

Tableau Assignment 5

Introduction:

This dataset contains the daily data of covid-19 pandemic for the individual states in the US. This tracking project was done from 13th January ,2020 to 7th March ,2021 across 56 states of US.

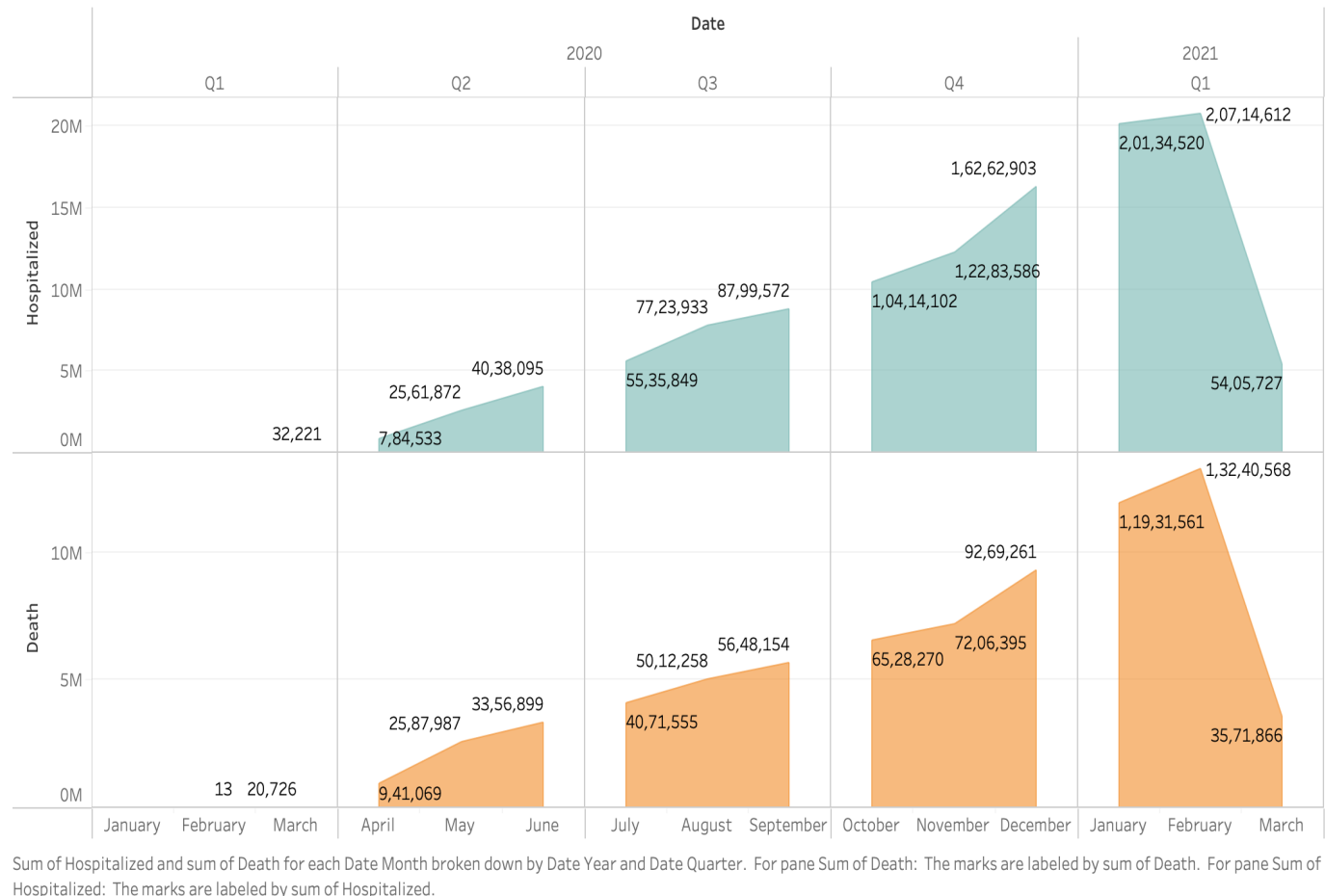
Link: <https://covidtracking.com/data/download>

This dataset contains the following information:

- Date-The dates between which the tracking project was performed.
- States- The US states involved.
- Death-The number of deaths that occurred due to covid-19.
- Death Increase- The increase in the number of deaths that occurred due to covid-19.
- Hospitalized Cumulative- The cumulative number of people hospitalized due to covid-19.
- Hospitalized Currently- The total number of people currently hospitalized due to covid-19.
- Hospitalized Increase- The increase in the number of people hospitalized due to covid-19.
- Positive-The number of total test cases that were found to be positive.
- Positive Increase-The increase in number of test cases that were positive.
- Negative-The number of total test cases that were found to be Negative.
- Negative Increase- The increase in number of test cases that were Negative.
- In ICU cumulative-The cumulative number of people kept in ICU's (Intensive care unit) as they were diagnosed with covid.
- In ICU currently- The total number of people currently kept in ICU's (Intensive care unit) as they were diagnosed with covid.
- On ventilator cumulative- The total number of people on ventilator as they were diagnosed with covid.
- On ventilator currently- The total number of people currently on ventilator as they were diagnosed with covid.
- Total tests viral- The total number of people who have at least once been tested using PCR testing.
- Total tests Antibody-The total number of antibody tests that we collected.
- Total tests Antigen- The total number of people who have at least once been tested using antigen testing.
- Total test results-The total number of covid tests conducted.
- Recovered – the total number of people who have recovered.

1. Time series analysis done by plotting Sum of hospitalized and sum of deaths vs time divided Yearly, quarterly, and monthly.

Sheet 1

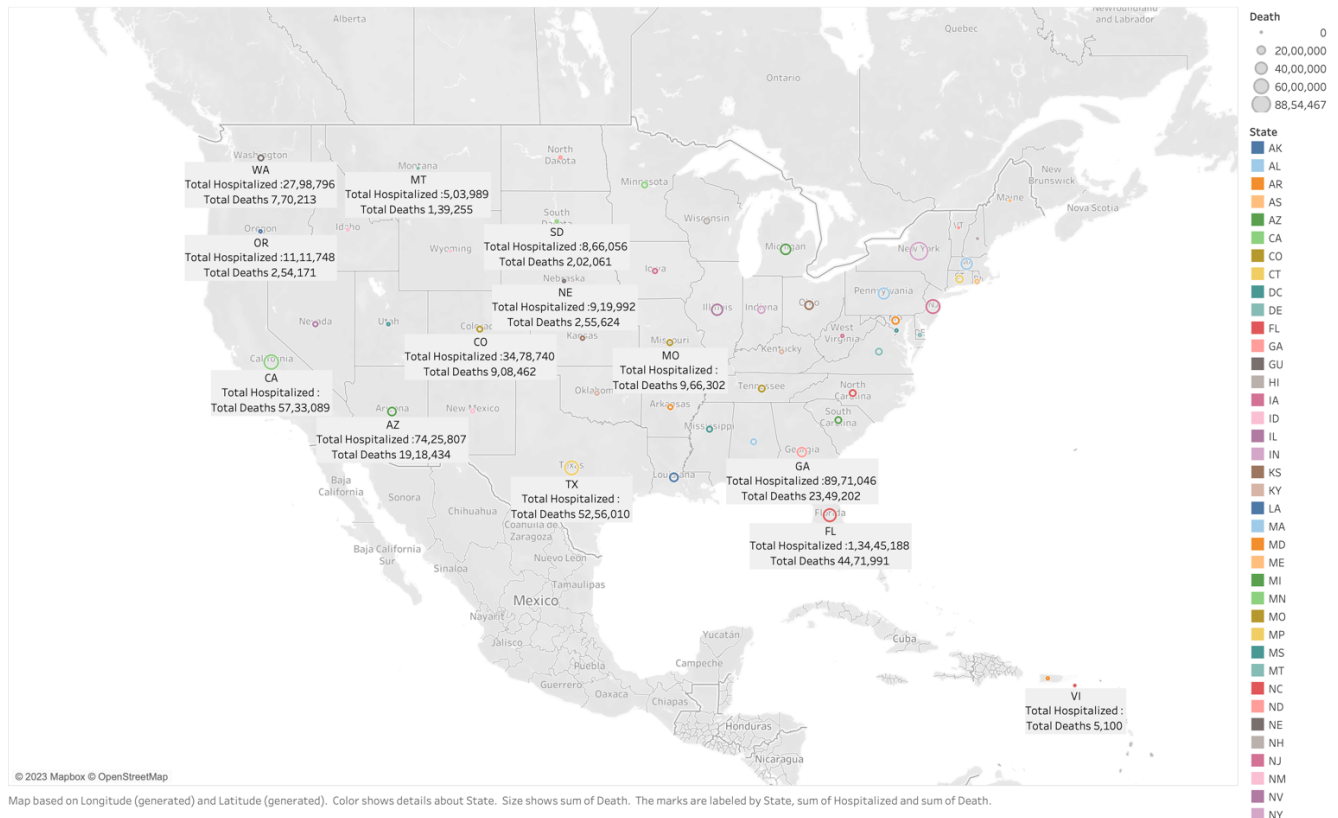


The above Time series analysis is done by plotting an area chart of Sum of hospitalized and sum of deaths vs time divided into first Yearly then quarterly and finally monthly.

- First , we see there has been significant increase throughout the year 2020 with the highest being 16mn hospitalized cases and 9.2mn deaths due to covid.
- The highest number of hospitalized reported were 20mn cases in the month of February which is the maximum among all the months.
- And the highest number of deaths reported were 13mn cases also in the month of February being the maximum among all the months.
- Then there was a sharp sudden decline in the number of cases as we go from February 2021 to March 2021 with a hospitalized cases being 5.4mn and deaths reported being 3.5mn.
- We can see the sudden increase had started from November and went till February.
- The sharp decline from February could mean that proper vaccinations were conducted.

2. Plotting a geographical map based on latitude and longitude in each state of US labelled with the total number of Hospitalized and deaths and size of the circle in each state showing the number of deaths.

Sheet 2

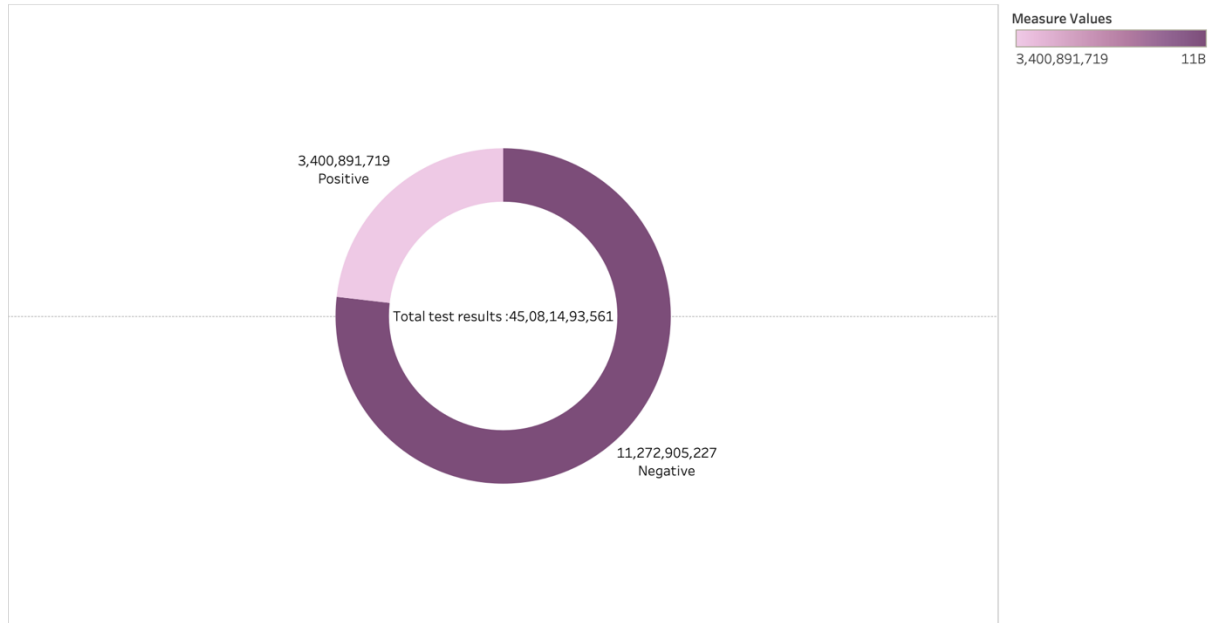


As we have plotted a geographical map using latitude and longitude. The size of the circle shows the total number of deaths. The larger the circle, the more are the number of deaths. The states have been color coded.

- We can see that some of the states like Texas, California, Missouri, Michigan and New York do not have the number of hospitalized values.
- Now, the states Wyoming, Alaska and Maine have less than 100k deaths.
- Rest all the states have deaths over 100k.
- The number of hospitalized is very low (1696 patients) in Iowa compared to the number of deaths at 650k.
- Coming to the east side, we see that number hospitalized is in millions except in Maine and the maximum in Florida with 4.4mn deaths and 13mn hospitalized.
- Now coming to the west side, we see that the number of hospitalized is the highest in Arizona with 7.4mn and the California with the highest number of deaths with 5.7mn.
- Overall, Geographical maps give an idea as to which areas were affected the most so that the required measures can be taken so as to what can be done to improve it.

3. Making a doughnut chart to analyse the totalTestResults , positive and the negative results.

Sheet 4

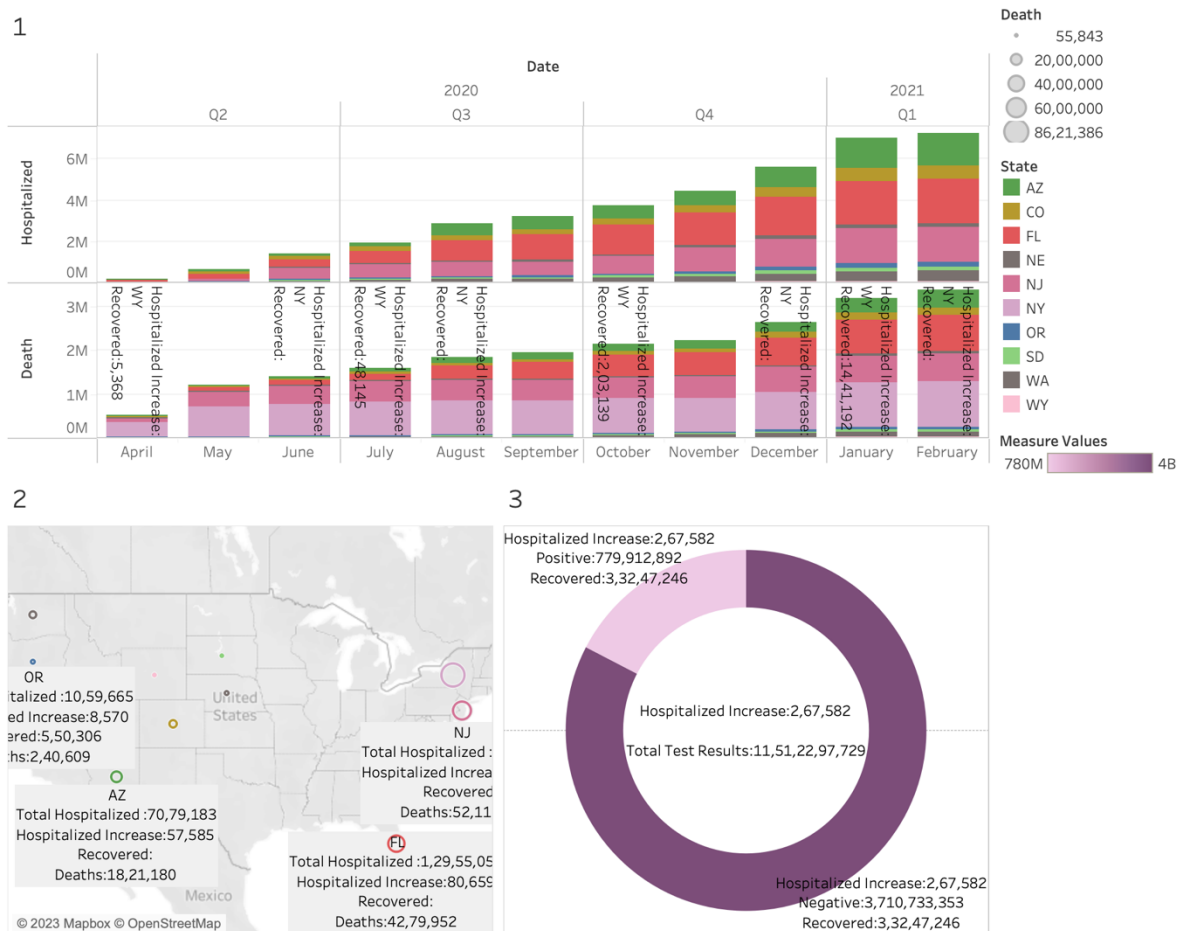


Zero and Zero. The marks are labeled by Negative, Positive and Zero. For pane Sum of Zero: Color shows Negative, Positive and Zero. The marks are labeled by Negative, Positive, Zero, Negative, Positive and Zero. Details are shown for Negative, Positive and Zero. For pane Sum of Zero (2): The marks are labeled by sum of Total Test Results, Negative, Positive and Zero.

- We see in the above doughnut chart that the total number of test results are 45billion meaning a significant number of tests were conducted as the population of US was around 329mn in 2020.
- Now coming to positive test out of the total number of test results which was 45bn, the number of positive test results were 3.5Bn , which is around 7.5% of the total test results .
- Now coming to the negative test results, we see that the number of negative test results are 11.2 bn which is around 24.8 , rounding of to 25% . There we can say that 1 in 4 test results turned out to be negative.

4. Creating a dashboard using the above 3 plots and putting my choice of parameter.

The parameter that we have chosen for better analysis is hospitalized increase and recovered over 10 states from 1st March 2020 to 1st March 2021



For this dashboard , I have used filter pane to include only the dates from 1st March 2020 to 1st March 2021. I have also taken only 10 states to give a better analysis. The 10 states are Arizona (AZ), Colorado(CO),Florida(FL), Oregon (OR), Wyoming(WY), New Jersey(NJ), Washington(WA),New York(NY) , South Dakota(SD) and Nebraska(NE). Only for the First graph, we have excluded the month march for clearer analysis.

The parameter that we have chosen for better analysis is hospitalized increase and recovered over these 10 states within 1st March 2020 to 1st March 2021

From the above dashboard we can conclude that, from graph 1 ,

- we see that Wyoming has had the most increase in Deaths.
- And Florida has had the greatest number of Hospitalizations over the entire year which increased significantly over the year, coming to a close second is the state New Jersey.
- The number of recoveries has also been the highest in the month of February with 1.4mn.

From graph 2 ,

- we see that most recoveries are from Nebraska with 15mn and the increase in hospitalized is 6632.
- Florida has the highest hospitalized increase with 80k but lesser recoveries than Nebraska with 1.2mn.
- NY showed the lowest hospitalized increase with 0 patients.
- The highest number of deaths occurred were in NY with 8.6mn.

In graph 3,

- The total test results in these 10 states over this time period 1st March 2020 to 1st March 2021 was 11.5Bn .
- The total number of positive cases were 779mn which is 6.7 % of the total test results and negative was 3.7bn which is 32.2% of the total test results.
- The total number of Hospitalized increases was 267k.
- The total number of recovered were 33mn.