Project 1 Group 3

PROJECT PROPOSAL:

1. **Project Title:**
    a. **Pandemic-related Changes in Alcohol Consumption for U.S. States Using Available Data (i.e. 16 states and 4 months)**
2. **Team Members:**
    a. **Phil, Rob, Diane, Gurupdesh**
3. **Project Description/Outline**
    a. Alcohol consumption has been an important part of American culture since inception, data is available that allows us to analyze consumer preference-based on a variety of variables, race, income, region etc. During this time of COVID liquor stores were identified early on as essential services and remained open. We are interested in understanding the impact shutdowns due to COVID spread have had on alcohol sales, consumption and/or distribution.
    b. **Scope Statement: Compare alcohol consumption stats from NIH.gov (see below) for 16 states, years 2017-2020, months March, April, May and/or June 2020.**
4. **Research Questions to Answer**

    Question #1: What, if any, non-Pandemic year alcohol consumption rate changes occur with a given year? (i.e. Seasonality)

    Question #2: Did alcohol consumption change in 2020/Pandemic months from the prior year? (i.e. Variance)

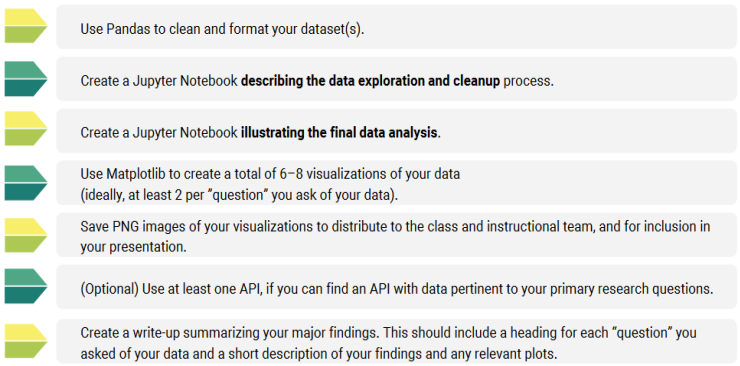
    Question #3: How did COVID shutdowns and infections rates affect alcohol consumption – if at all?

    Question #4: Which regions of the U.S. saw increases in alcohol tax revenue?

    Question #5: Did regions not affected by shutdowns see any significant changes in alcohol consumption?

    Question #6: Were alcohol consumption rates associated with infection rates? Or vice versa.

    a. **Task 3: Assign data review by state (4 states each for 16 states total) for NIH/Consumption, COVID Tracking and New York Times Shutdowns**
        i. Phil: Alaska, Arkansas, Colorado, and Connecticut
        ii. Diane: Florida, Illinois, Kansas, and Kentucky
        iii. GG: Louisiana, Massachusetts, Missouri, and North Dakota
        iv. Rob: Oregon, Texas, Virginia, and Wisconsin
5. **Project Requirements:**
    a. **Requirement #1: Pull, clean, merge data sources in Pandas**
    b. **Requirement #2: Jupyter Notebook 1: describing data exploration and cleanup process**
    c. **Requirement #3: Jupyter Notebook 2: illustrating the final data analysis**
        i. Perform statistical analysis on merged data, look for correlations in Jupyter Notebook
        ii. Pull insights relevant to research questions from data



d. **Requirement #4: Use Matplotlib to create 6-8 visualizations, or 2-3 per "question"**
    i. Including a map
e. **Requirement #5: Create and save PNG images of visualizations**
f. **Requirement #6: Prepare/draft presentation write-up summarizing major findings**
    i. Use PowerPoint, or Word, save as PDF but also in the Markdown README.md?
    ii. Include a heading for each "question" you asked of your data and a short description of your findings and any relevant plots.

6. **Rough Breakdown of Tasks**
    a. **Task 1: Identify data sources – first thing Saturday [DONE]**
    b. **Task 2: Describe Data (pull data definitions/explanations from the following websites:**
        i. **Datasource #1:**
            1. **NIH/NIAAA Alcohol Consumption by state for March, April, May and June 2020**
            2. URL: https://pubs.niaaa.nih.gov/publications/surveillance-COVID-19/COVSALES.htm
                a. Identify seasonality patterns of alcohol consumption for prior years, by month.
                b. Run statistical analysis on pandemic-specific months, (Mar – Jun 2020) to identify pandemic-related changes in alcohol consumption.
            3. NIH data on Pandemic Alcohol Sales: This file contains data on per capita alcohol sales from 16 states (Alaska, Arkansas, Colorado, Connecticut, Florida, Illinois, Kansas, Kentucky, Louisiana, Massachusetts, Missouri, North Dakota, Oregon, Texas, Virginia, and Wisconsin) by type of alcoholic beverage from January 2017 through June 2020. For 2020, currently available for the months of March, April, May and/or June 2020.
        ii. **Datasource #2:**
            1. **OPTIONAL API: The COVID Tracking Project, testing, infection, hospitalization and death rates by state.**
                a. URL: https://COVIDtracking.com/data/api
                b. Testing Rates by State
                c. Positive Test Rates by States over time
                d. Total Cases by State – Infection Rates
            2. The COVID Tracking Project: COVID, Jan 22, 2020 thru Oct 9, 2020 (updated daily) testing, infection, hospitalization and death rates
            3. The COVID Tracking Project | Data API
            4. The COVID Tracking Project | Our Data
        iii. **Datasource #3:**
            1. **The New York Times Shutdowns and Reopenings by State for the level of shutdowns**
                a. URL: https://www.nytimes.com/interactive/2020/us/states-reopen-map-coronavirus.html
                b. Create a score for level of shutdown by state (see: NYTimes link)
                c. Compare two states with the most and least level of shutdown?
        iv. **Others as time permits!**
            1. The Cocktail DB
                a. URL: https://www.thecocktaildb.com/api.php
            2. Open Brewery
                a. URL: https://www.openbrewerydb.org/
        v. Others as needed!