**DATA 230 Project Report: COVID-19 in the U.S.**

Vincent Zhang

SID:016741595

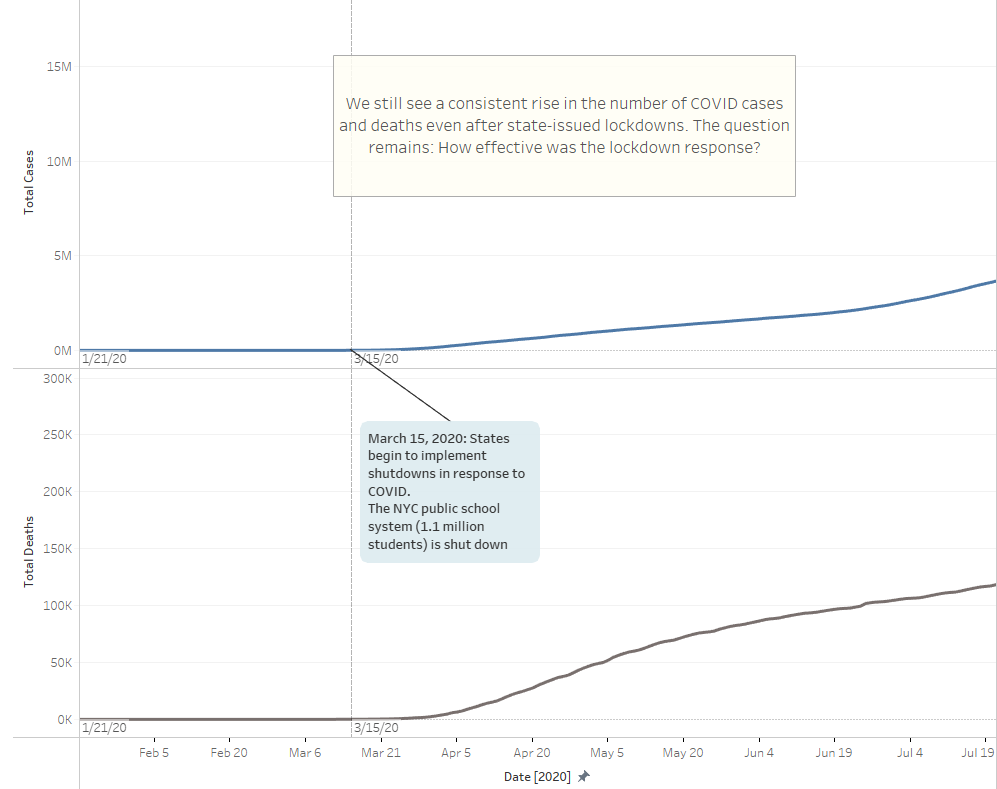
GitHub Code: <https://github.com/ErdMocha/Data230.git>

Abstract:

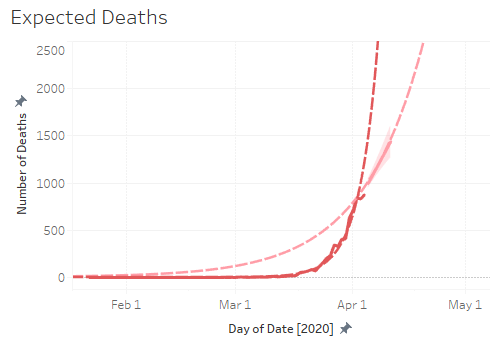
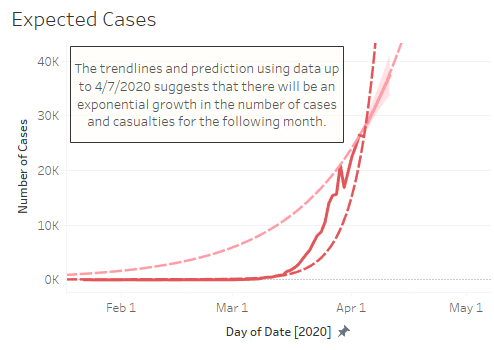
Having worked in the healthcare field, I want to use healthcare related data to showcase useful insight that can positively influence healthcare and patients in the future. Over the past 3 years, Coronavirus Disease 2019 (COVID-19) had a significant impact on American healthcare, economy, and even lifestyle. I created a visualization to break down COVID-19 cases in the U.S. to analyze the impact of major developments within the virus’s timeline as well as how the number of cases and deaths changed in response to government intervention. I started by gathering data from numerous sources and articles, obtaining my main data source from Kaggle and CDC while also referencing articles and excel files published by internet users for data cleaning and profiling. I used Tableau Prep Builder to aid with my data wrangling and I used Tableau Desktop to visualize the data and created an interactive dashboard giving users the option of narrowing in on data within a specific state. I hope to utilize this visualization to give users a perspective on how COVID spread throughout the U.S. and compare the impact that it had on each state.

Patterns and Findings:

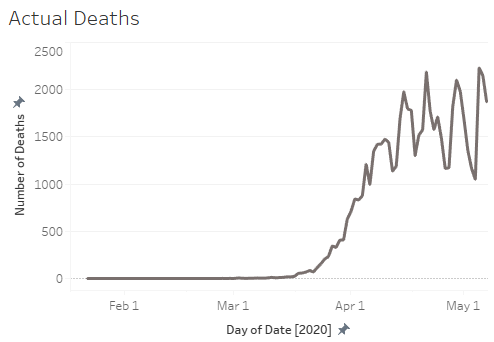
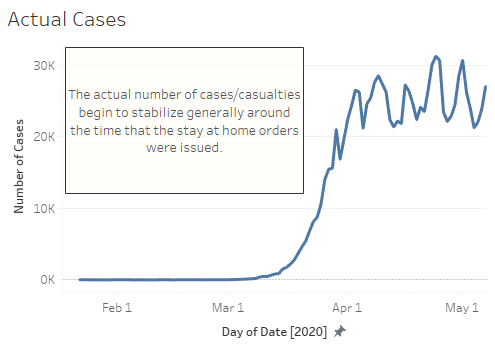
1. Stay-at-Home Orders
   1. Following the first State-issued stay-at-home order on Mar. 15, 2020, we still see a steady climb in the number of total cases.



* 1. However, if we change our perspective to look at the number of new cases and deaths up to April 7, 2020 (last stay-at-home order by South Carolina), the number of expected cases and deaths follows an exponential trend and is expected to hit over 40000 cases and 2500 deaths per day by May 7.

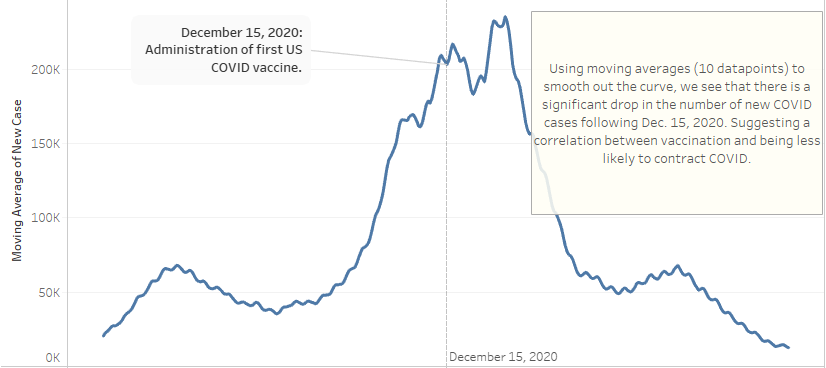


* 1. Compared to the actual data, the number of new cases averages out at around 25000 per day over the next month. Similarly, the number of deaths averaged out at around 1600 per day for the next month.

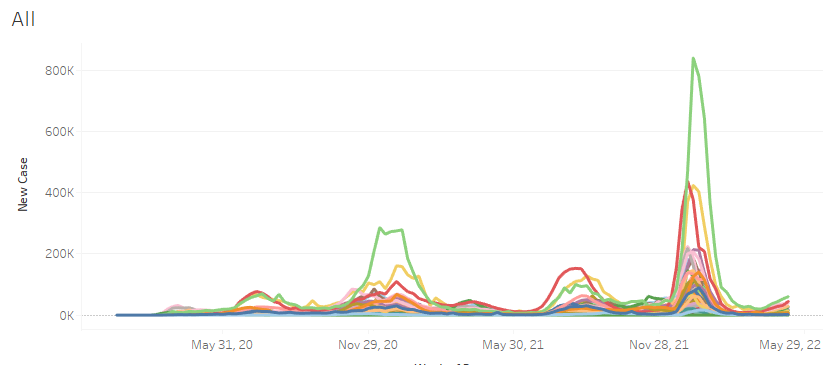


* 1. The data suggests that the government stay-at-home orders did hinder the spread of COVID-19.

1. Impact of the COVID-19
   1. With the following visualization, we can see a drop in the number of new COVID cases following the start of vaccination. This suggests there is a correlation between COVID vaccines and enhanced immunity against the virus.

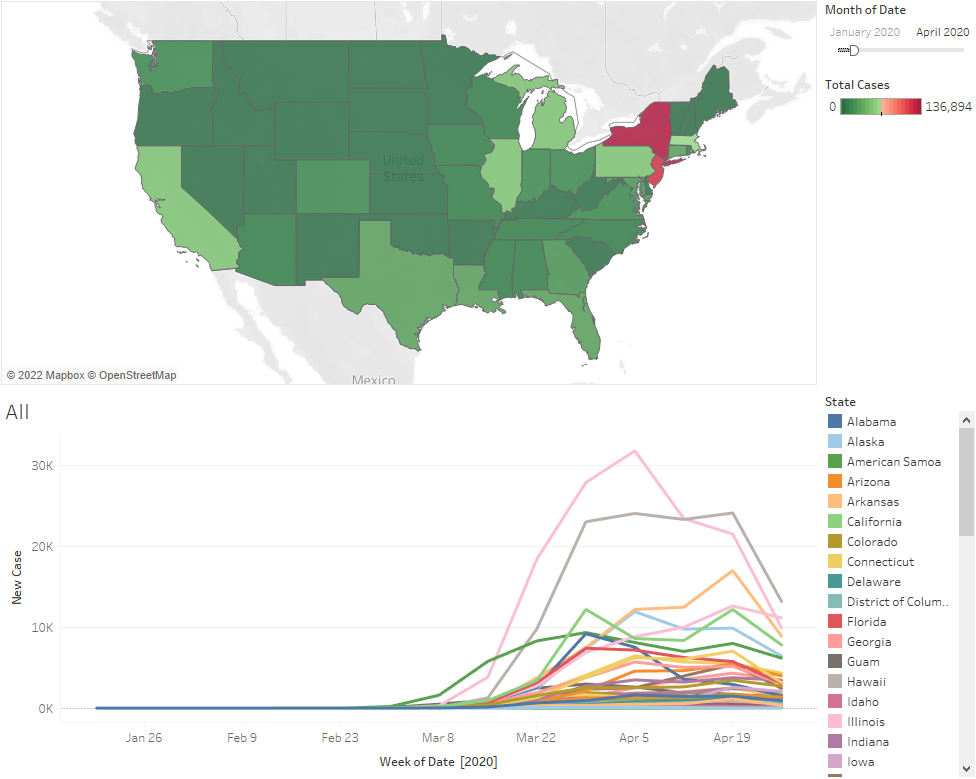


* 1. However, the spike and drop in recorded cases also correspond to the pattern of spikes in new cases during the winter holiday season.

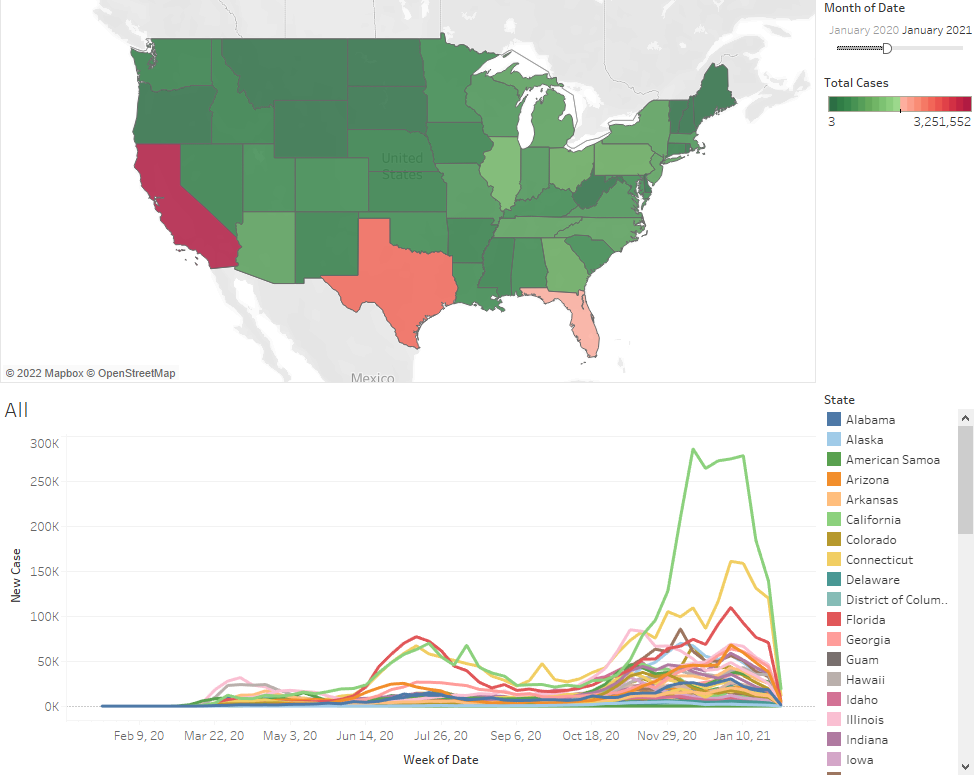


* 1. Although the COVID vaccine does seem to correlate with improvements in immunity against the virus, we cannot entirely attribute the drop in cases to the vaccine.

1. COVID Impact by State
   1. For the first 4 months of COVID, we can see that New York and New Jersey were the most heavily impacted states.



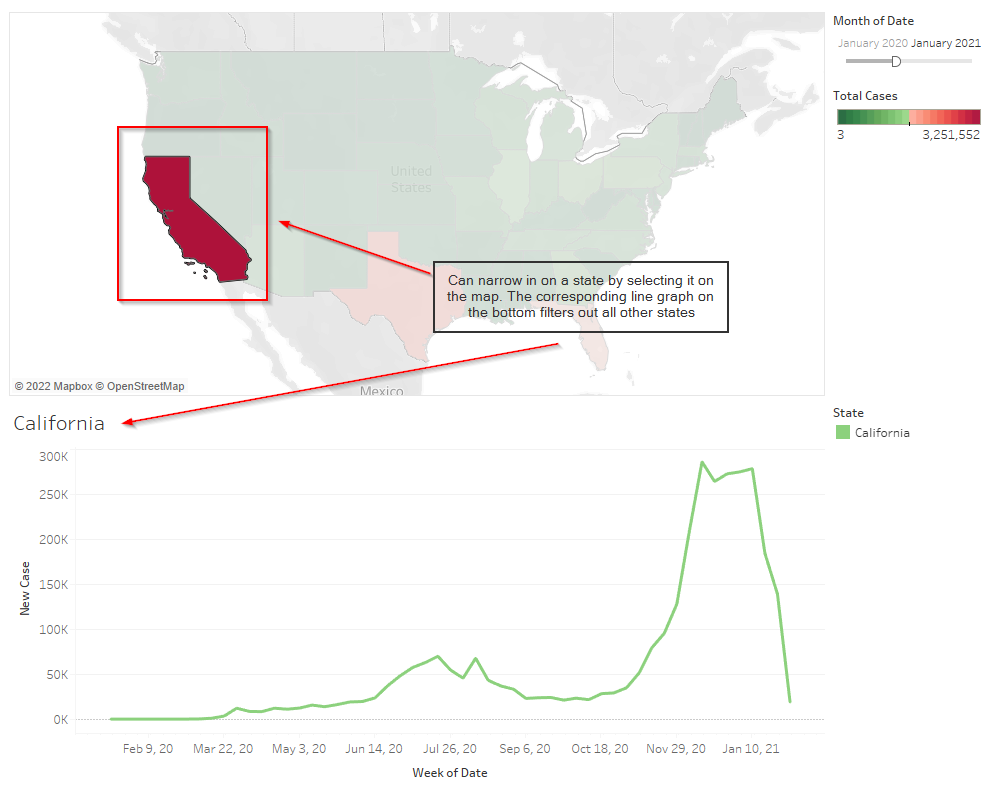
* 1. However, as we extend our timeline to one year, the most heavily impacted states shift to California, Texas, and Florida



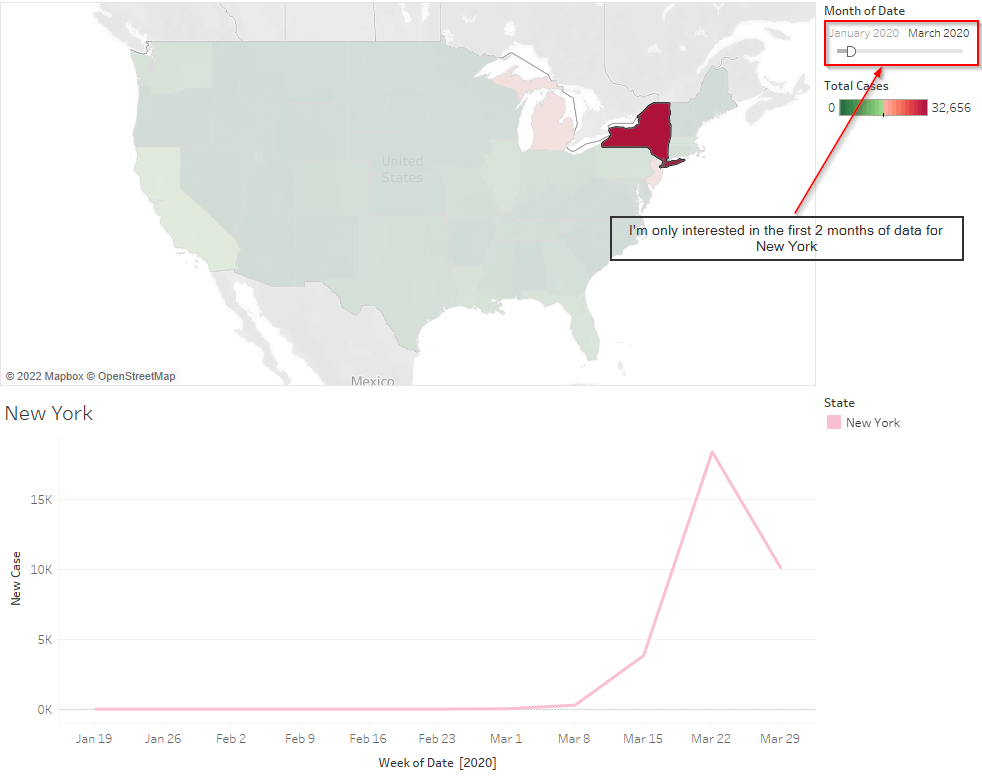
* 1. From this visualization comparison we can draw the conclusion that New York’s COVID policies were more efficient at slowing the spread of COVID compared to that of California, Texas, and Florida.

Use Cases:

1. Users can narrow-in on a specific state for detailed analysis:



1. Users can indicate their interest in a specific period of time after the COVID outbreak (from Jan. 21 onwards).



Data Source:

* <https://www.kaggle.com/datasets/sudalairajkumar/covid19-in-usa>
* <https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36/data>
* <https://www.ajmc.com/view/a-timeline-of-covid19-developments-in-2020>
* <https://community.tableau.com/s/question/0D54T00000C6egqSAB/display-state-names-instead-state-acronyms>
  + 