

COVID-19 ANALYSIS FOR SOUTH CAROLINA STATE

KEVIN YU

01/20/2021

BACKGROUND

- COVID-19 pandemic caused many illness and deaths of human. In the US, the situation is still very bad since the spread is not under control so far. Every state is fighting hard for the COVID-19 including the tracking and actively testing.
- In this project, we will follow the active cases daily and total cases for the South Carolina state. The data will be viewed through choropleth map to show the **hotspots** in the state. Then we will recommend the closest the clinic for the **testing**.
- This project will give the local community (our audience) visual information for the current Covid status. Also, this program can be scaled to small local area or other local communities based on the dataset.

METHODOLOGY

- Exploratory data analysis was conducted to understand the COVID case distribution in the states. The **folium map with choropleth** to visualize the COVID **hotspot**, which is quick and direct way to show the audience. Then **linear regression** analysis was conducted to understand the relationship between total cases and the population. **Polynomial regression** was further tested to check which is a better fitting for the data.
- After the data analysis, **Foursquare API** was utilized for searching the COVID test places around hotspot area. The locations are displayed on the Folium map to help people find the closest places.

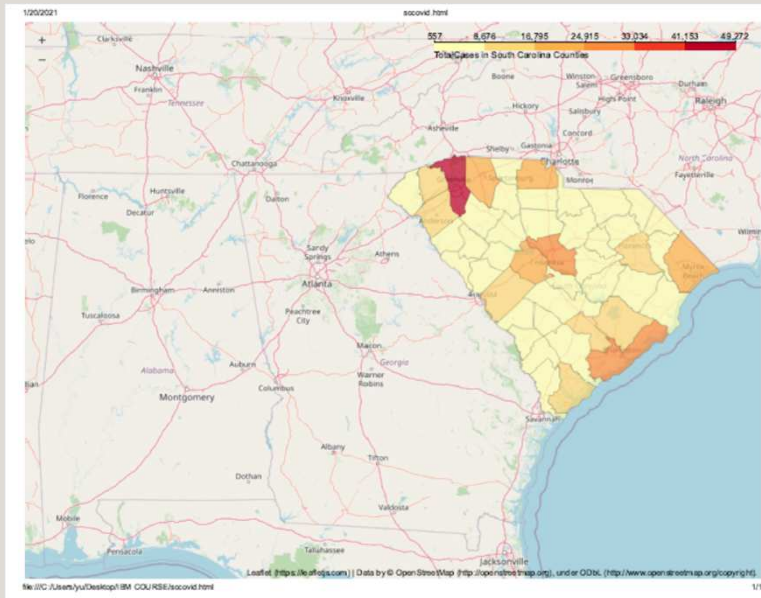
DATA SOURCE

- The COVID-19 data will be scraped from [worldometer](#) into data frame and cleaned. The data will be categorized into each county in the state. The data will include the 'total cases', 'new cases', 'total deaths', and 'new deaths'.
- The population data was scraped from 'https://www.southcarolina-demographics.com/counties_by_population'

DATA TABLE HEAD SAMPLE

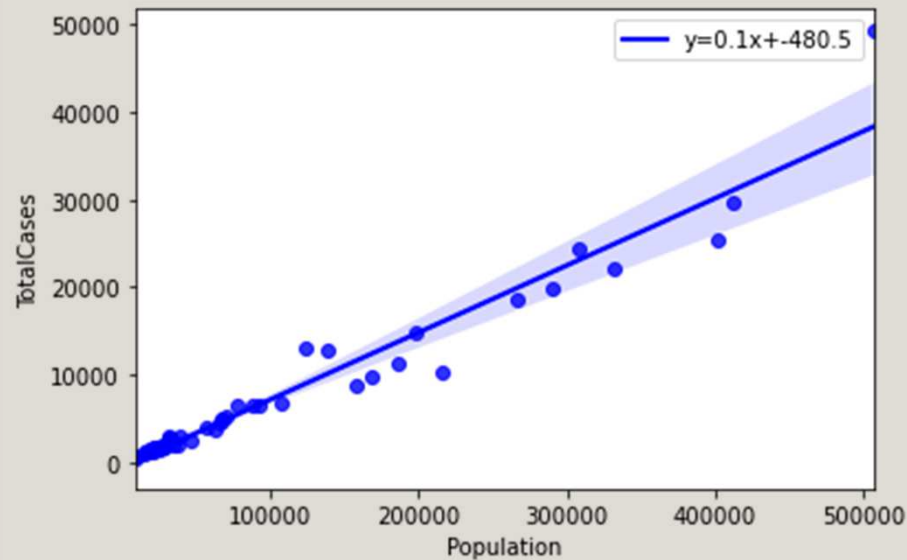
Date	County	TotalCases	TotalDeaths	Population
1/20/2021	Greenville	49272	589	507003
1/20/2021	Richland	29592	357	411357
1/20/2021	Charleston	25376	319	401165
1/20/2021	Spartanburg	24310	446	307617
1/20/2021	Horry	22096	307	332172
1/20/2021	Lexington	19815	304	290278
1/20/2021	York	18693	195	265872
1/20/2021	Anderson	14876	310	198064
1/20/2021	Pickens	13044	170	124029
1/20/2021	Florence	12735	298	138475
1/20/2021	Beaufort	11418	121	186095

COVID CASES FOR EACH COUNTY WITH FOLIUM MAP



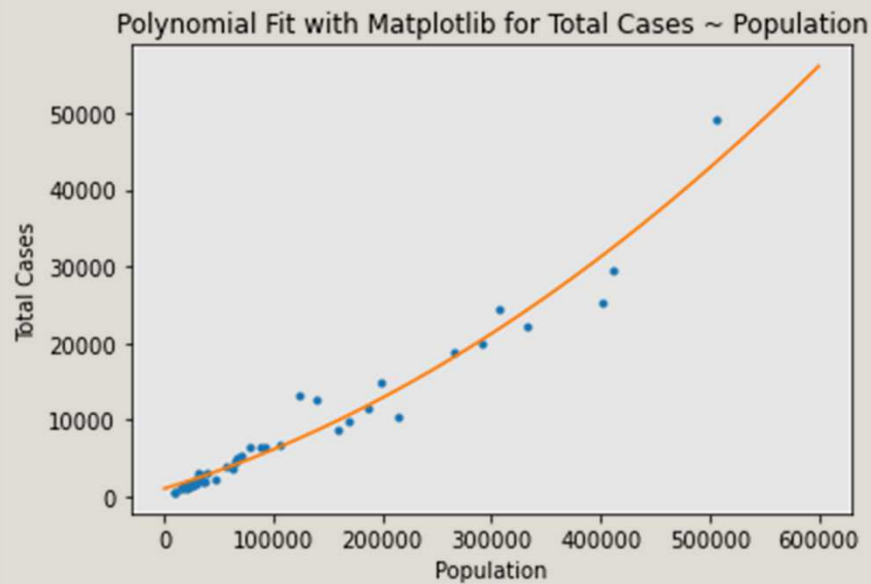
Greenville County is the “hot spot” for the South Carolina

LINEAR REGRESSION MODEL



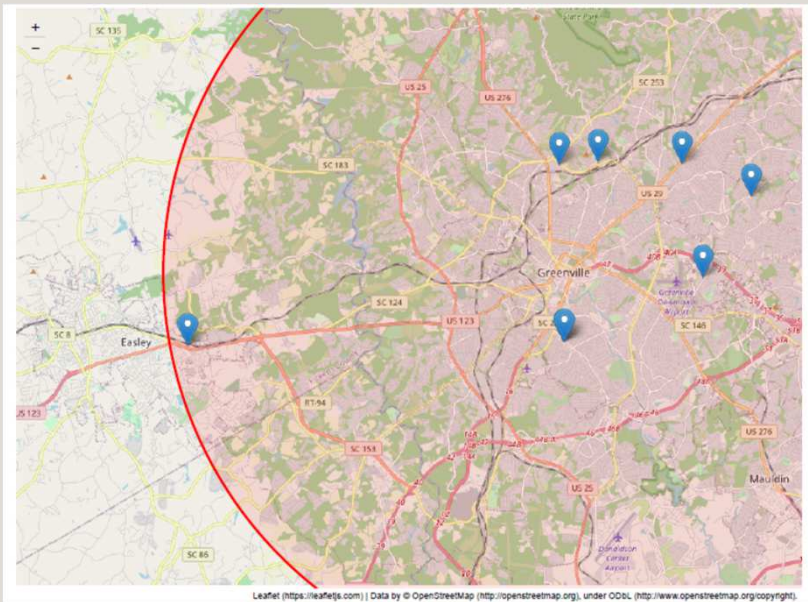
Linear regression model was further applied and the slope is round 0.1, which means that 10% of population were affected by the COVID-19 in the most of area within the South Carolina state.

POLYNOMIAL MODEL



Polynomial fitting is much better than the linear regression especially on the high population area

MAP OF URGENT CARE FOR COVID TEST



Foursquare API was used to search the Urgent care facility for the COVID testing. The location was mapped by folium image.

CONCLUSION

- We observed that the Greenville county is the Covid **hotspot** for the South Carolina County. There is a simple linear relation between the total cases versus the population. So far, we observed 10% infection rate in the state. Further data analysis, we observed that high population area has high infection rate compared to the lower population area. **Polynomial** fitting is better to show the hotspot area. Through **Foursquare API**, we can easily locate the covid test for the local community. This will be a very useful to show the visual info and give the suggestion or alarm to the local neighborhood.