

COVID-19 Analysis for South Carolina State

Xinfei Yu

Introduction¹

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

The data will be scraped from [worldometer](https://www.worldometer.com/) 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 data will be reviewed through choropleth map to find the hotspot. Exploratory data analysis and linear regression will be conducted to find the relationship between population and the total Covid cases

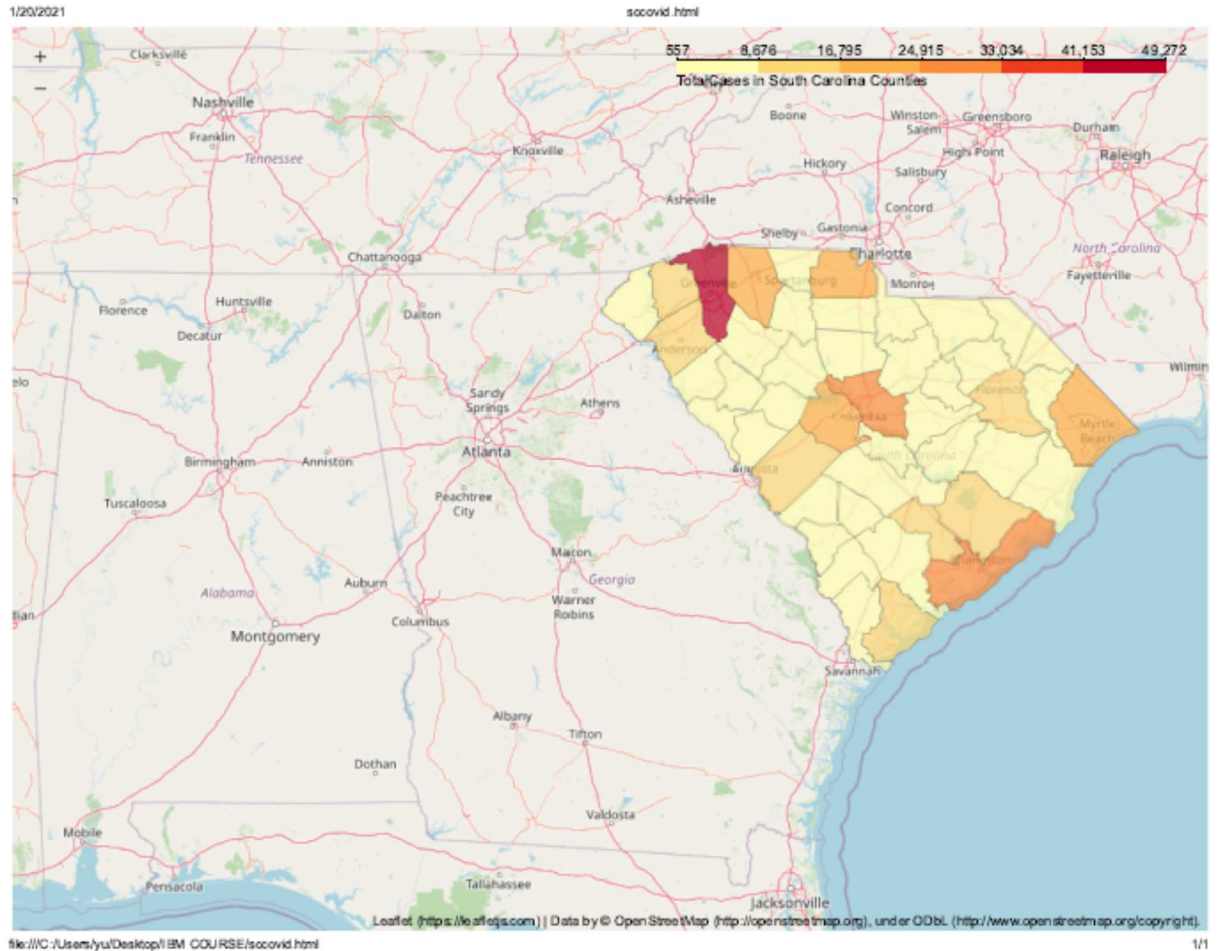
For the hotspot area, Urgent care or pharmacy store will be located through **Foursquare API** to give the recommendation for testing. Based on the longitude and latitude info of the hotspots, the hospitals and urgent care will be listed for selection.

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
1/20/2021	Berkeley	10235	113	215044
1/20/2021	Aiken	9701	136	168301
1/20/2021	Dorchester	8759	122	158299
1/20/2021	Sumter	6727	122	106757
1/20/2021	Oconee	6539	78	77528
1/20/2021	Orangeburg	6508	170	87687
1/20/2021	Lancaster	6408	95	92308
1/20/2021	Greenwood	5278	99	70411
1/20/2021	Darlington	4949	104	67027

1/20/2021	Laurens	4704	99	66846
1/20/2021	Kershaw	4456	74	65112
1/20/2021	Cherokee	3912	94	56895
1/20/2021	Georgetown	3666	74	61952
1/20/2021	Newberry	3094	70	38194
1/20/2021	Dillon	2997	60	30689
1/20/2021	Chester	2779	46	32311
1/20/2021	Williamsburg	2604	56	31324
1/20/2021	Marion	2486	74	31308
1/20/2021	Chesterfield	2357	67	45953
1/20/2021	Clarendon	2075	77	33957
1/20/2021	Colleton	2062	63	37585
1/20/2021	Marlboro	2010	30	26753
1/20/2021	Union	1813	49	27490
1/20/2021	Edgefield	1782	19	26927
1/20/2021	Barnwell	1739	33	21346
1/20/2021	Jasper	1653	28	28657
1/20/2021	Fairfield	1610	54	22565
1/20/2021	Abbeville	1413	21	24627
1/20/2021	Lee	1358	44	17365
1/20/2021	Hampton	1208	33	19564
1/20/2021	Saluda	1112	26	20303
1/20/2021	Bamberg	1101	43	14376
1/20/2021	Calhoun	957	23	14663
1/20/2021	Allendale	615	9	9024
1/20/2021	McCormick	557	8	9531

Results

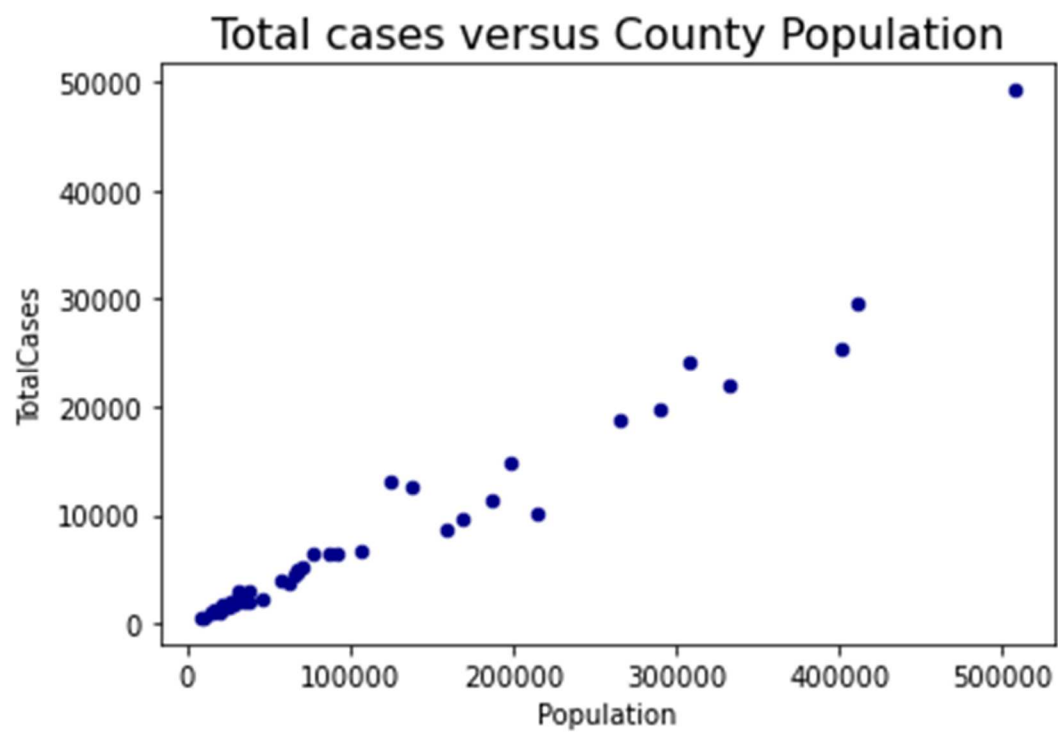
Exploratory data analysis was conducted through the folium map. The image below shows the total cases of COVID-19 on each county. Clearly, Greenville county has the most total cases so far, which is the hotspot for the south Carolina state.



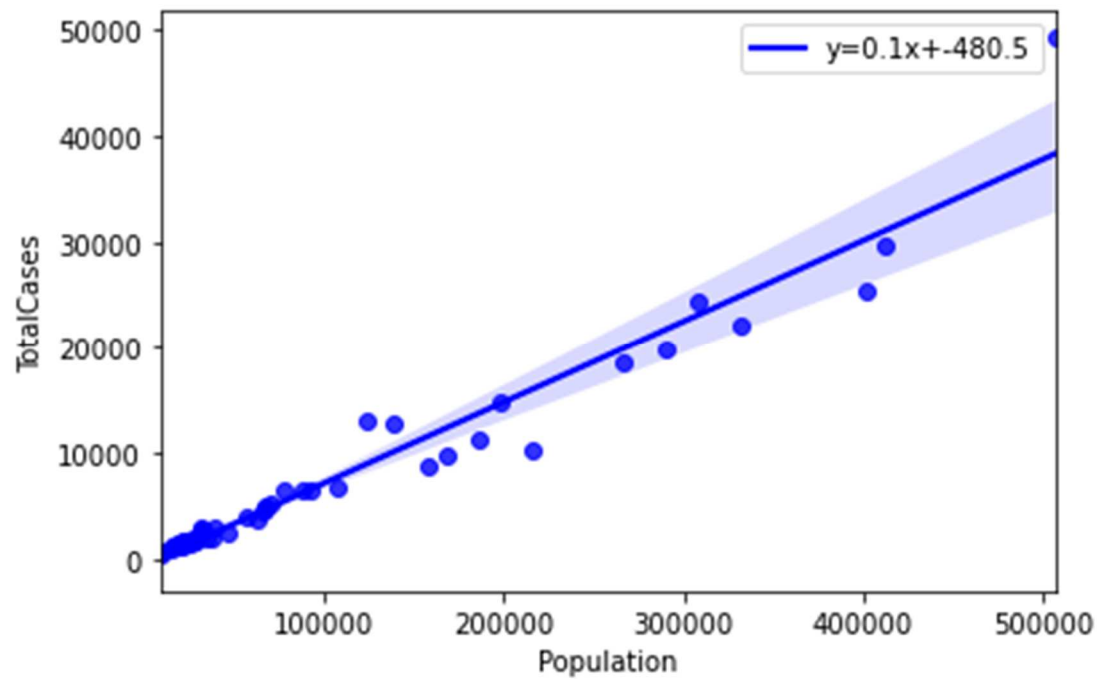
Discussion

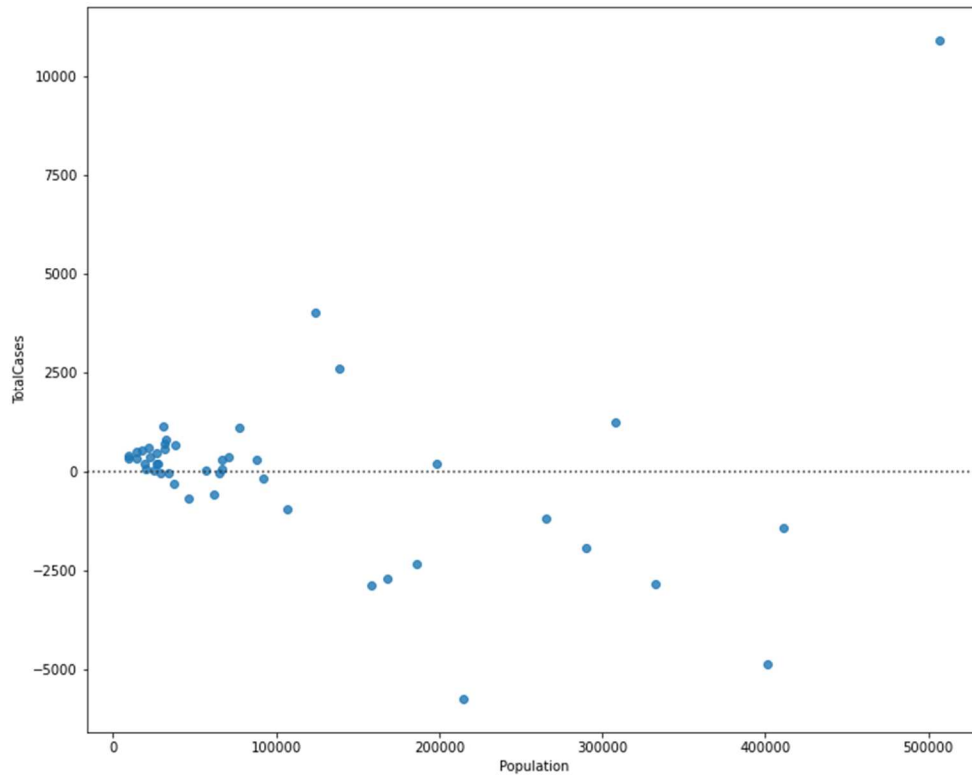
The scatter plot was used to study the relationship between total cases versus the populations of each county. The graph shows the linear relationship for the low population area. Linear regression model was further applied and the slope is round 0.1, which means that 10% of population were affected by the COVID-19 in the most of area within the South Carolina state.

From the linear regression plot, we also observed high residue for the high population area, which means that linear regression does not fit well for high population area. For the south Carolina state, Greenville county does not in that linear range.

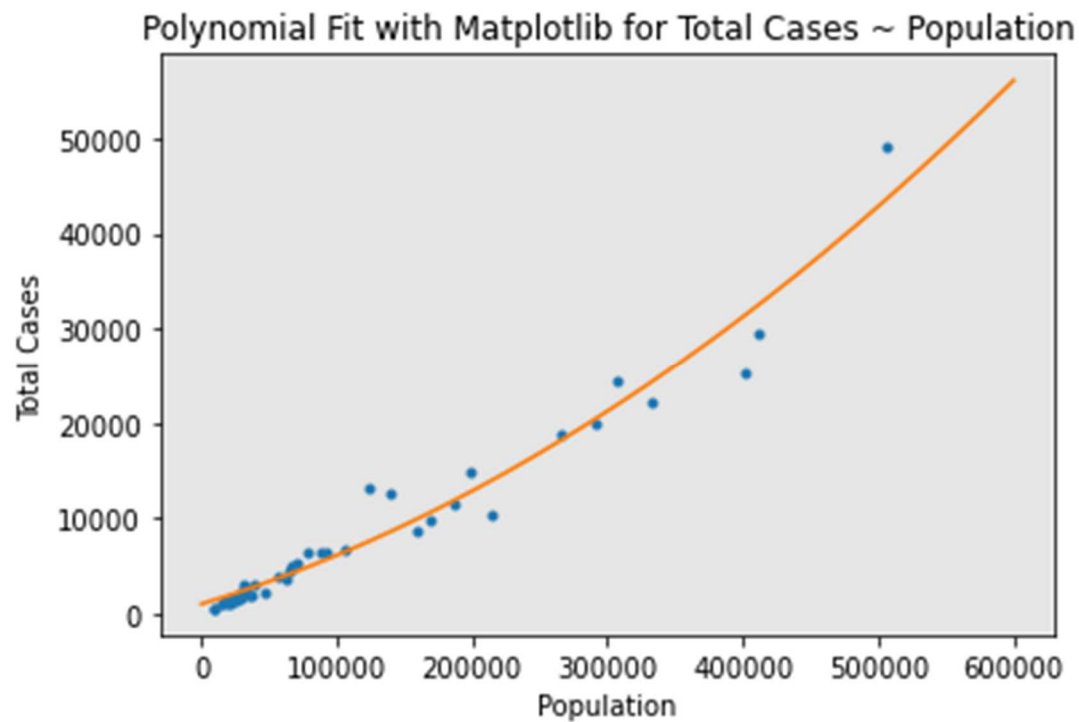


Data was fitted through linear Regression and display through graph below



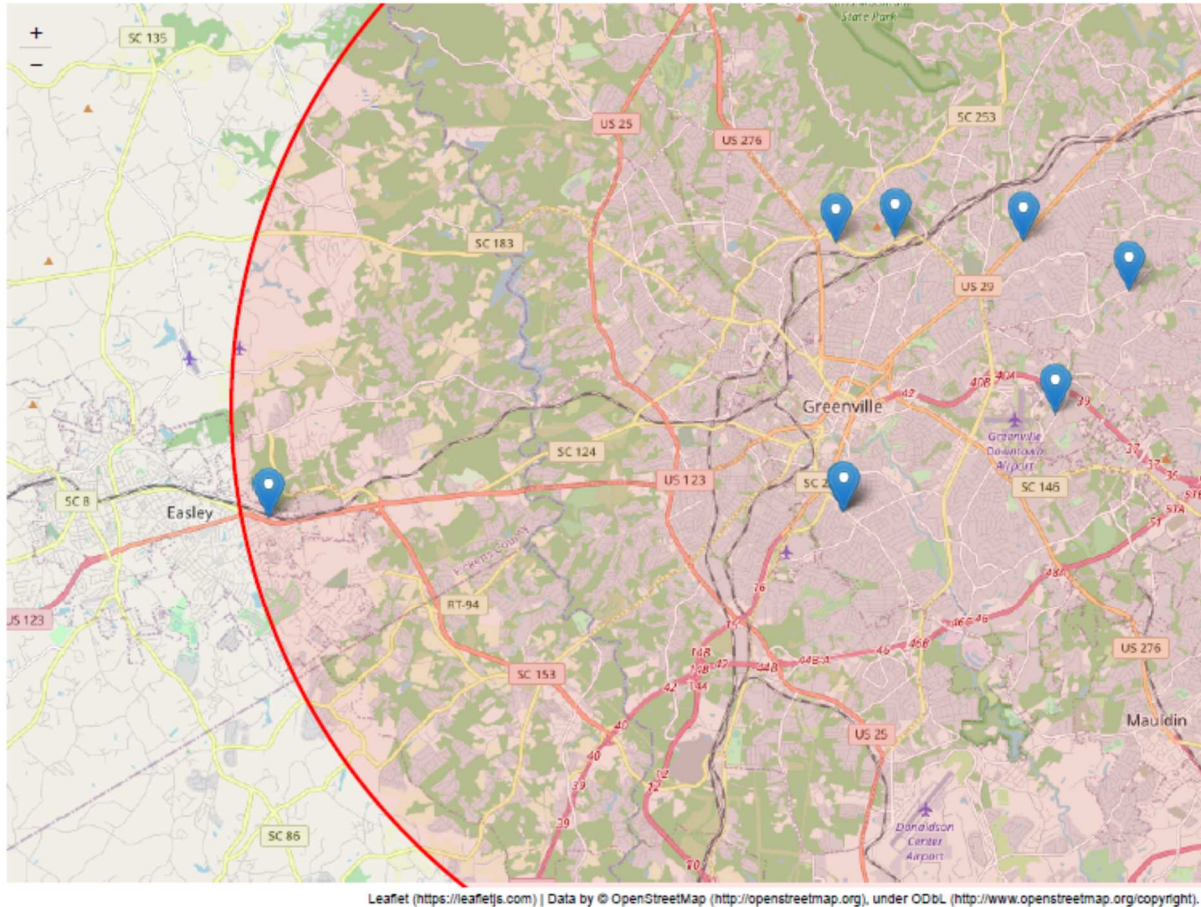


Polynomial fitting is much better than the linear regression especially on the high population area



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

name	categories	address	postalC ode	City	state
AFC Urgent Care Bon Secours â€” Augusta Road	Hospital	1800 Augusta St Ste B	29605	Greenvi lle	SC
AFC Urgent Care Bon Secours - Cherrydale	Hospital	3213 N. Pleasantburg Dr.	29609	Greenvi lle	SC
AFC Urgent Care Bon Secours â€” Wade Hampton	Hospital	2310 Wade Hampton Blvd	29615	Greenvi lle	SC
AFC Urgent Care	Urgent Care Center	1800 Augusta St Ste B	29605	Greenvi lle	SC
AFC Urgent Care Bon Secours â€” Pelham Road	Hospital	2709 Pelham Rd Ste B	29615	Greenvi lle	SC
AFC Urgent Care Woodruff Road	Hospital	1467 Woodruff Rd Ste C	29607	Greenvi lle	SC
AFC Urgent Care Bon Secours â€” Greer	Hospital	1494 W Wade Hampton Blvd Ste C	29650	Greer	SC
AFC Urgent Care Easley, SC	Urgent Care Center	5208 Calhoun Memorial Hwy	29640	Easley	SC
AFC Urgent Care Bon Secours â€” Simpsonville	Hospital	3930 Grandview Drive, Suite B	29680	Simpso nville	SC
AFC Urgent Care Bon Secours Haywood Road	Hospital	649A Haywood Rd	29607	Greenvi lle	SC
Advanced Urgent Care	Medical Center	2498 N Pleasantburg Dr	29609	Greenvi lle	SC
Urgent Care	Urgent Care Center		29615	Greenvi lle	SC
Doctors Express Urgent Care - Greer	Urgent Care Center	1494 W Wade Hampton Blvd Ste C	29650	Greer	SC



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.