Quiz 2 Notes

Videos:

- Quantitative vs. Qualitative
    - Quantitative: Numerical Analysis/Data Collection, what?
    - Qualitative: Open-ended/Why questions, why?
- Exploratory Data Analysis
    - Method to observe trends and patterns within data, not rigid or formal
    - Common Tools: Scatterplot, Boxplot, Histograms
- Tips for Finding Data
    - Availability
    - Data Cleanliness
    - Good Documentation
- Probability Distributions:
    - Bernoulii: P(success) + P(failure) = 1, single trial with only two outcomes
    - Binomial: essentially multiple Bernoulli samples strung together
    - Geometric: Number of failures of needed before success when stringing together Bernoulli Trials
    - Normal Distribution: most common (bell curve), 95% of sample with one standard deviation, 97.5% within two
    - Poisson: used with events over time, specifically measures number of events over a time period
    - Exponential: measures time between events(not needed for quiz)
- Hypothesis Test:
    - Come up with hypothesis and then come up with a statistical test to see if true
        - Create Research Question
        - Change Question into a Hypothesis (Null and Alternative)

        - Collect Data
        - EDA
        - Decide on Statistical Level
        - Run Stat Test
        - Examine Output/Check Reasonability
        - Report Key Stats/Results
- T and Z Tests
    - Statistical Tests for data samples
    - Test depends on given information
    - Depending on number of samples, etc, test could change
    - With each test, there are a number of assumptions for the data
- P-Value and Confidence Intervals
    - Statistics that can determine if data is significant
        - If P-Value is more than 0.05, then data wouldn't be significant
        - If Confidence Interval contains 0, then data wouldn't significance
- ANOVA Test
    - Used for 3 or more samples
        - One Way: 1 variable differs across the populations
        - Two Way: How 2 categorical variables differ and categorical factors affect each other
        - Each carries their own assumptions