

SAIPAVAN TADIKONDA

MEMBER TASKS:

1. Generate weekly statistics for a specific state:

I have chosen North Carolina State. I was able to find the individual new cases and new deaths by using the diff function and converted the daily new cases and new deaths into the weekly data. And on this weekly data I determined mean median and mode for weekly new cases and weekly new deaths.

2. Compare the chosen state data with other five states:

The States which I have chosen for comparing are as follows:

- 1) California (CA)
- 2) Florida (FL)
- 3) New York (NY)
- 4) South Carolina (SC)
- 5) Texas (TX)

I obtained the individual new cases and new deaths across the selected state and normalized the obtained new cases and new deaths by population per 100000 as every state differs in the population. Then I converted the daily data into weekly data for further comparisons. The mean and median for weekly normalized new cases is as follows:

	State	mean	median
0	CA	18.598905	8.224941
1	FL	23.057242	16.455578
2	NC	19.905846	14.507621
3	NY	19.586660	9.408487
4	SC	24.523143	16.819734
5	TX	19.973268	13.540347

Fig. 1

From the above Fig1. We can observe that the mean is more for the South Carolina state followed by the Florida state. The mean of the California state is least among the six states.

We can also observe that the median for South Carolina is more followed by Florida.

The mean and median for weekly normalized new deaths is as follows:

	State	mean	median
0	CA	0.281679	0.173545
1	FL	0.296292	0.262065
2	NC	0.233098	0.174348
3	NY	0.474664	0.182119
4	SC	0.351741	0.246941
5	TX	0.315405	0.197072

Fig. 2

From the above Fig. 2 we can observe that the mean is more for the New York than other states. The North Carolina has least mean. Coming to the median the florida has the highest median.

Identify the peaks:

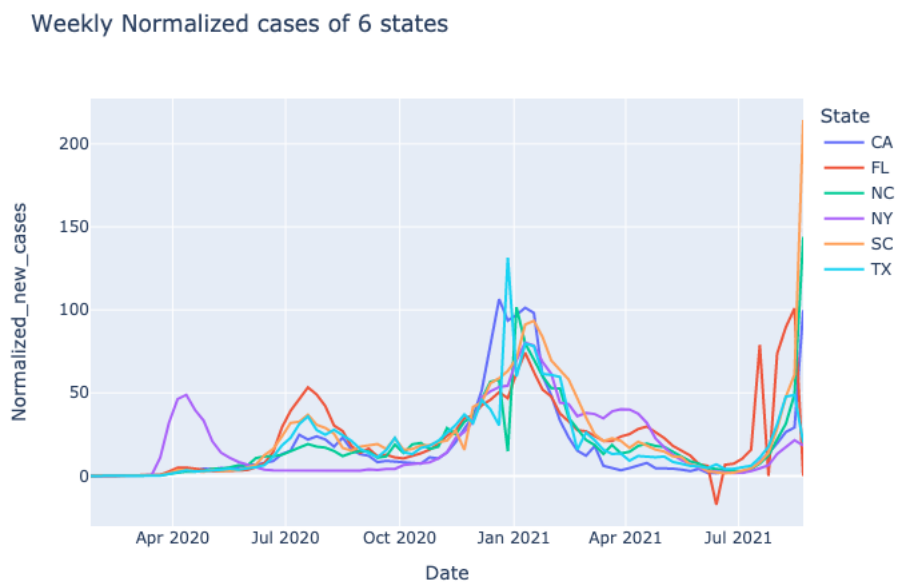


Fig. 3

From April 2020- July 2020 there is peak for the Normalized new cases for the state New York. From July 2020- October 2020 Florida State registered its peak. But from Nov 2020- March 2021 all the six states have registered peaks in between these months. After April 2021 there is a downfall in the cases across all the states. Again, after July 2021 we can observe the increase in cases. So, we can understand up to an extent that the spread of covid cases across these states is more in winter months compared to remaining months.

Weekly Normalized deaths of 6 states

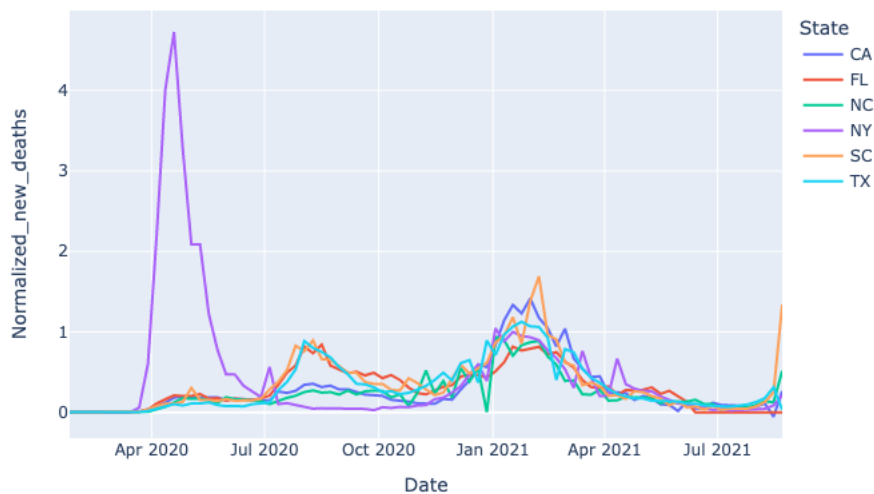


Fig. 4

From the above Fig. 4 we can observe that for the State New York in between April 2020- June 2020 there is large peak in the death. It registered its highest peak in between these months. And the states like Florida, Texas, South Carolina showed some peaks in between August 2020- September 2020. All the six states registered their peaks in between the Jan2021-March 2021.

Do they compare with US patterns?

Yes, they can be compared with the US patterns. If we observe US case patterns there are a greater number of cases and deaths in between November 2020-March 2021. The above taken six states also showed their peaks in between these months. So, we can say that these patterns matched with US patterns.

- 3) Identify top five counties in NC having high cases and death rates:
The top counties having high cases is as follows:

	countyFIPS	County Name	Population	new_cases	new_deaths	Normalized_new_cases	Normalized_new_deaths
77	37155	Robeson County	130625	19542.0	295.0	1496.0	23.0
81	37163	Sampson County	63531	8976.0	115.0	1413.0	18.0
47	37095	Hyde County	4937	690.0	9.0	1398.0	18.0
83	37167	Stanly County	62806	8716.0	142.0	1388.0	23.0
23	37047	Columbus County	55508	7697.0	162.0	1387.0	29.0

Fig. 5

Robeson County in NC had the highest cases.

The top counties having high death rates is as follows:

	countyFIPS	County Name	Population	new_cases	new_deaths	Normalized_new_cases	Normalized_new_deaths
61	37123	Montgomery County	27173	3609.0	97.0	1328.0	36.0
80	37161	Rutherford County	67029	8270.0	225.0	1234.0	34.0
65	37131	Northampton County	19483	1994.0	61.0	1023.0	31.0
51	37103	Jones County	9419	969.0	29.0	1029.0	31.0
23	37047	Columbus County	55508	7697.0	162.0	1387.0	29.0

Fig. 6

Montgomery County in NC had the highest death rate.

4) Plot the weekly trends for the top 5 infected counties:

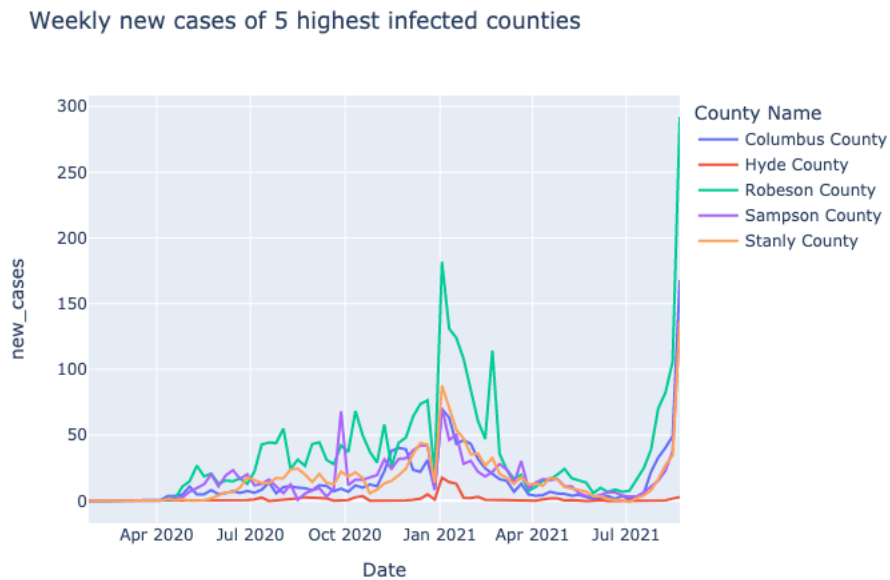


Fig.6

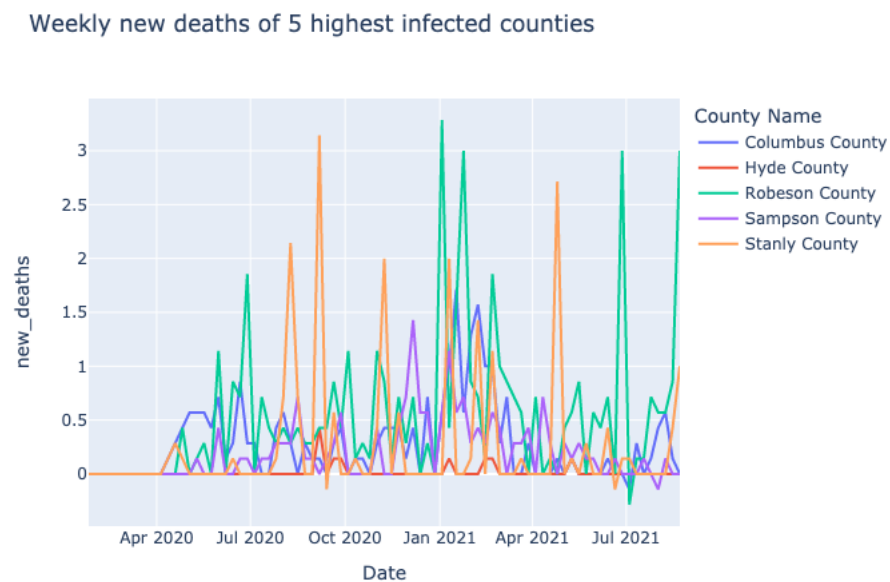


Fig. 7

From the above figures 6 & 7 we observe that all the top 5 infected counties registered their highest peaks in between Jan 2021 and March 2021. The reason might be because of cold climatic conditions the spread might be huge among these counties. After April

2021 there was complete downfall in the cases. And again, after July 2021 there is rapid increase in the cases. This might be because of a new variant among these counties. Coming to the deaths among these counties different county's registered their peaks in different months. For example, Stanly county had registered its peak in September 2020. This might be because of lack of Medical facilities in that county. Robeson County had registered its highest peak in January 2021. This might be because of cold climatic conditions.

Do the counties follow State pattern?

Yes, all these five counties follow State patterns. Even if we look into the NC state patterns highest peaks of cases and deaths are registered in between Jan 2021- March 2021. Even in the top5 counties the highest peaks are in between Jan2021-March 2021.