James Yang

DATA 512

Part 1 – Common Analysis

Explanation for Visualization

The visualization shows the dates vs the confirmed amount of cases with the corresponding rate on a color coded map of New Jersey in the county Bergen. The visualization itself contains all moments of masking policies in effect because New Jersey never had a moment where it didn't have masking policies through the term of COVID. The rates increase a bit through mid November 2020 to May 2021 before flattening out again.

In order to standardize and clean the three datasets, we needed to link the raw us deaths with the state associated to the New Jersey state and the columns associated with the Bergen county. Another thing we needed to clean is the dates associated with each of the states and the FIPS value that links with the mask use by county. Lastly, the number of deaths and the date associated with each date needs to be standardized across each of the datasets so they can be interpretable. These are the inner workings of the plot that was produced.
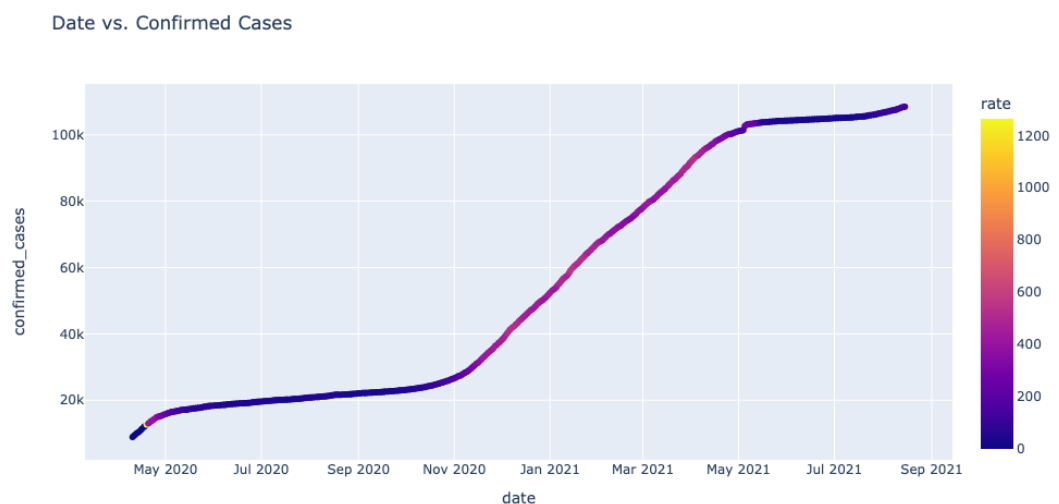

Figure 1.1: Depicts the visualization.

Reflection Statement (on what I got out of the collaborative assignment)

Something that I learned about answering this research question is that there is a creativity aspect to developing an interpretable plot that gives meaningful information. Originally, I had planned on using a ruptures module to create predicted segmentation. However, the plot I produced with that method wasn't that interpretable and didn't give great results. I moved on to a simplistic rate plot and realized that this plot gave me a much better interpretation of the rate with respect to the confirmed cases.

The possibility of collaboration on this part assisted in avoidance of pitfalls or other redundant experiment that others encountered. Being able to read other people's interpretations of different methods allowed for my general hypothesis that complicated does not always necessarily return better outputs. However, collaboration can sometimes pigeon-hole thought process and different perspectives on a problem. It's easier to direct a solution towards someone else's than it is to independently think of a solution. I still believe that collaboration is more representative of real world experience and industry typically encourages collaboration rather than individualism. The reason behind this is because it can produce legacy code, dropped projects, or a lot of lost money.

I believe that this section of the project was a good thought exercise with low stakes that allowed students to express their creativity and showcase their thoughts.