

ANALYSIS ON COVID 19 AND EXISTING HEALTH CARE SYSTEM

SUMMARY

COVID - 19 is an ongoing pandemic affecting the lives of millions of people worldwide. America has been touted to have one of the best healthcare systems in the world, but sadly, it is the country with the maximum number of cases in the world. In our dashboard, we analyze data from hospitals across all states to visualize if the American healthcare system can handle the pandemic.

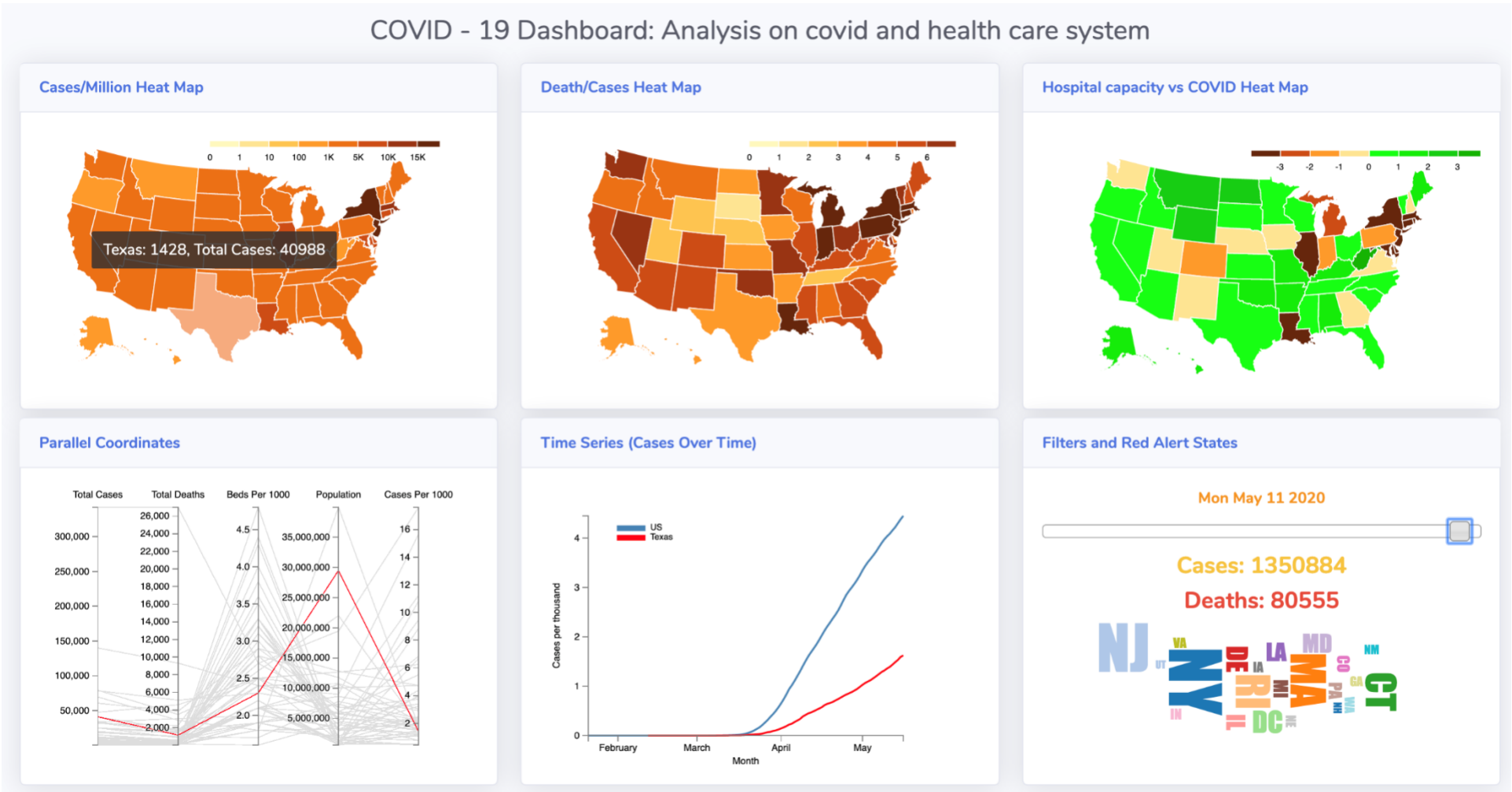
DATASET

We have fused 3 different datasets for the purpose of our analysis:

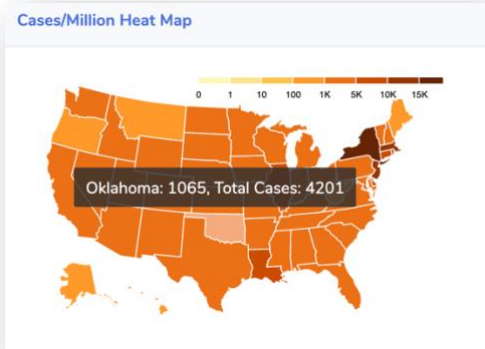
- 1. Covid: <https://github.com/nytimes/covid-19-data>
- 2. Population: <https://worldpopulationreview.com/states/>
- 3. Hospital beds: <https://worldpopulationreview.com/states/>

Data Field	Dataset	Description
State	Covid	The name of the state
Cases	Covid	The number of cases reported for the day
Deaths	Covid	The number of deaths reported for the day
Date	Covid	The date corresponding to the data
State	Population	The state and its corresponding population in the next column
2020 Pop.	Population	The population for the state
Location	Hospital Beds	The state corresponding to the hospital beds
Total	Hospital Beds	Number of beds/1000 in considered state

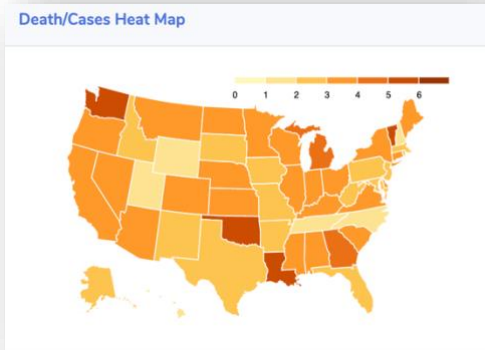
LAYOUT



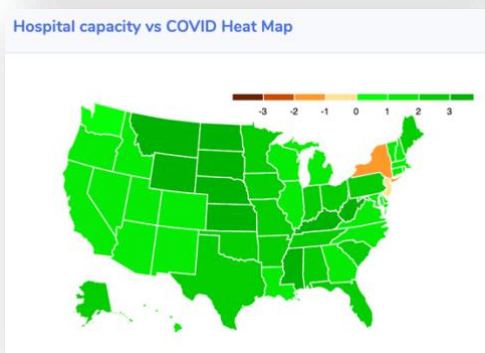
VISUALIZATION ELEMENTS



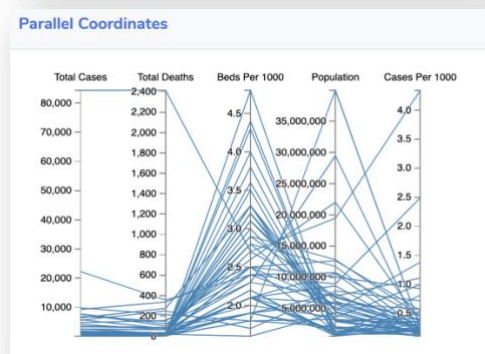
We observe that states with high cases doesn't necessarily mean higher alert. As we can see from the density map, it is cases/population that is apt and it remains pretty much the same for most of the states (ex. California and Arizona)



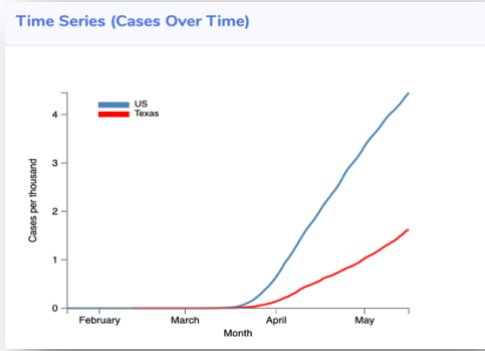
This shows the death % of people with COVID. This in correlation with hospital heat map explains if there is a correlation between hospital occupancy and death rate. Also NY, etc had higher death rates at the end of May 15



This shows how well can the hospitals handle current cases, given their bed capacity. A state is shown green in color if the number of cases < number of hospital beds. Comparing this with chart 2, we see that there is a +ve correlation between death rate and occupancy status i.e., greater the cases than hospital beds, higher the death rate. Ex. NY, Illinois, Michigan etc



A parallel coordinate chart will help us to better compare multiple critical components for each state. For example, we can find out which states are doing better/worse by filtering on the beds per 1000 attribute and also the cases per 1000 attribute. For example, NY/California, NY/Florida

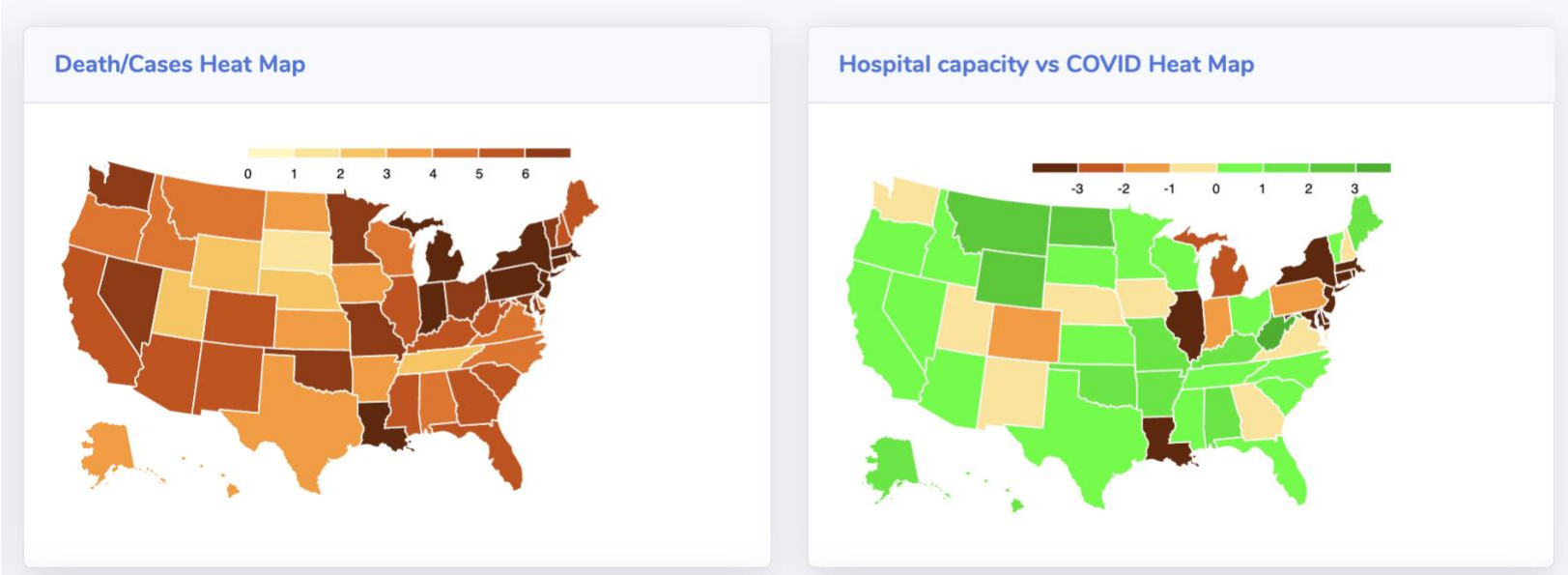


This time series plot will reveal the growth of cases per 1000 when compared to the national average and can help us track the trend of cases in each state. For example, the growth in Nebraska was not as steep in April as it is in May, but it has overtaken the national average as well now, which could suggest of an increase in the number of tests, or the actual number of cases.

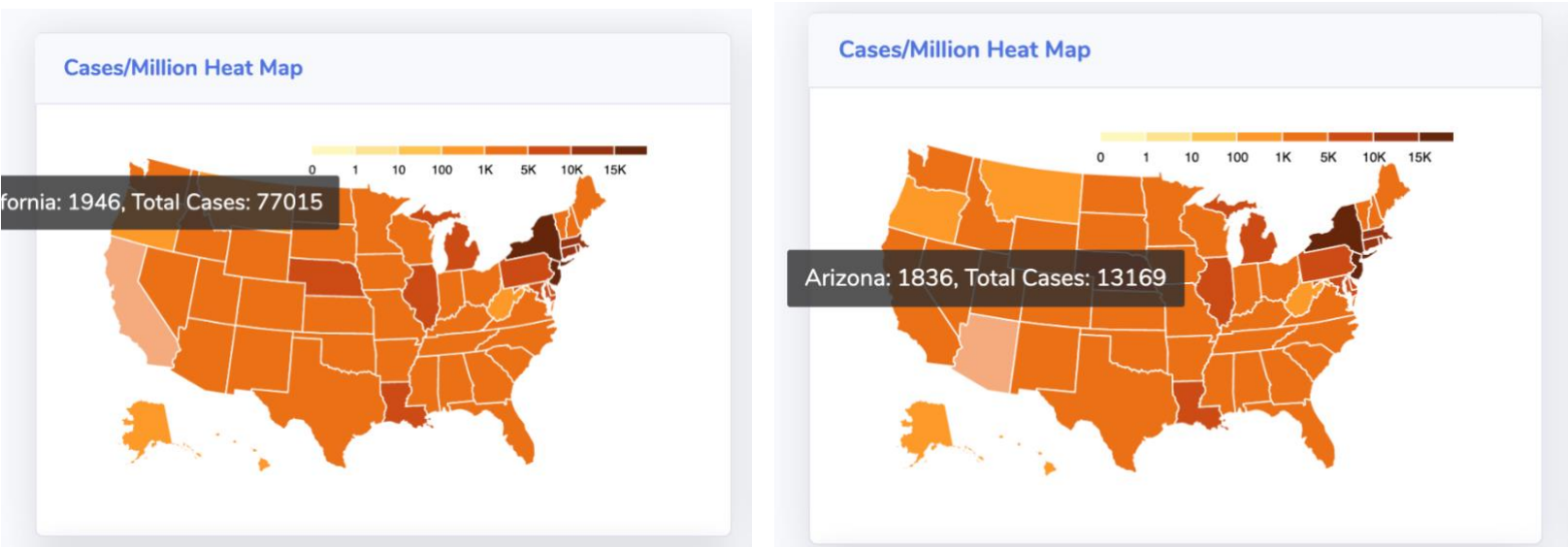


This has date filter for entire dashboard. It shows current cases and deaths along with the highly risked states of corona in a wordcloud, which could put it's residents at a higher risk.

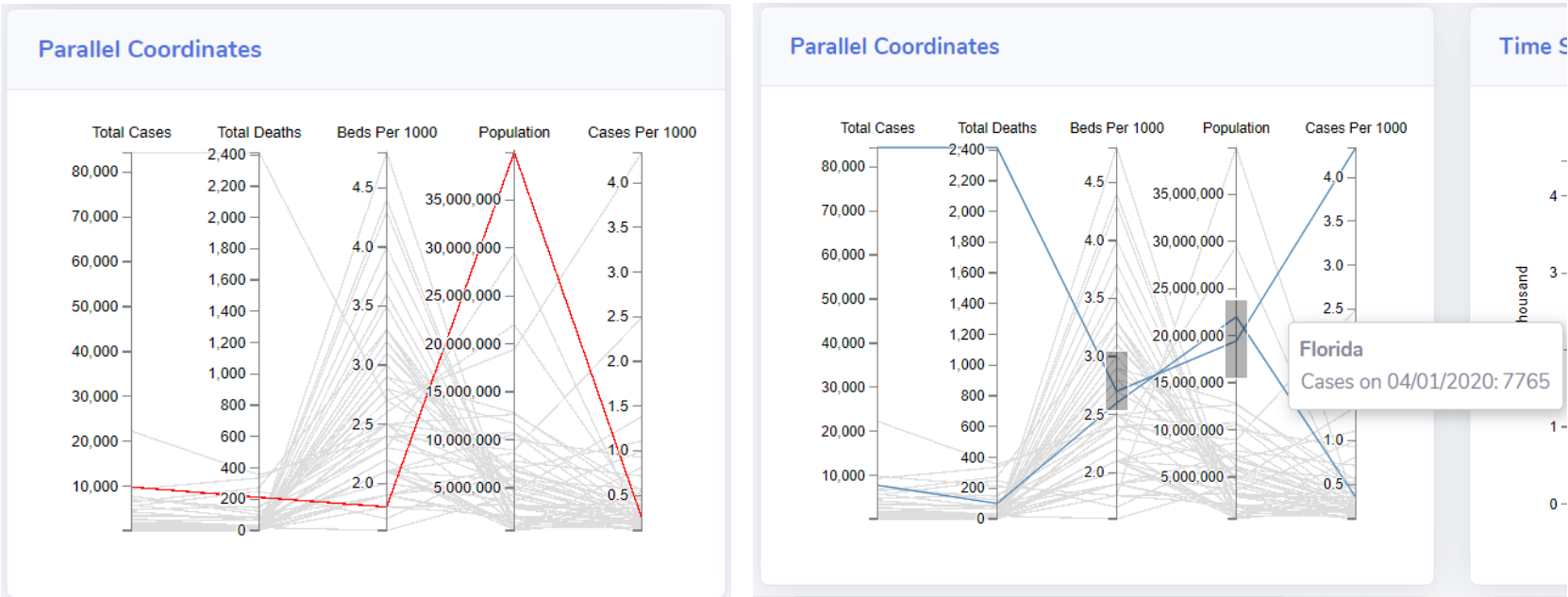
OBSERVATIONS



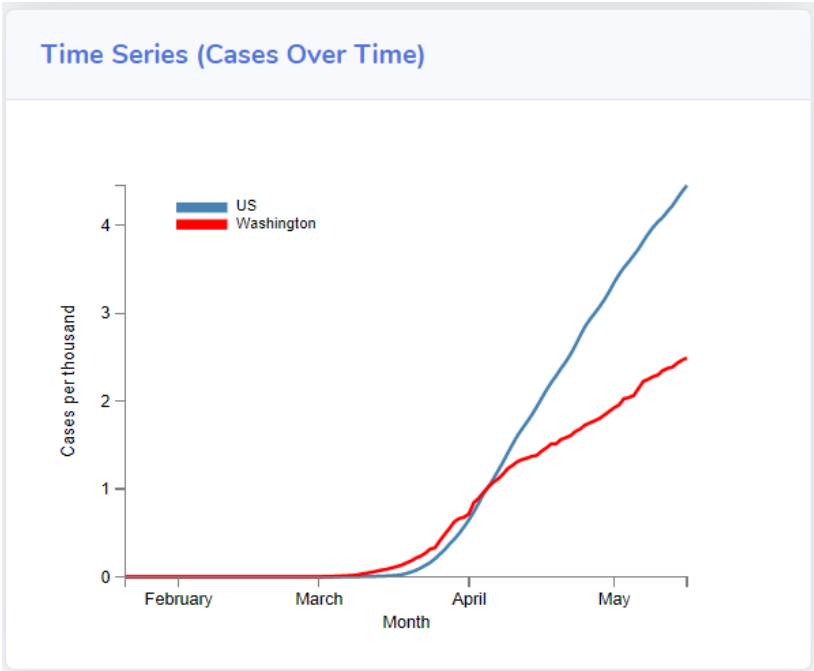
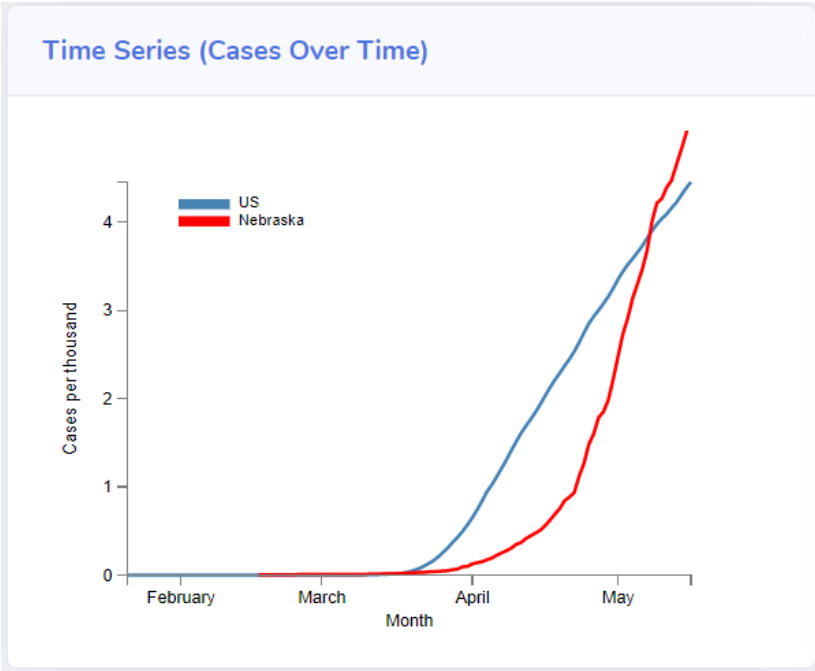
- From the death/cases heat map and hospital beds vs COVID heat map, we see that a higher death percentage is positively correlated with hospital bed shortage i.e., greater than difference between current cases and hospital beds available, higher is the death percentage. It is evident from states like New York, Michigan, New Jersey in both of the above heat maps.



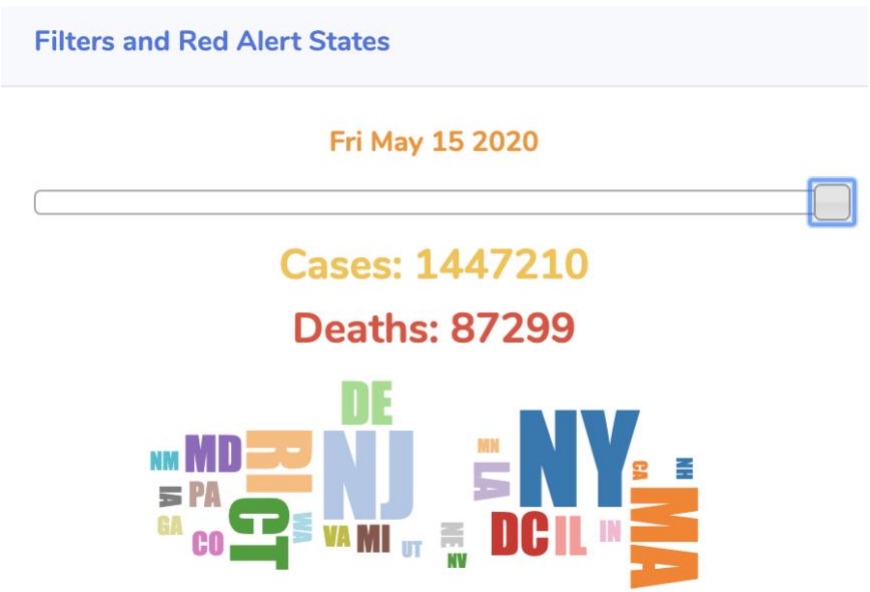
- Though California seems to have 7x number of cases compared to Arizona, the average cases per 1000 is still the same for both of them suggesting that they are at an equal danger level. The same is the state for many other states. They have the number of cases in proportion to their population leaving them equally susceptible to infections.



- The Parallel Coordinates chart is being used to visualize multiple aspects of the data at the same time. We can draw various observations by filtering out a variety of data over here. One interesting observation that we found was the the number of beds per 1000 of California was very low as compared to other states, and also it is the most populated state in America, but the number of cases per 1000 was surprisingly low as compared to other states with better healthcare to population ratios.
- Another observation was that for the states of Florida and New York. Both states had similar healthcare systems and similar population as well, but the number of cases per 1000 have a huge difference for both the states. This raises questions about the handling of the pandemic by both the states, and what went right for Florida but wrong for New York



- The time series plots here plot the cases per 1000 for each selected state and compare it with the national average. One observation was made for Nebraska, in which the cases were not so many till mid-April, and then we see a spike in cases in May, which even overtakes the national average at the start of May.
- Also in Washington, the average was greater till the start of April, but then reduced a lot after that owing to great measures taken. Thus we can track the growth of the virus in individual states using this time series graph.



- By the end of our period of analysis, from the word cloud, we observe that New York is at the highest risk due to Covid infections. It is followed by New Jersey, Massachusetts, Rhode Island, Connecticut, Delaware. It is better to avoid these locations at the current moment.

TECHNOLOGIES AND TOOLS USED

- Bootstrap cards for the 6 card layout
 - Specifically card-deck for the alignment of cards in each row
- D3 js for visualizations
 - topojson for d3 geo maps
 - d3-tip for tooltips
 - d3-cloud for wordcloud

GITHUB LINK:
<https://github.com/krishshanmukh/Covid-Visualization>

YOUTUBE LINK:
<https://youtu.be/kPVRALynOK8>

MEMBERS:
 Sri Krishna Shanmukh Medapati (112688738)
 Saurav Singh (112733214)