Summary:
It's been two years since the 2020 Election and six years for the 2016 Election. I think it's a great time to start analyzing connections between Election and COVID-19. Politics connects a lot of things in our life but lots of people barely care about it. A couple of questions I would like to find out is: Is there a relationship between how many people got COVID vs which part they vote for? Does the seriousness of COVID-19 affect how people vote in a certain way? Is there any chance that the policies made by administrations between 2016-2020 affect the voter's decision?

Datasets:

1. 2016 Election
    a. Size: 50 MB
    b. Location: Kaggle
    c. CSV files
    d. Download into a local computer, then load using Pandas
2. 2020 Election
    a. Size: 3 MB
    b. Location: Kaggle
    c. CSV files
    d. Download into a local computer, then load using Pandas
3. COVID 19 data
    a. Size: 75 MB
    b. Location: Kaggle
    c. CSV files
    d. Download into a local computer, then load using Pandas
4. COVID 19 Data (JSON)
    a. Size: 446 KB
    b. Location: GitHub
    c. JSON files
    d. Download into a local computer, then load using JSON

Preprocessing:
Fortunately, the datasets from Kaggle are usually well-cleaned before publishing. For the JSON one, I have to clean the data a little bit because it contains a lot more irrelevant data that I will not be using such as "reproduction_rate" or "icu_patients". All of them will be stored in a folder, I plan to group data based on countries, then I can map COVID 19 data with voter's votes based on State or County. Once I merged the CSV files, I was able to have data that relates to elections and COVID. I can then perform analysis on the data I found.

Data visualization:
I will create a bar chart that shows the voter's decision for both party and COVID 19 cases specifically in 2016 and 2020. I believe I will have more results once I started working on it. But

so far just by glancing at the data I have. I think this would be my first step on visualizing the data.