Chicago COVID 19 Analysis

Mandakini Majumdar

May 31st, 2020

1. Introduction

1.1 Background

COVID 19 brings an unprecedented challenge for our society and locality. Being a Chicago resident, I want to understand impact of COVID 19 outbreak for Chicago localities. Furthermore COVID 19 forced many business and restaurants to be closed for dine-in. People are now relying on delivery and takeout offerings.

1.2 Problem

Some restaurant owners have already shut down operations due to safety risk. Still, some restaurants continue to serve to give Chicagoans some semblance of normalcy as they face economic uncertainties. I used Foursquare location data to find out restaurants those are still open for delivery and takeout offerings.

1.3 Audience and Interest

Audience of this project is people of Chicago city or people who have to come Chicago for some urgency. This project will guide people to understand COVID 19 impacts in different zip codes. Also this project will help them to find out where to order food takeout or delivery for their respective zip code.

2. Data Acquisition and Cleaning

2.1 Data Sources

I have collected COVID 19 data for Chicago from 'healthdata.gov'. This site dedicated to making high value health data more accessible to entrepreneurs, researchers, and policy makers in the hopes of better health outcomes for all. COVID 19 data set provides weekly count of tests, positive cases and deaths for a zip code due to COVID 19 pandemic.

I used latitude and longitude data from 'public.opendatasoft.com'. Opendatasoft is a French company that offers data sharing software and platform. Based in Paris, it also has a subsidiary in Boston (United States) and an office in Nantes (France). Zip data set from Opendatasoft provides latitude, longitude, county, city and state details of a zip code.

I used Foursquare API to get restaurant recommendations those are still open for corresponding ZIP code. I also retrieve category of restaurant to guide my audience for their food selection.

2.2 Data Cleaning and Processing

Data has been downloaded from websites using read CSV functions and stored in data frames. There were COVID 19 data with unknown zip code. Those data had been removed from the data frame. Chicago's COVID 19 data was merged with Zip Code latitude and longitude data to get geolocation of affected areas. Data columns were modified to remove black spaces from column names.

	Zip	City	State	Latitude	Longitude	Timezone	Daylight savings time flag	geopoint	Zip_Code	Cases	Tests	Deaths
0	60651	Chicago	IL	41.901485	-87.74055	-6	1	41.901485,-87.74055	60651	1463.0	4625	59
1	60644	Chicago	IL	41.881331	-87.75671	-6	1	41.881331,-87.75671	60644	1019.0	3776	66
2	60646	Chicago	IL	41.995331	-87.76010	-6	1	41.995331,-87.7601	60646	335.0	1725	34
3	60654	Chicago	IL	41.888627	-87.63538	-6	1	41.888627,-87.63538	60654	132.0	766	1
4	60616	Chicago	IL	41.847400	-87.63126	-6	1	41.8474,-87.63126	60616	385.0	2512	34

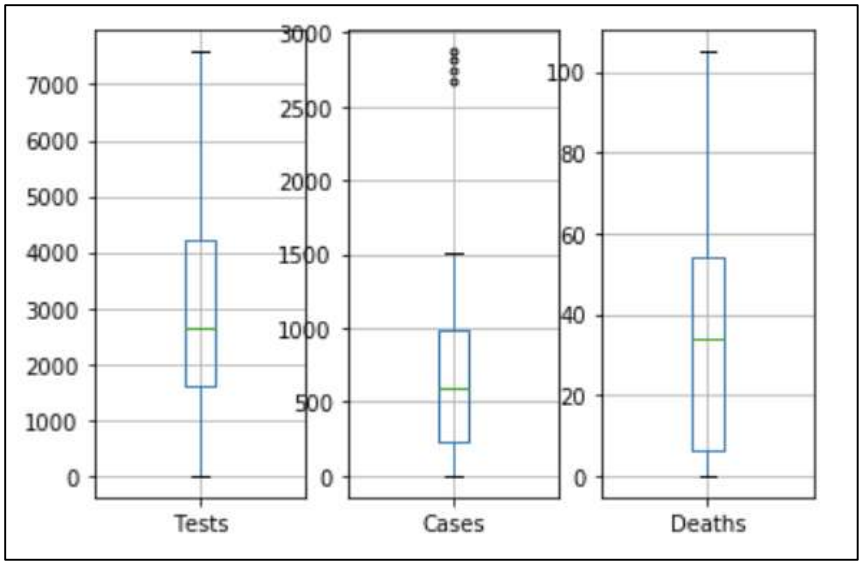
3. Exploratory Data Analysis

3.1 Descriptive statistics of data

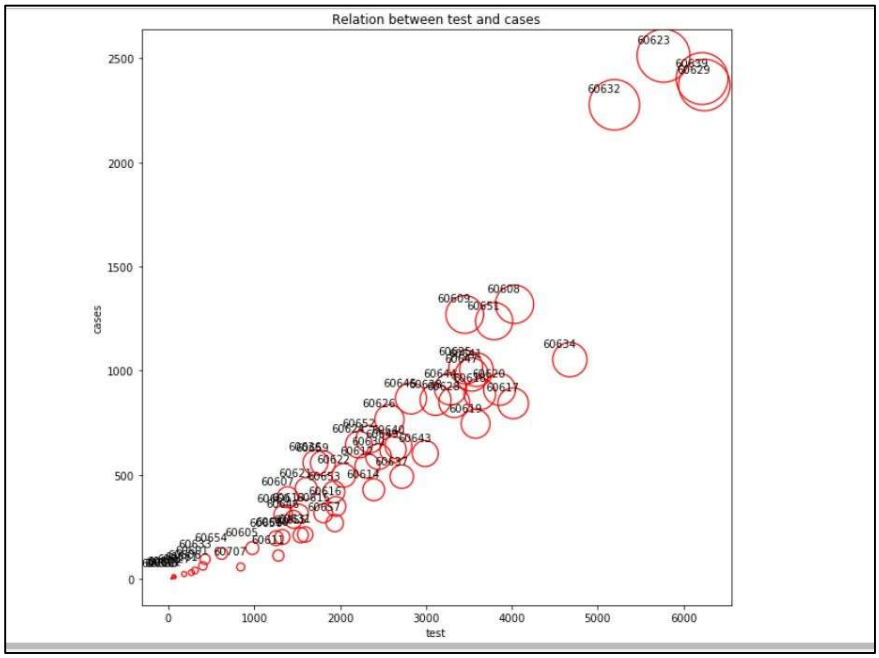
Descriptive statistics of data was obtained as follows.

Measures	Values
count	58.000000
Mean	724.534483
Standard Deviation	699.431705
Min	0.000000
25%	232.250000
50%	592.000000
75%	981.250000
Max	2871.000000

Based on percentile spread of data, all the affected Zip Codes were classified in low, medium, high and very high intensity. A label column was added to mark the intensity of Zip Code. Boxplot was plotted to analyze how the values in our data are spread out for COVID 19 tests, positive cases and deaths



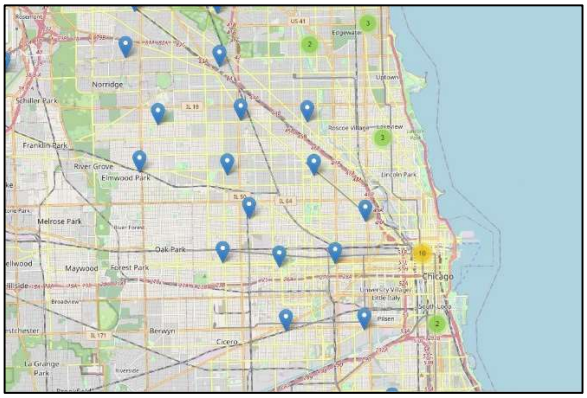
Then plot a scatter graph was plotted to establish relationship between test and cases where circle size represents the number of positive cases.



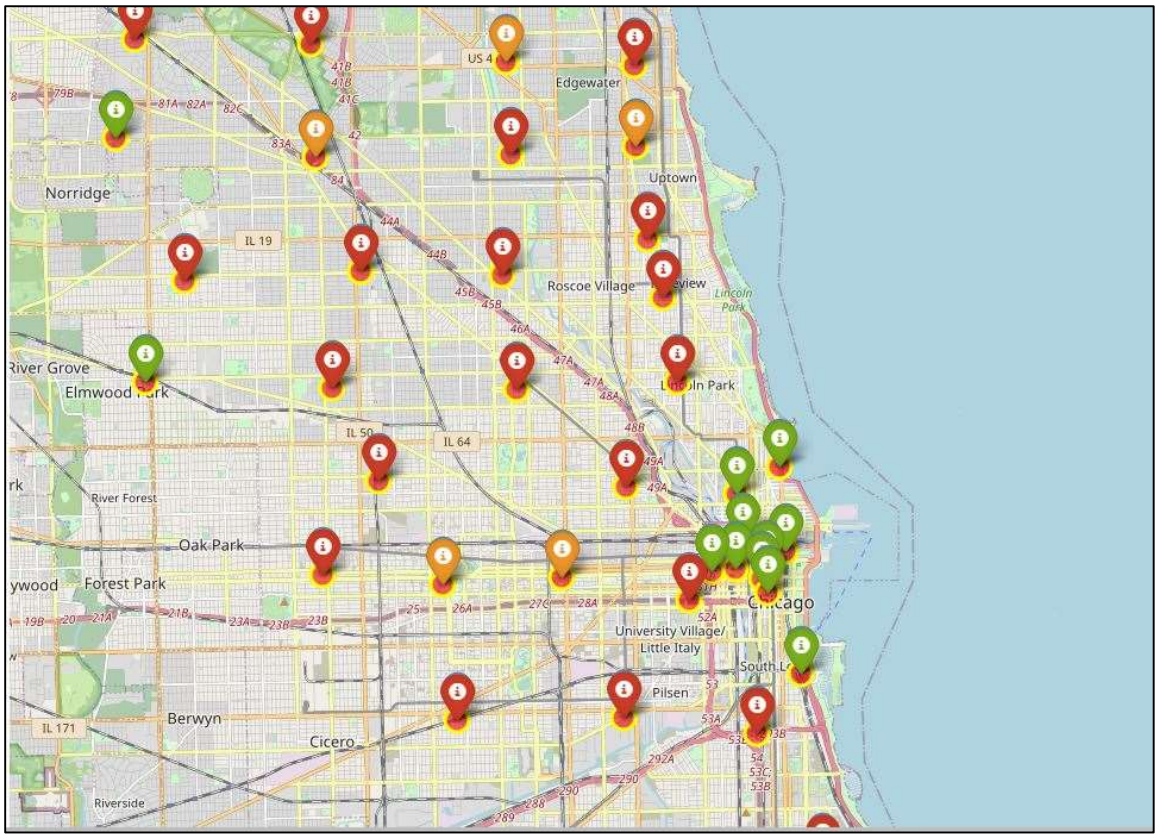
The plot clearly show increasing pattern of positive cases with the increment of testing. Also there are some highly affected zip codes such as 60623, 60629, 60632 and 60639. People can avoid visiting those Zip codes if it is not absolutely required. Also people from those Zip codes should be extra careful during going out.

3.2 Plotting affected zip codes on Chicago map

Subsequently affected Zip Codes were plotted on Chicago map using Folium library. This map clearly shows high density of positive COVID 19 cases around Chicago city. There are clusters in city areas.

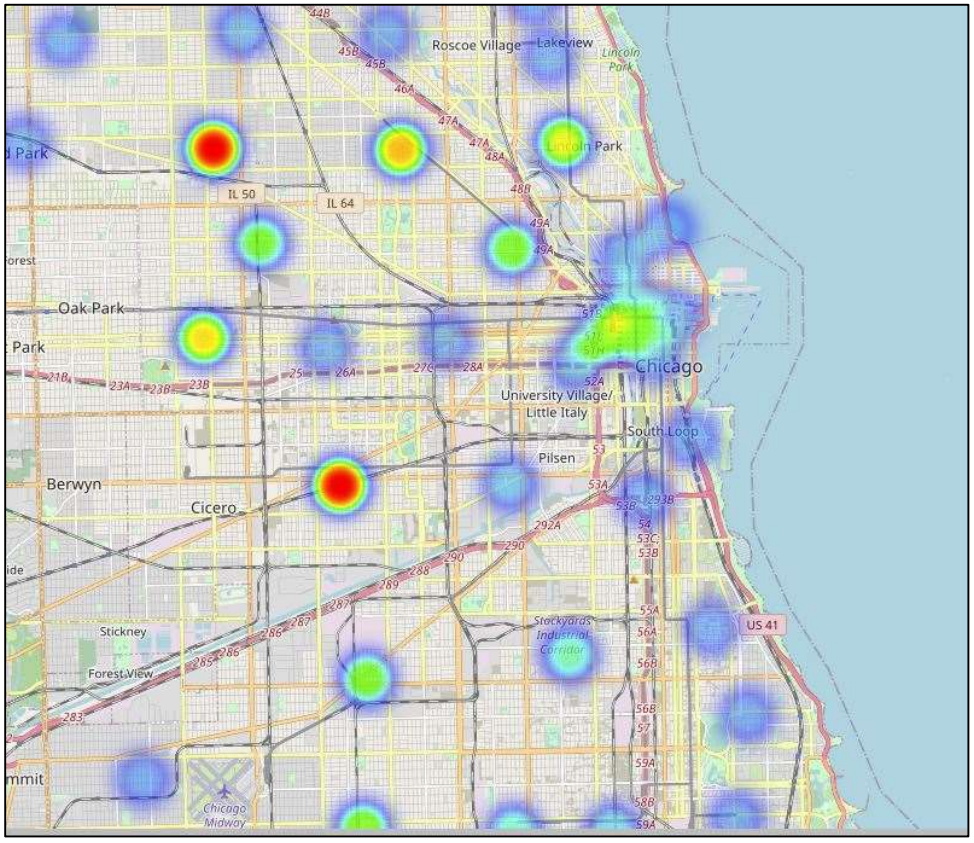


Then intensity of affected Zip Codes were plotted with appropriate color (Low intensity = Yellow, Medium intensity = Yellow, High intensity = Orange and Very high intensity = Red)



From the intensity plotting it can be observed that most of the Zip Codes in Chicago are now under highly or very highly affected zones

Similarly, heatmap has been plotted to analyze COVID 19 deaths in Chicago region. Few places outside Chicago city has significantly higher death rates compare to other Zip Codes.

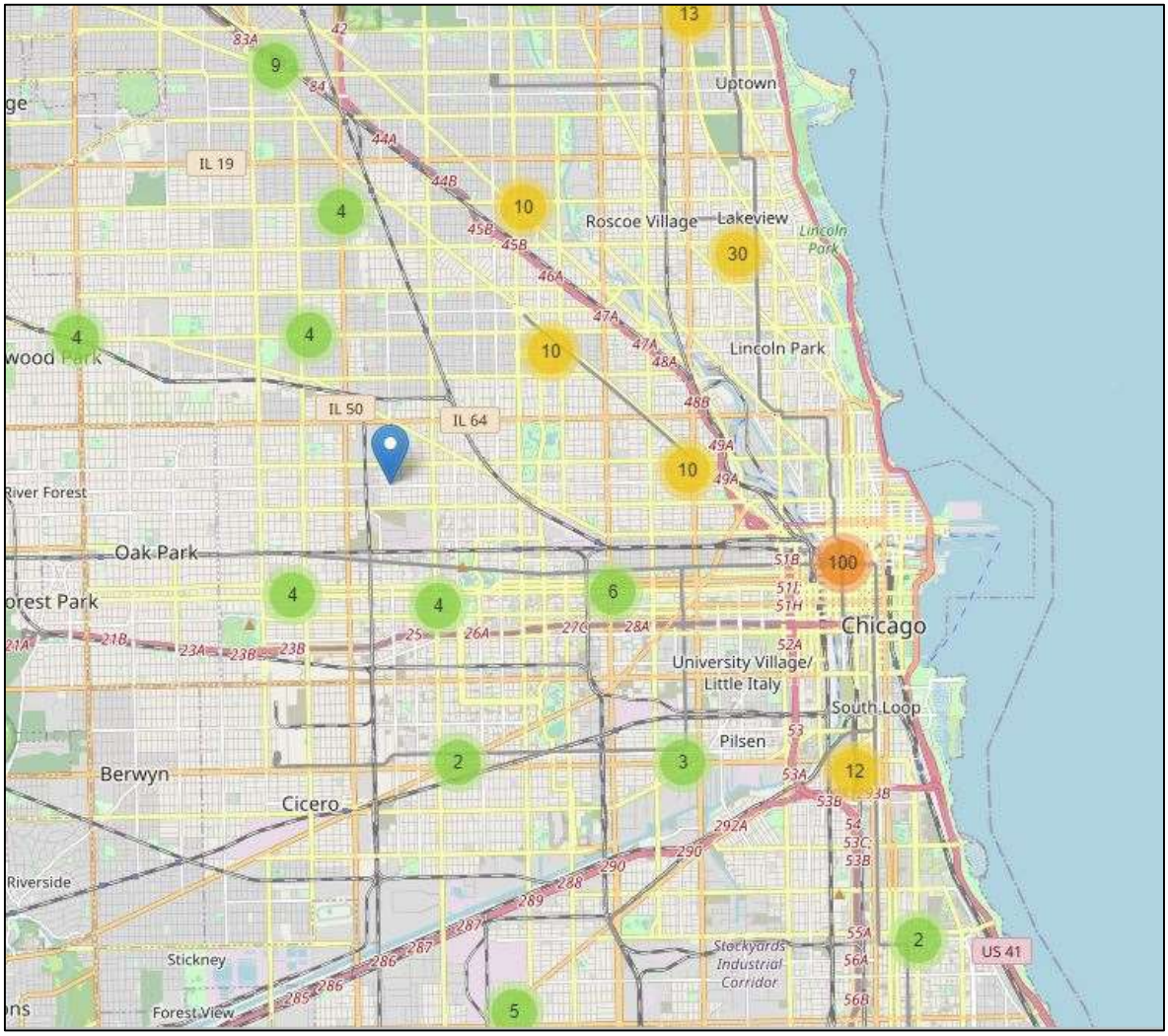


4. Use Foursquare API to explore open restaurants

As described in background section of this report, I am trying to guide Chicagoans where to order to get their food under this locked down situation. Foursquare API is used to obtain all the open restaurants those are still operating for take out or delivery. Once I got the restaurants, I used categorizations to reflect type of restaurants or cuisine.

	zip_code	Zip Latitude	Zip Longitude	Venue	Venue_Latitude	Venue_Longitude	Venue_Category
0	60651	41.901485	-87.74055	Wingstop	41.910045	-87.737101	Wings Joint
1	60651	41.901485	-87.74055	Taco Bell	41.909116	-87.739787	Mexican Restaurant
2	60644	41.881331	-87.75671	Seafood Junction	41.880618	-87.757804	Seafood Restaurant
3	60644	41.881331	-87.75671	MacArthur's Restaurant	41.880611	-87.760757	Southern / Soul Food Restaurant
4	60644	41.881331	-87.75671	KFC	41.880592	-87.756757	Fried Chicken Joint
5	60644	41.881331	-87.75671	McDonald's	41.880179	-87.750711	Fast Food Restaurant
6	60644	41.881331	-87.75671	Red Snapper Madison	41.879926	-87.757363	Fried Chicken Joint

Then the open restaurants were plotted on Chicago map to guide the audience visually about their available options to order



5. Conclusions and Future Direction

In this study, I analyzed how COVID 19 pandemic out broke in Chicago area. This analysis can increase awareness of this deadly pandemic among Chicagoans. Also, this analysis will help them to find out food options in this locked down situation. Similar analysis can be done for other cities to guide the citizens.