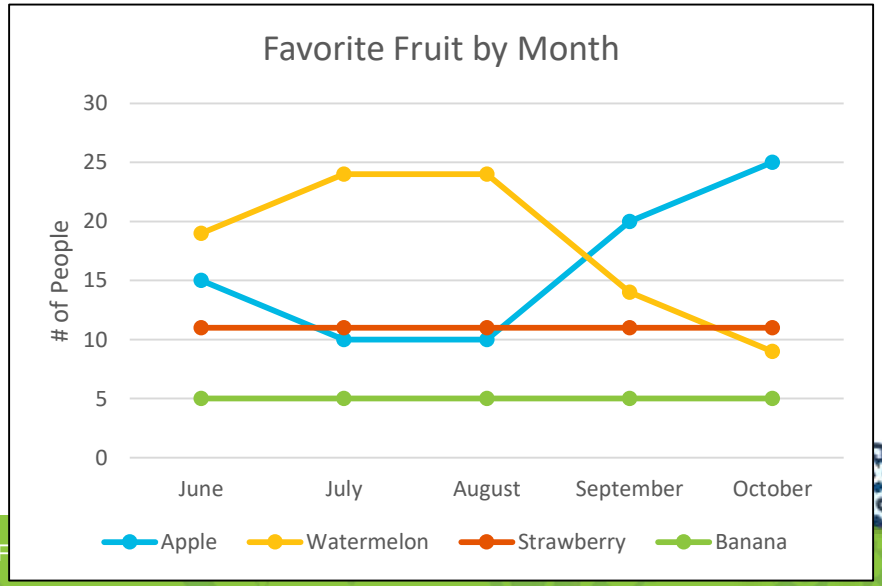




Change Over Time

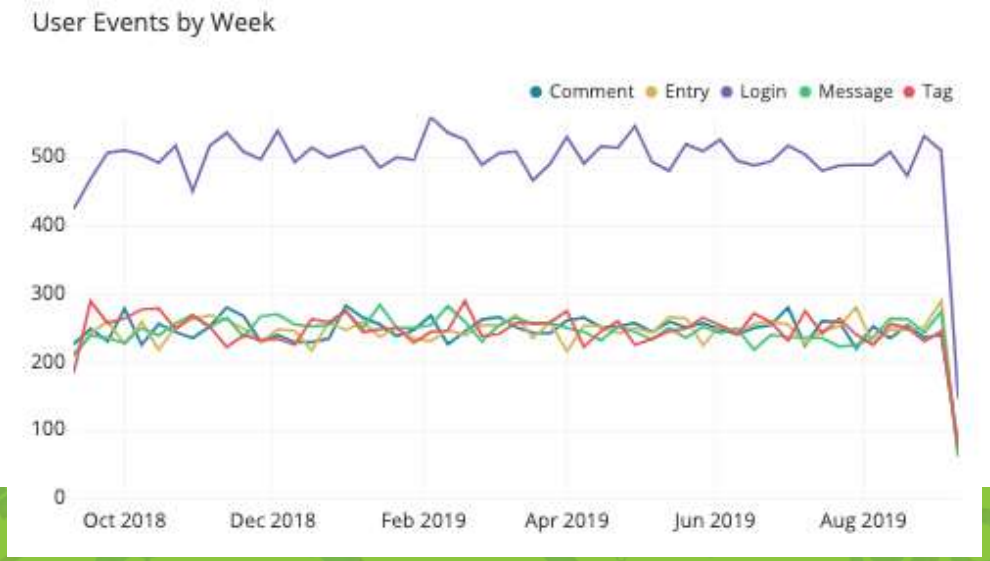
- Line Chart and Scatter Plot w/ lines
  - Good choice if your independent variable is numerical
    - Line chart (left)
    - Scatter plot w/ lines (right)





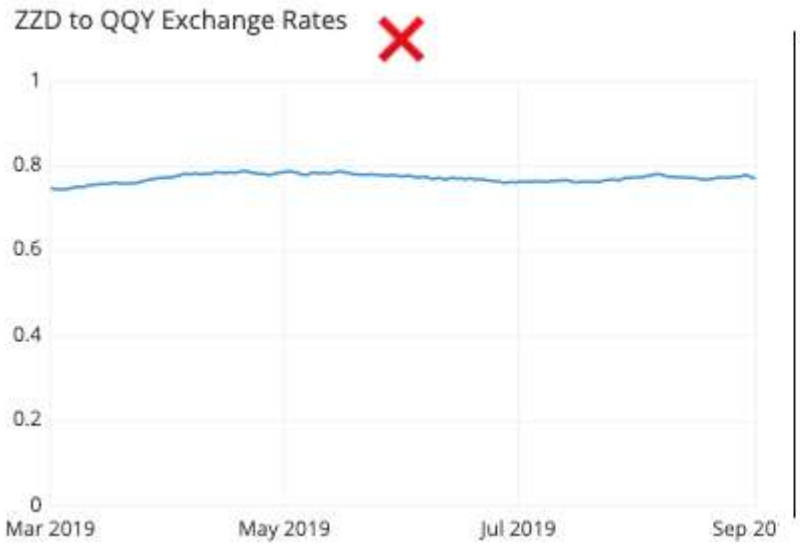
Hints – More is Not Always Better

- Think about what you want to communicate
  - Good: clearly able to see logins are greater than other activities
  - Bad: Difficult to determine differences between Entry, Message, xxxx



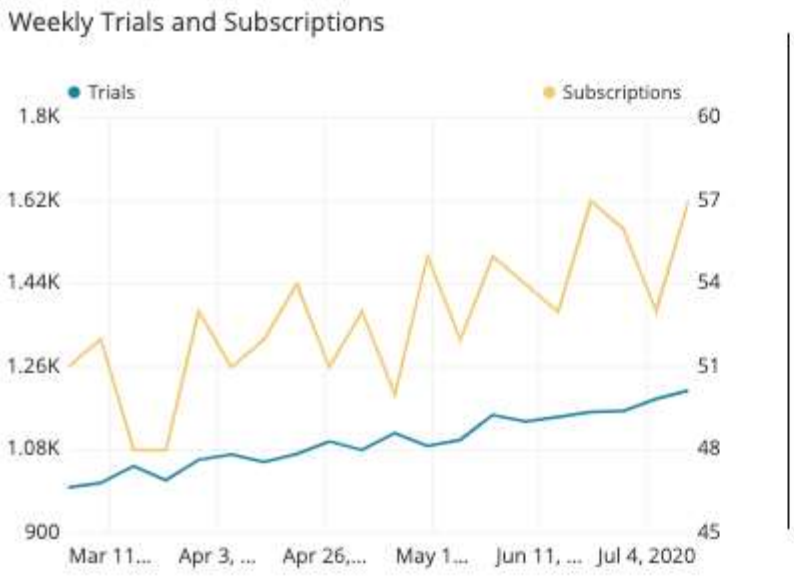
Hints – Baseline Value

- Think about what you want to communicate
  - For Line Chart: Emphasize changes in value



Hints – Dual Axis

- We can compare different dependent variable variables on a 1 graph
  - Be consistent with your axis scaling (don't place a large offset on the scaling)



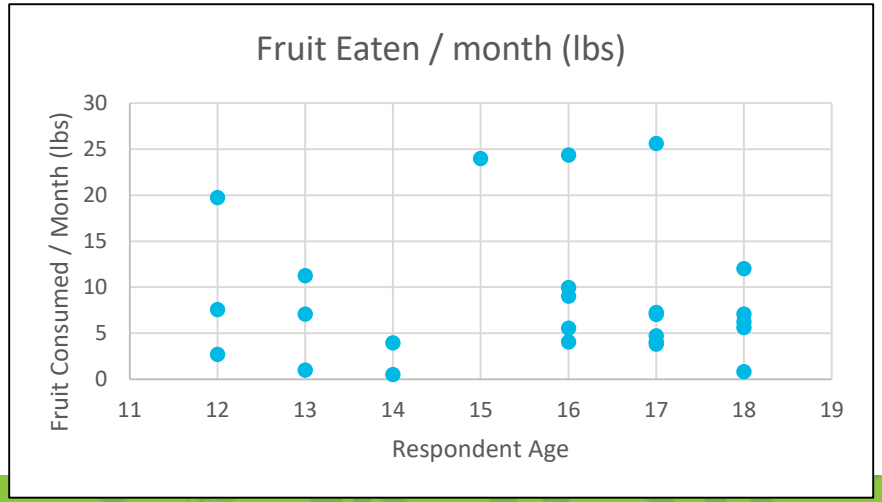
What is our goal?

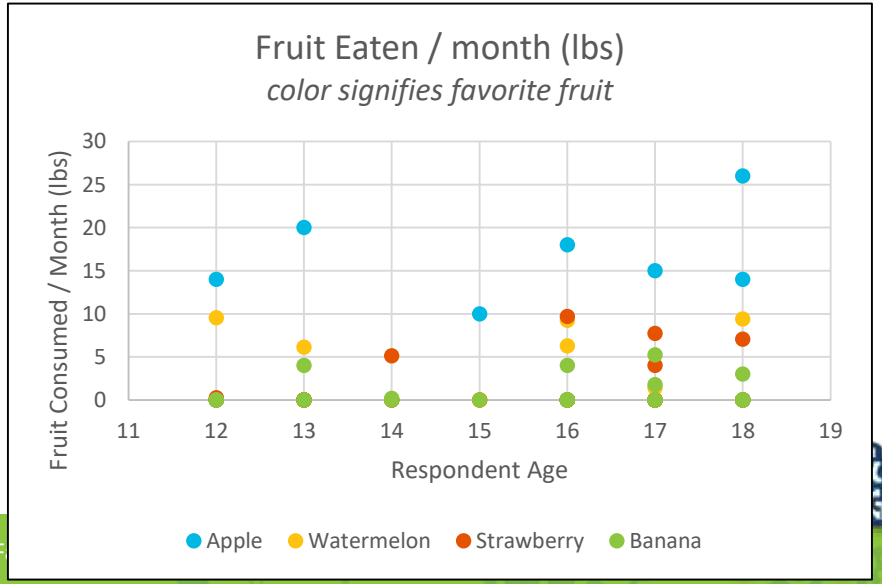
- Communicating: Observe relationships between groups
  - Scatter Plot
  - Bubble Chart
  - Grouped Bar Chart

- Examples are data from a survey of 50 people
  - Survey sent each month
  - Ask participants to choose their favorite fruit
  - Goal: Does someone's favorite fruit change throughout the year

Relationship Between Groups

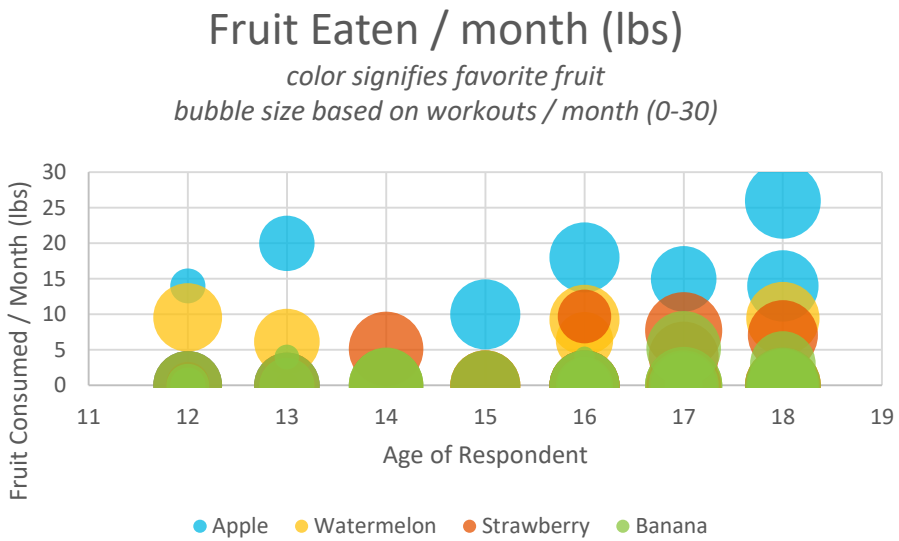
- Scatter Plot https://chartio.com/learn/charts/what-is-a-scatter-plot/
  - Good choice for determining correlation between groups, finding outliers, versatile
    - Scatter Plot (left)
    - Scatter plot w/ color for favorite fruit





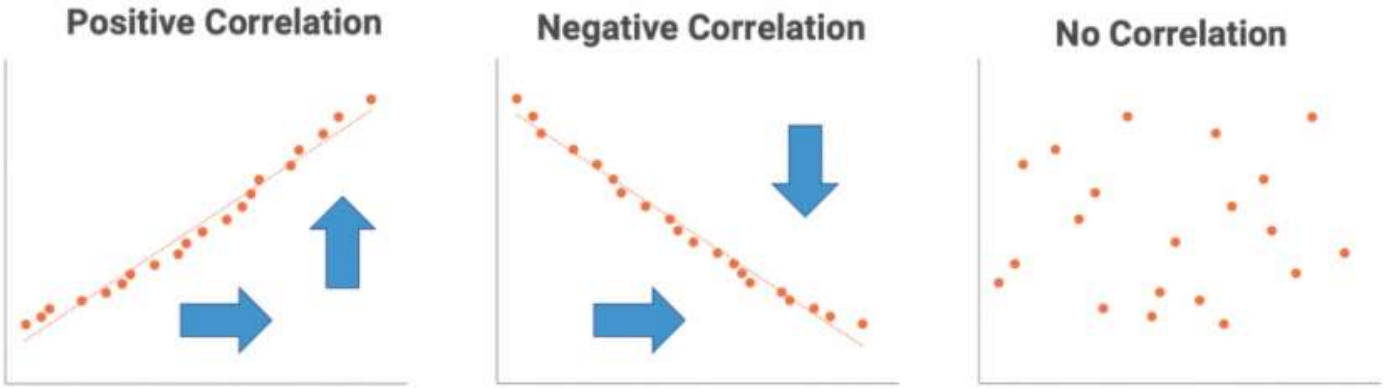
Relationship Between Groups

- Bubble Chart
  - Good choice for determining correlation between groups, finding outliers, versatile



Correlation

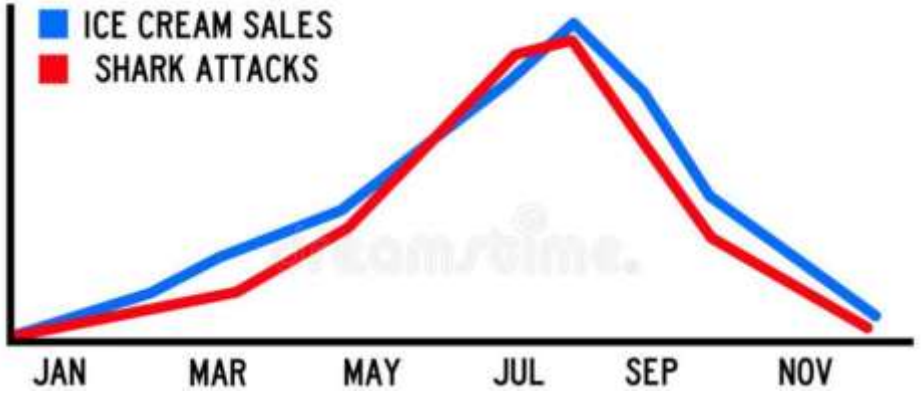
- Correlation: the relationship between the two variables
  - How much does one variable affect the other?
  - Positive correlation: Both variables move in the same direction
  - Negative correlation: Variables move in opposite directions
  - No correlation: No link between the two variables



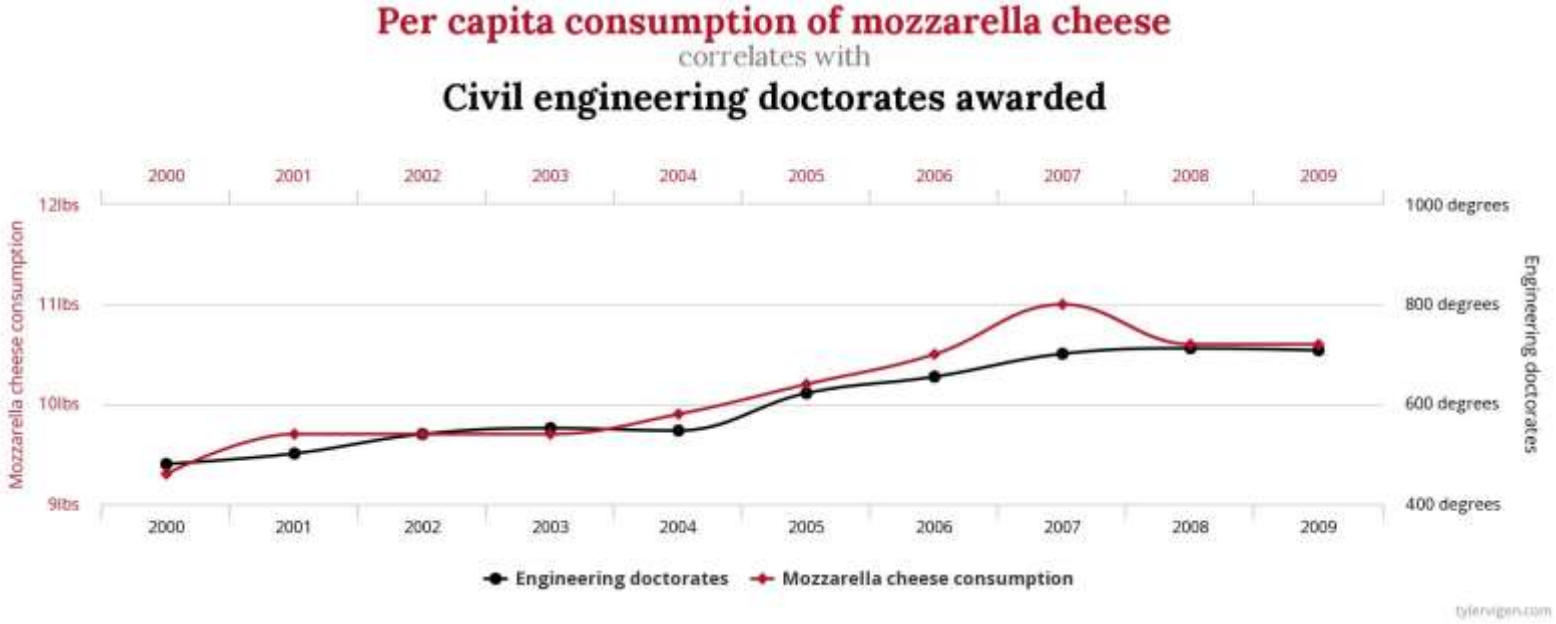
https://mylearningsinaiml.wordpress.com/2018/11/21/scatter-plots/
https://astutesolutions.com/blog/articles/causation-vs-correlation

Correlation =/ Causation

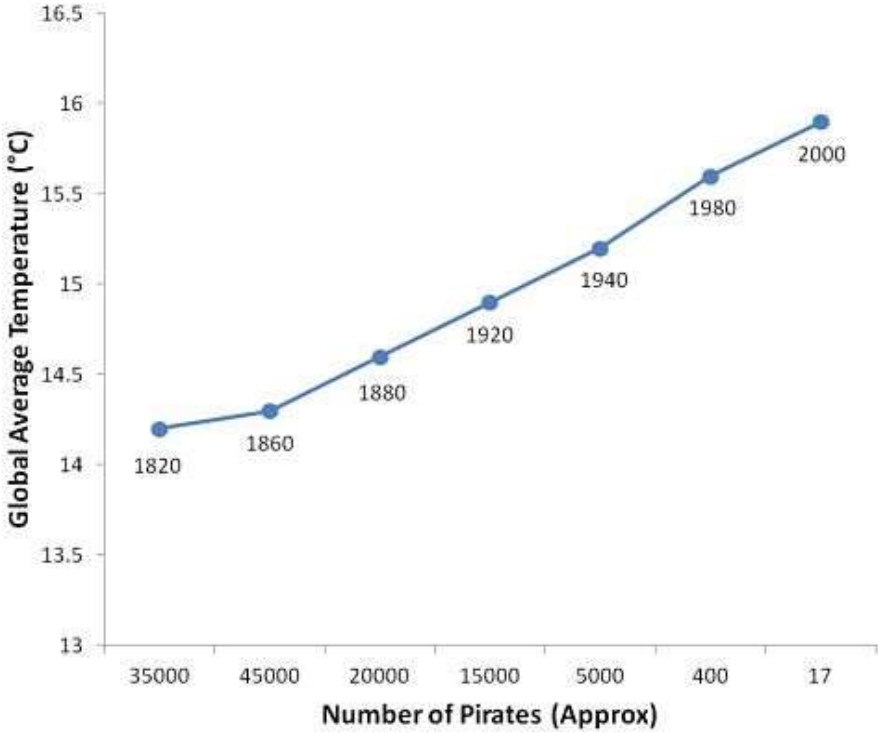
- Correlation: A change in one variable mirrored by a positive or negative change in the other.
  - Spurious Correlation: strong relationships between variables that are not caused by one another.
- Causation: One variable is changing as a result of the other variable.



Correlation =/ Causation



Correlation =/ Causation



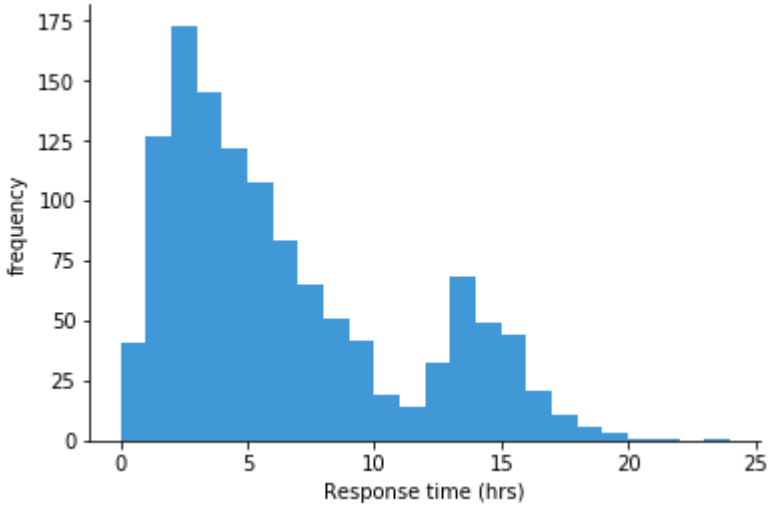
What is our goal?

- Communicating: How our data is distributed
  - Bar Chart
  - Histogram
  - Density Curve
  - Box Plot (advanced, will show at end)

- Examples are data from a survey of 50 people
  - Survey sent each month
  - Ask participants to choose their favorite fruit
  - Goal: Does someone's favorite fruit change throughout the year

Data Distribution

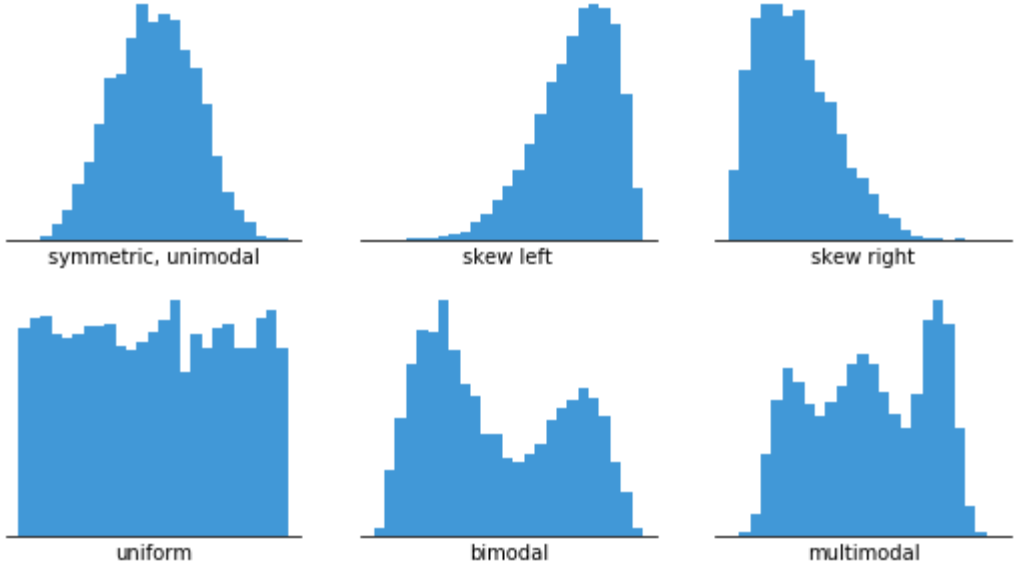
- Histogram
  - Plots the distribution of a numeric variable's values as a series of bars
  - The x-axis values are "binned" together (ex. each hour is binned together)



https://chartio.com/learn/charts/histogram-comp

Data Distribution

- Histogram
  - Very good at showing the distribution of our data



What is our goal?

- Communicating: Part to Whole Comparison (understanding the components that make up the total)
  - Pie Chart
  - Doughnut Chart (pie chart w/ the center missing)
  - Stacked Bar Chart
  - Stacked Area Chart

- Examples are data from a survey of 50 people
  - Survey sent each month
  - Ask participants to choose their favorite fruit
  - Goal: Does someone's favorite fruit change throughout the year

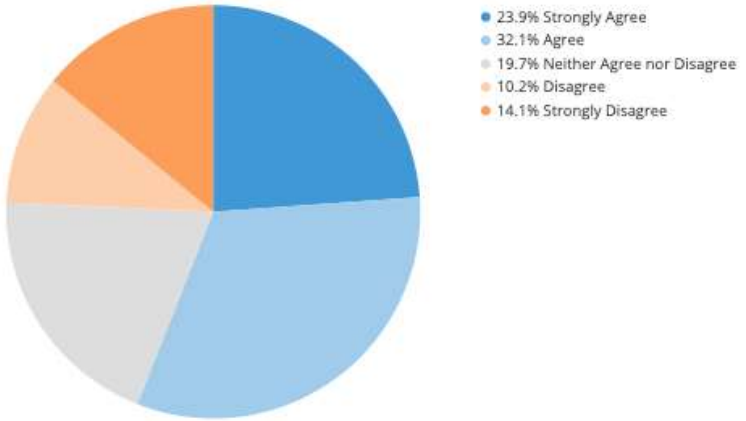
Part to Whole Comparison

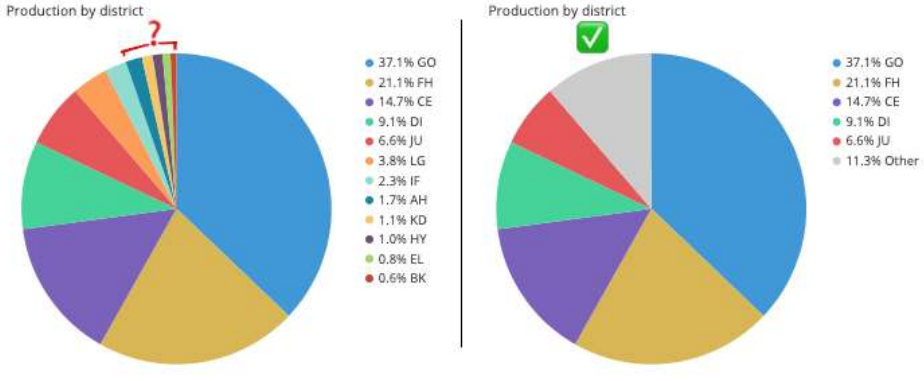
- Pie Chart
  - Comparing each variable relative to the whole data set
  - Only use for the above use case

https://chartio.com/learn/charts/pie-chart-complete-guide/

Hints – Pie Chart

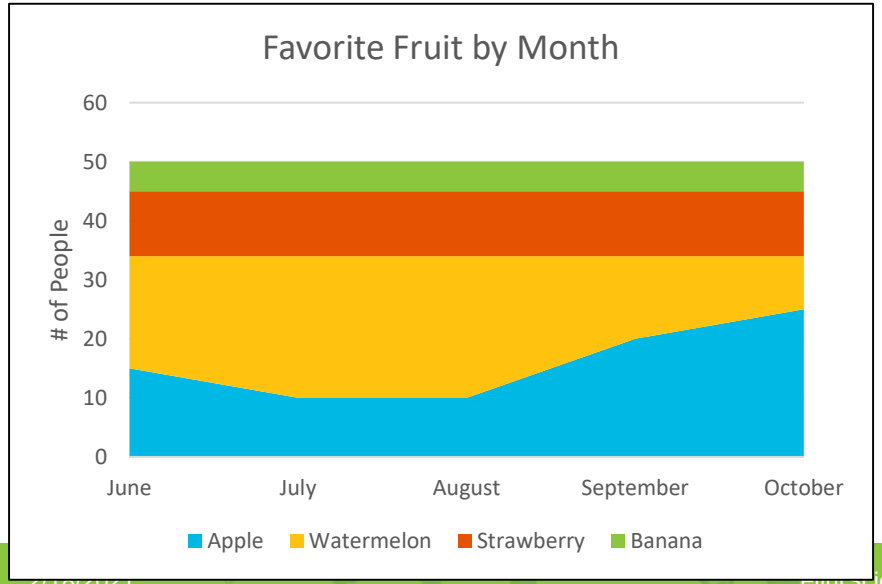
- Include annotations (%, value)
- Order slices by size
- Limit the number of slices
  - Group many “small” slices into “other”

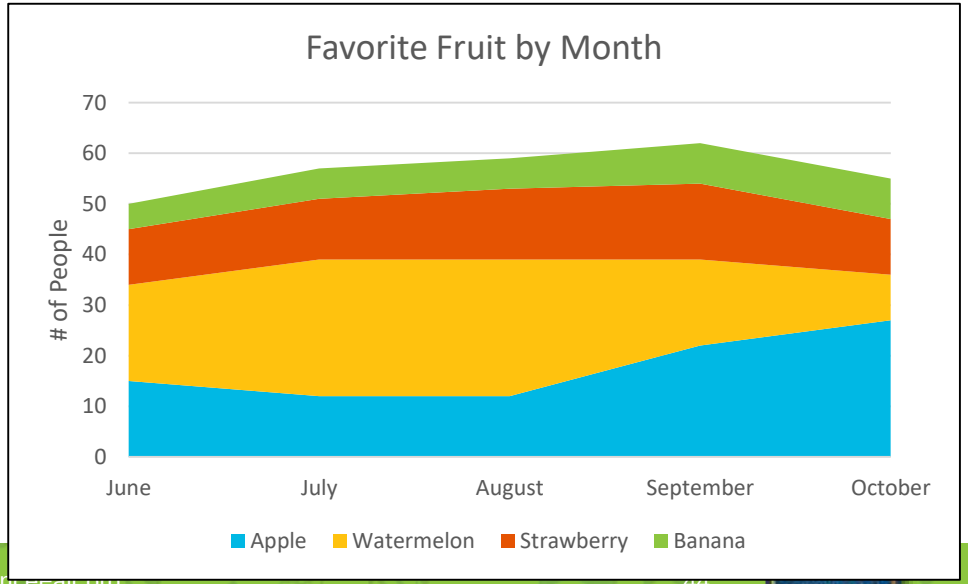




Part to Whole Comparison

- Area Chart https://chartio.com/learn/charts/area-chart-complete-guide/
  - Line Chart + Bar Chart
  - Very powerful if the whole is also changing (left: same # / month; right: different # / month)



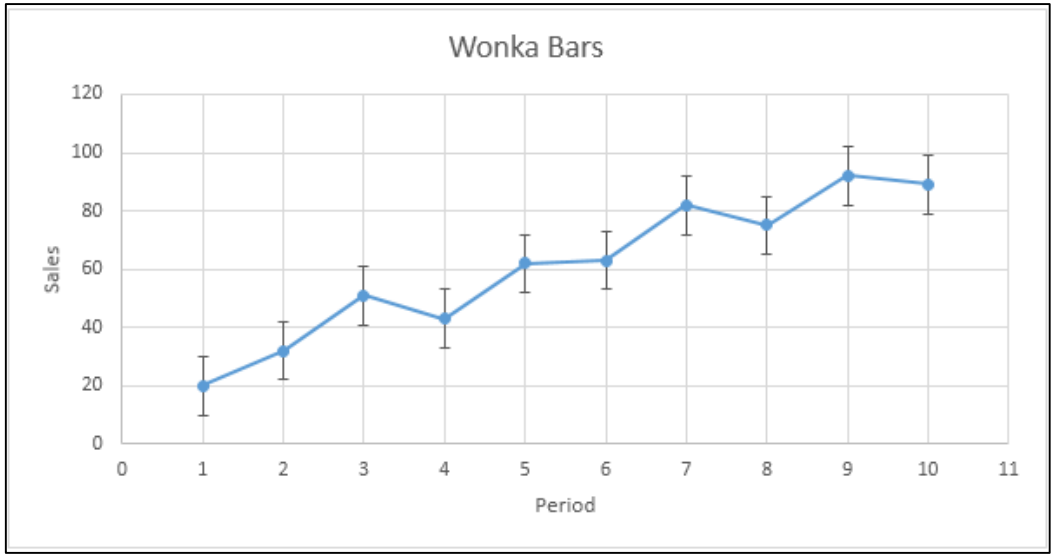


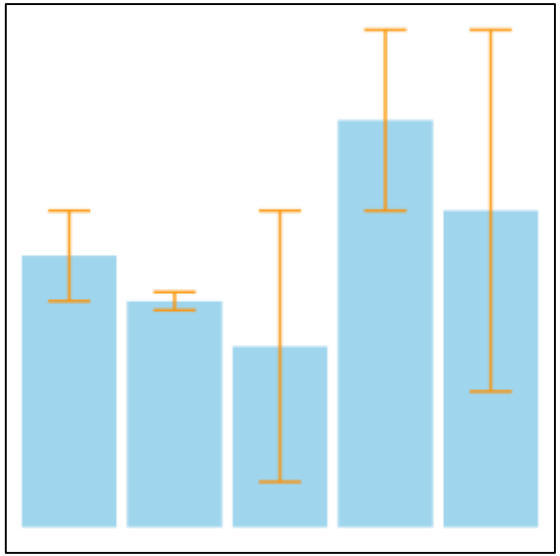
Uncertainty

- What is uncertainty?
  - The variability of our measurement.

- How do we communicate our uncertainty?
  - Uncertainty / Error bars
  - Uncertainty shading
  - Box plots

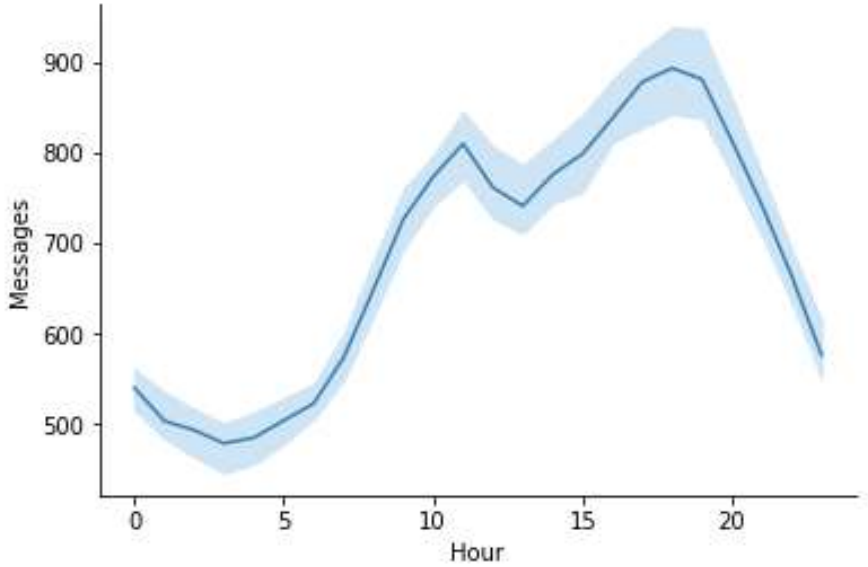
Error Bars





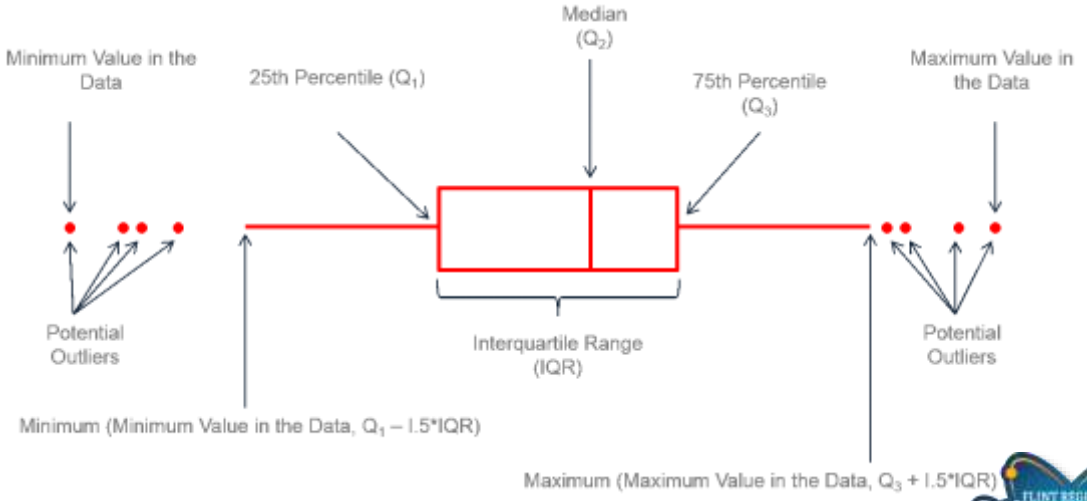
Shading for uncertainty

- An alternative to error bars is to add shading for uncertainty.

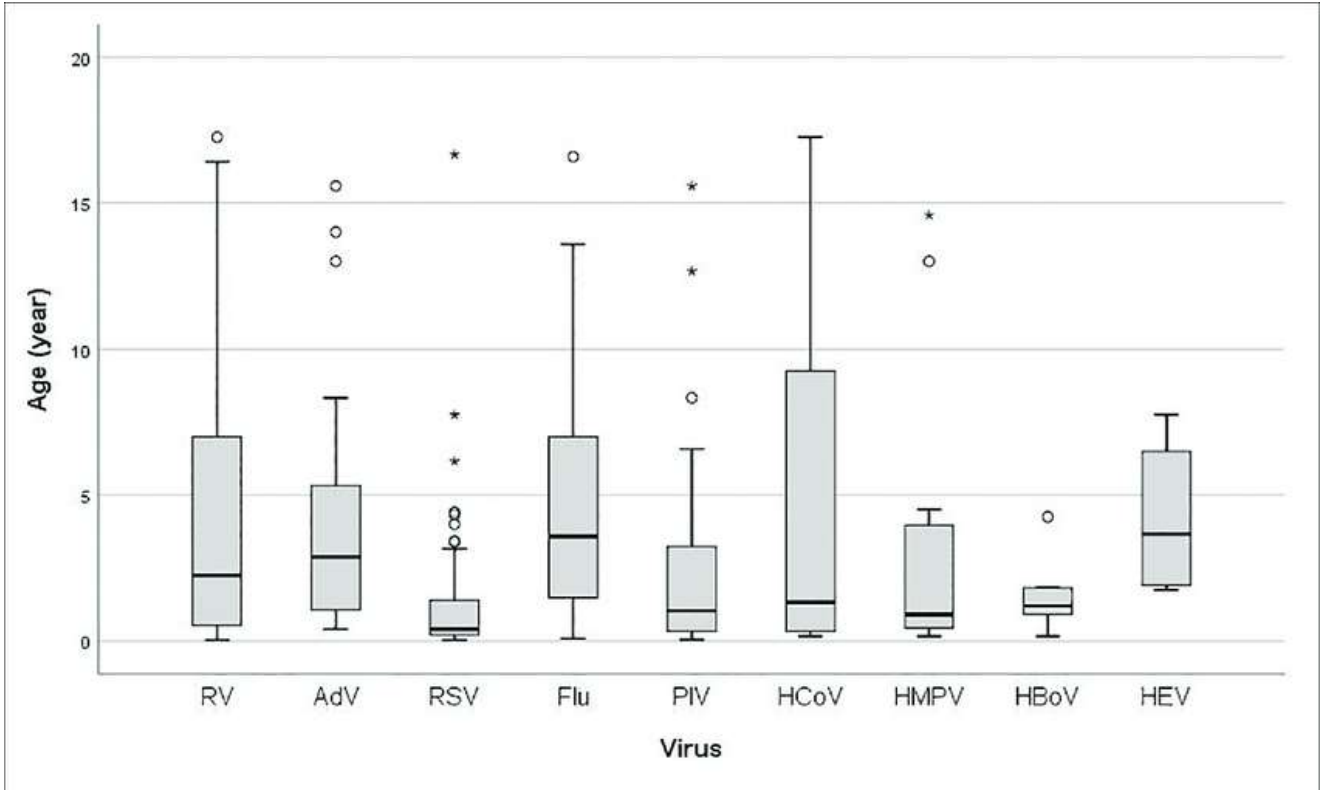


Box Plot

- Combines many concepts https://chartio.com/resources/tutorials/what-is-a-box-plot/
  - Provides a 5 number summary in 1 graph
    - Minimum
    - Maximum
    - Median (Average)
    - First Quartile (25%)
    - Third Quartile (75%)



Box Plot



Thank You!

Reach out anytime:

Jordan Krell
jkrell@flintsciencefair.org