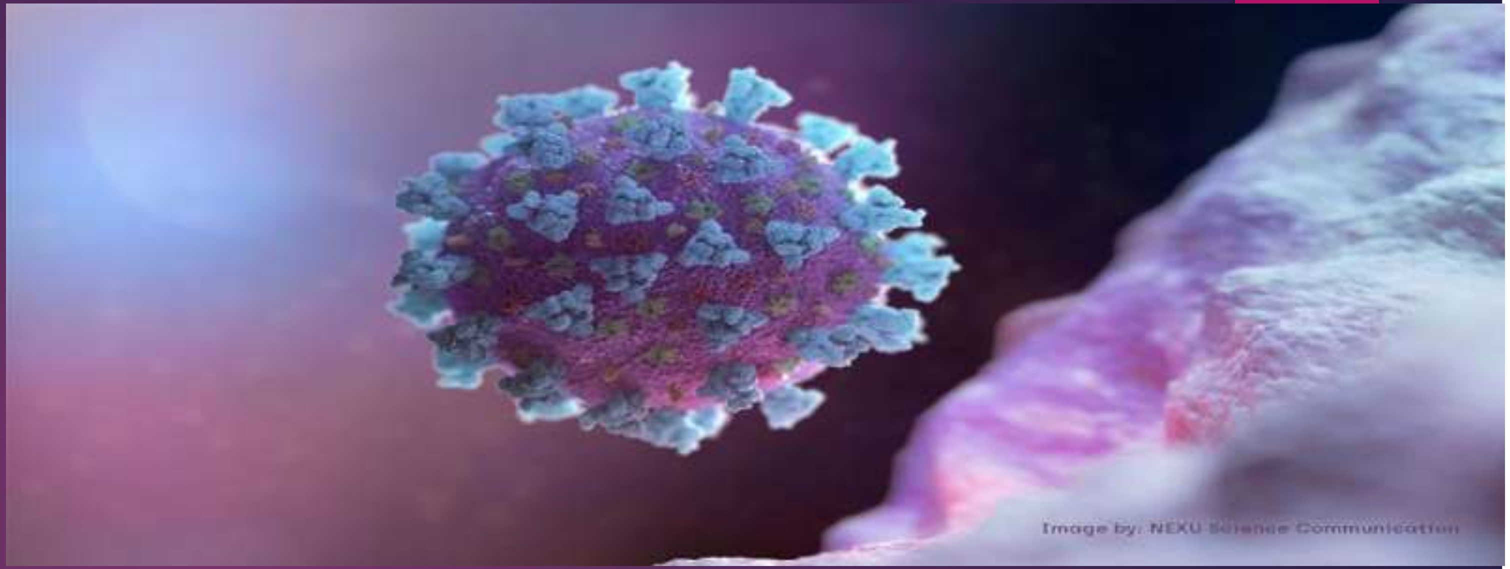


Capstone Project - The Battle of Neighborhoods (Week 2)

CORONAVIRUS GLOBAL PANDEMIC SPREAD IN THE CITY OF LONDON

Introduction

- **Background:** In current scenario of global pandemic, the coronavirus spread in more than 180 countries around the world. Its place of origin, cause of spread and after effects.
- **Problem:** To control the spread of coronavirus in one of the worst effected regions of any country like here we are taking London, UK.
- **Interest:** Government, Doctors and Health Care workers who are real audience at this time of crisis. This project will provide location of the hospitals and near testing centres for COVID-19 patients(effected by coronavirus).

Data acquisition and cleaning

- **Data sources**

The data acquired for the project is the combination data from three sources.

- The first source of the project uses London Coronavirus Infected cases in London.
- The second source of data is scraped from the Wikipedia page that contains the list of London Boroughs. The page contains additional information about the borough.
- The third data source is the list of Neighbourhoods in the Royal Borough of Brent as found on a Wikipedia page. This dataset is created from scratch using the list of neighbourhoods available on the site.

Data acquisition and cleaning

- **Data Cleaning**

The data preparation for each of the three sources of data is done separately.

- Initially, we took the data from showing total number of confirmed infections in each borough.
- The second data is scraped from a Wikipedia page using the Beautiful Soup library in python and get extracted in the tabular format.
- The two datasets are merged on the Borough names to form a new dataset that combines the necessary information in one dataset.
- After visualizing the coronavirus cases in each borough, we can find the worst effected borough. The third source of data is acquired from the list of neighbourhoods in the worst affected borough on Wikipedia.

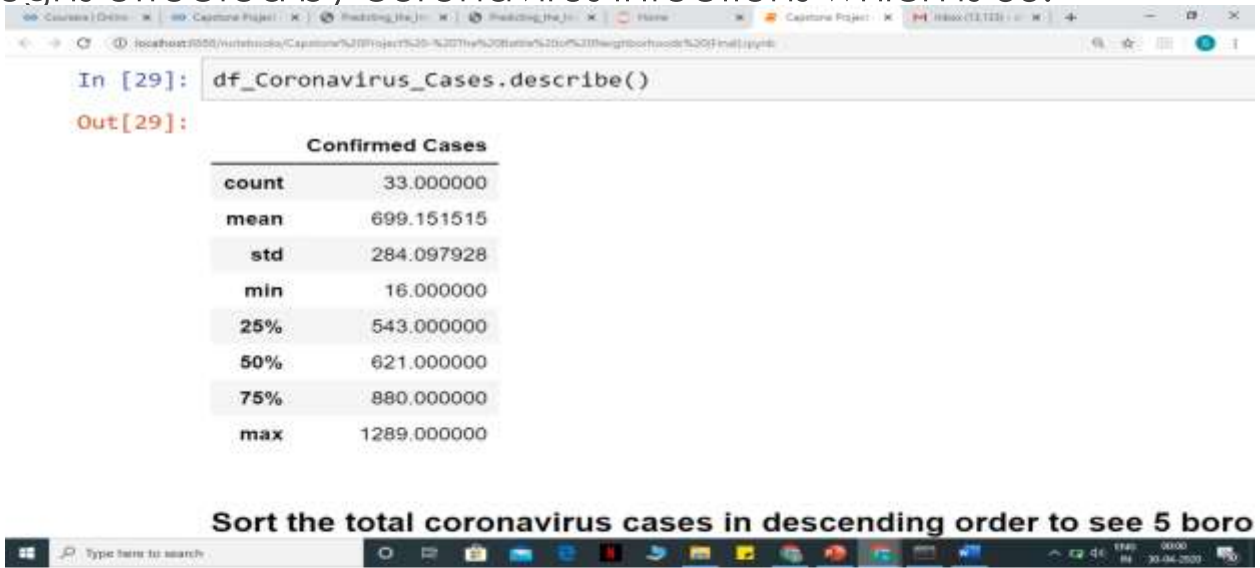
Relevant Solution

- One way is to analyse the worst regions of London and practice strict Lockdown there. Along with that the regions have to be critically monitored for the spread.
- using the **Foursquare API** we can help locate common venues with hospitals in the neighbourhoods. Since hospitals and other healthcare facilities seems to be very less in the neighbourhoods of Brent which is worst effected, therefore plan should be to convert certain places like malls and restaurants which are closed for now to quarantine places or temporary hospitals where large number of patients can be taken care.

Methodology

- **Exploratory Data Analysis**
- **Statistical Summary of Data**

We can get the statistics of London Coronavirus data by using the describe function. The function returns mean, standard deviation, minimum, maximum, 1st quartile(25%), 2nd quartile(50%) and third quartile(75%) and total number of boroughs effected by coronavirus infections which is 33.



In [29]:
```
df_Coronavirus_Cases.describe()
```

Out[29]:

	Confirmed Cases
count	33.000000
mean	699.151515
std	284.097928
min	16.000000
25%	543.000000
50%	621.000000
75%	880.000000
max	1289.000000

Sort the total coronavirus cases in descending order to see 5 boro

Methodology

- **Boroughs with maximum number of Coronavirus infected cases**

  Figure out the top five 5 boroughs with maximum number of coronavirus positive cases.

Methodology

- **Boroughs with minimum number of Coronavirus infected cases**

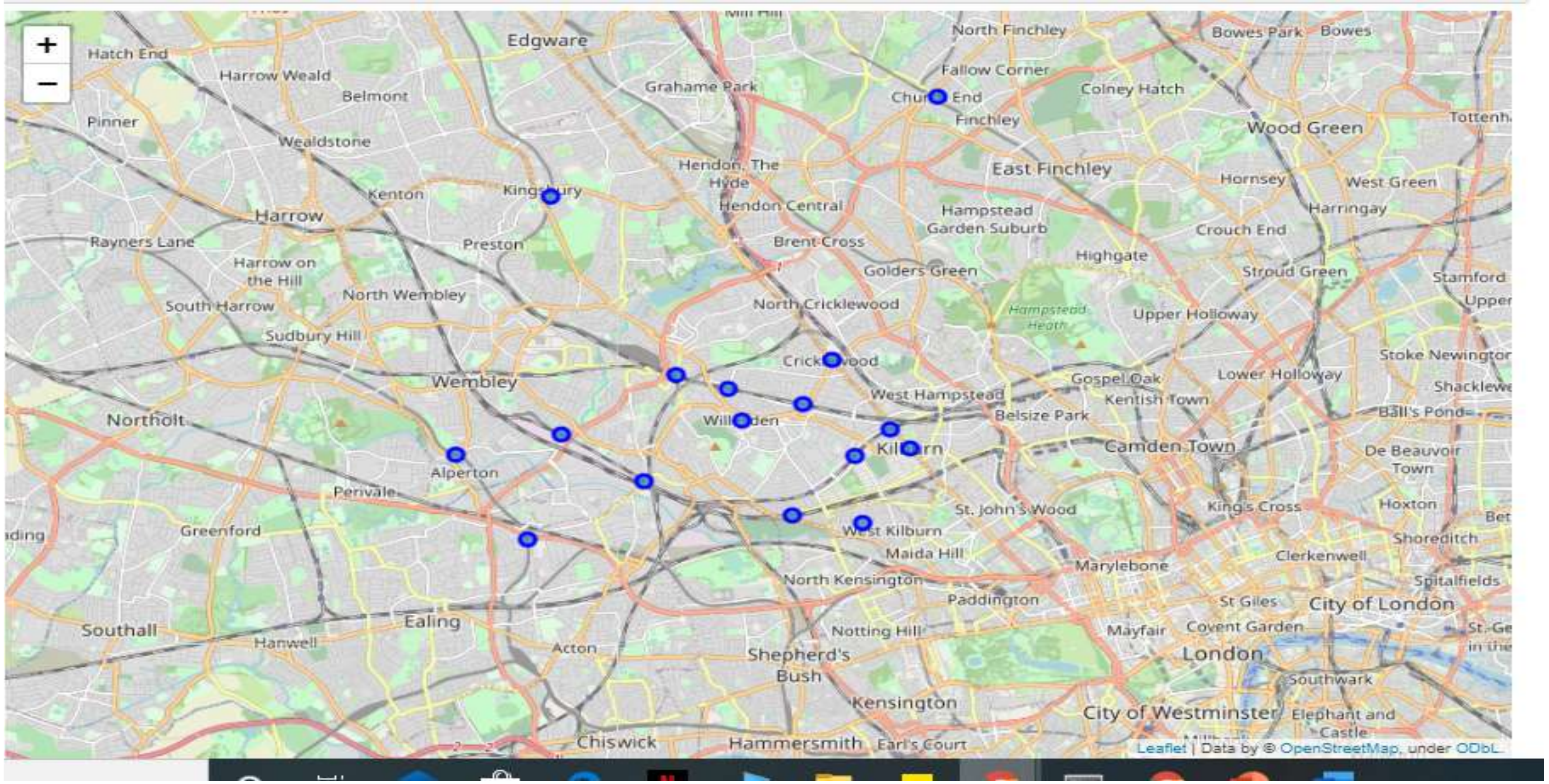
  Figure out the top five 5 boroughs with lowest number of coronavirus positive cases

Methodology

- After looking at the data we found that Brent is the place with the maximum number of coronavirus cases and this is the place where the strict government guidelines to be followed and measures to be taken to control the current spread of the virus.

Methodology

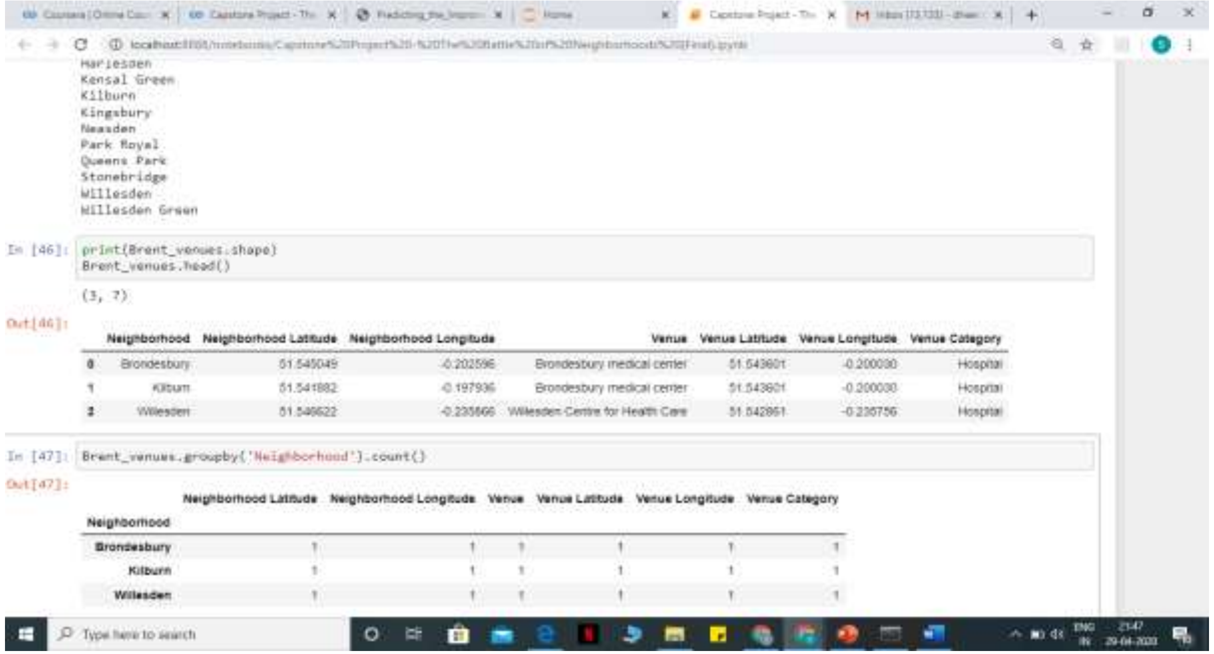
- Visualize the neighbourhood of Brent Borough



Methodology

- **Modelling**

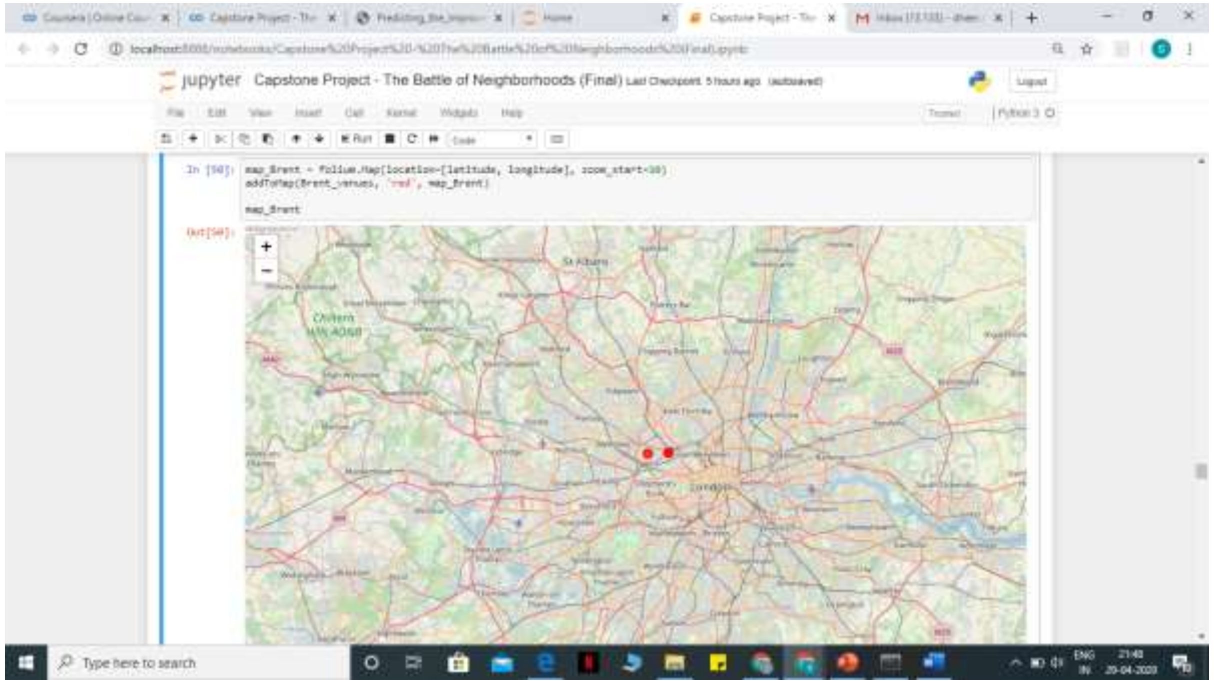
  Using the final dataset containing the neighbourhoods in Brent along with latitudes and longitudes, we can get the hospitals located within 500 metre radius using Foursquare API.



Methodology

- **Modelling**

  Visualizing the hospitals in the neighbourhood.

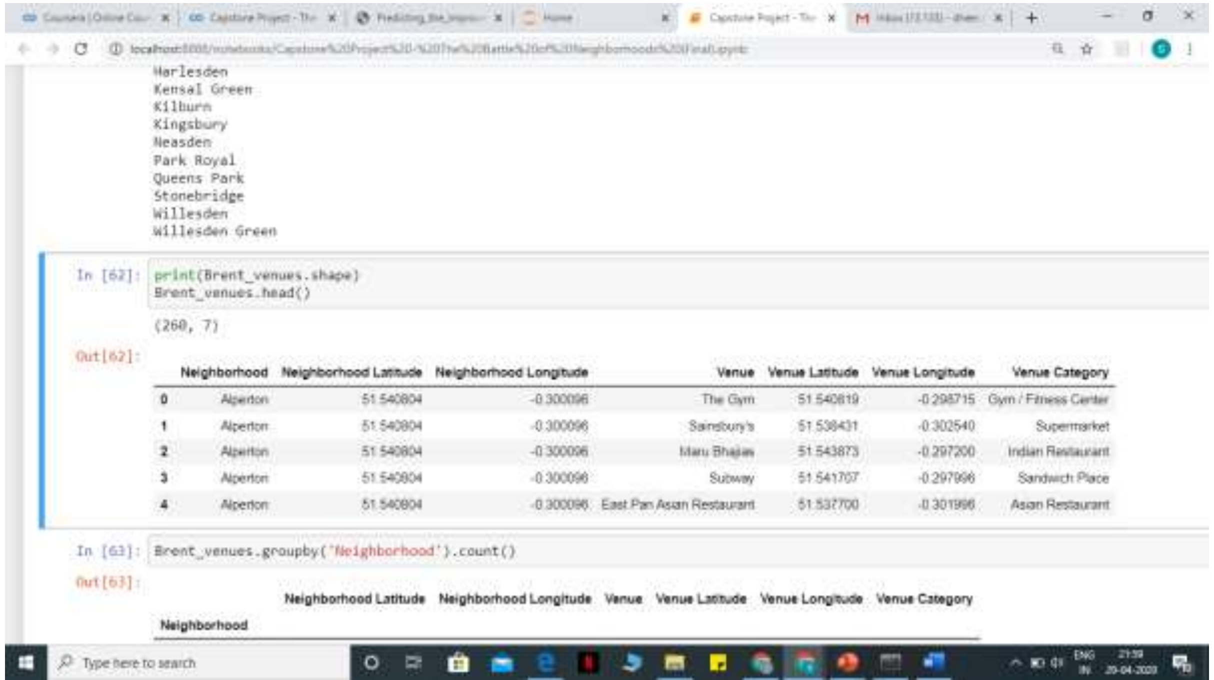


- **Modelling**

- One hot encoding on hospital data. The venue data is later grouped by neighbourhood and means of the hospitals are calculated.
- Clustering is done using k-means clustering i.e. unsupervised ML algorithm that cluster data based predefined cluster size. We are choosing 2 for hospitals since not many hospitals are not there located in the map.

- **Modelling**

  We perform the similar steps to figure out other venues which are voluntarily contributing converting themselves to health care centres and quarantine facilities for which we are choosing cluster size to 5.



```
Harlesden
Kensal Green
Kilburn
Kingsbury
Neasden
Park Royal
Queens Park
Stonebridge
Willesden
Willesden Green
```

In [62]:
```
print(Brent_venues.shape)
Brent_venues.head()
```

(260, 7)

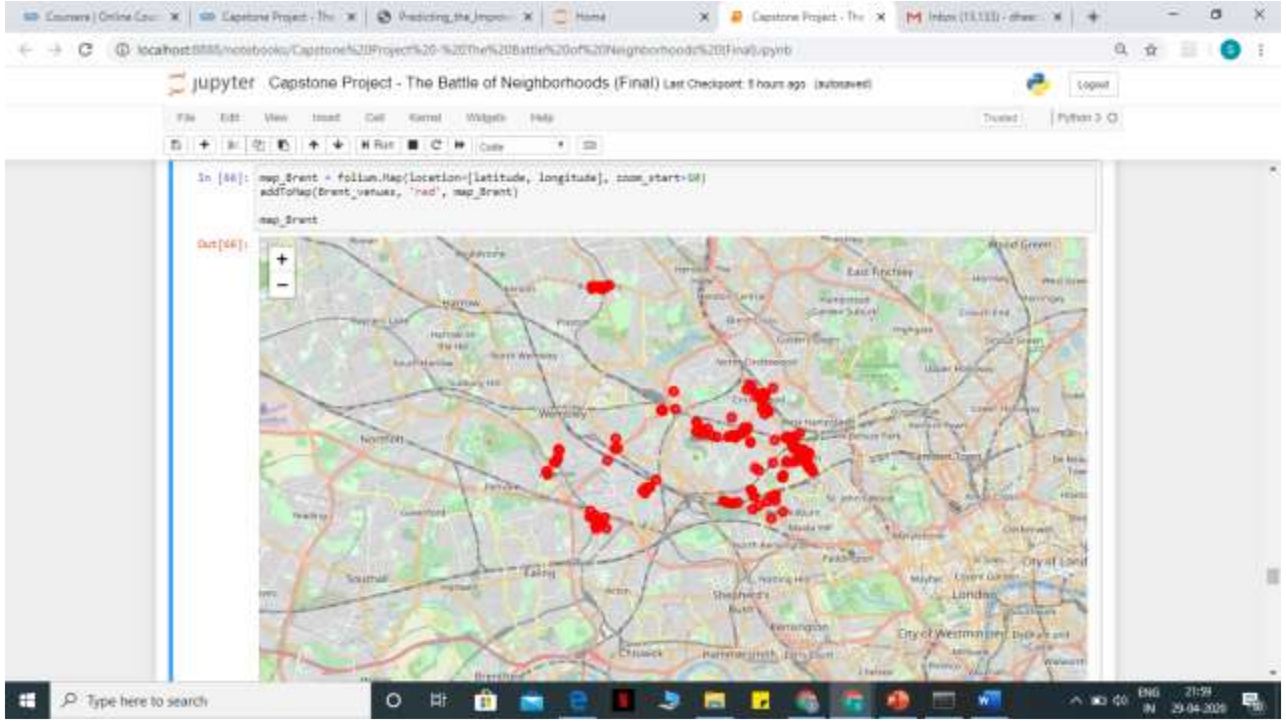
Out[62]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Alperton	51.540804	-0.300096	The Gym	51.540819	-0.298715	Gym / Fitness Center
1	Alperton	51.540804	-0.300096	Sainsbury's	51.538431	-0.302540	Supermarket
2	Alperton	51.540804	-0.300096	Maru Bhajias	51.543873	-0.297200	Indian Restaurant
3	Alperton	51.540804	-0.300096	Subway	51.541707	-0.297996	Sandwich Place
4	Alperton	51.540804	-0.300096	East Pan Asian Restaurant	51.537700	-0.301996	Asian Restaurant

In [63]:
```
Brent_venues.groupby('Neighborhood').count()
```

Out[63]:

	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighborhood						


jupyter Capstone Project - The Battle of Neighborhoods (Final)
map_Brent = folium.Map(location=[latitude, longitude], zoom_start=10)
addToMap(Brent_venues, 'red', map_Brent)
map_Brent
East Finchley
Wembley
Ealing
London
City of Westminster
Type here to search

Results

- After running the k-means clustering we can access each cluster to see neighbourhoods with hospitals and other venues belonging to each cluster.

1st Cluster of neighborhoods with hospitals.

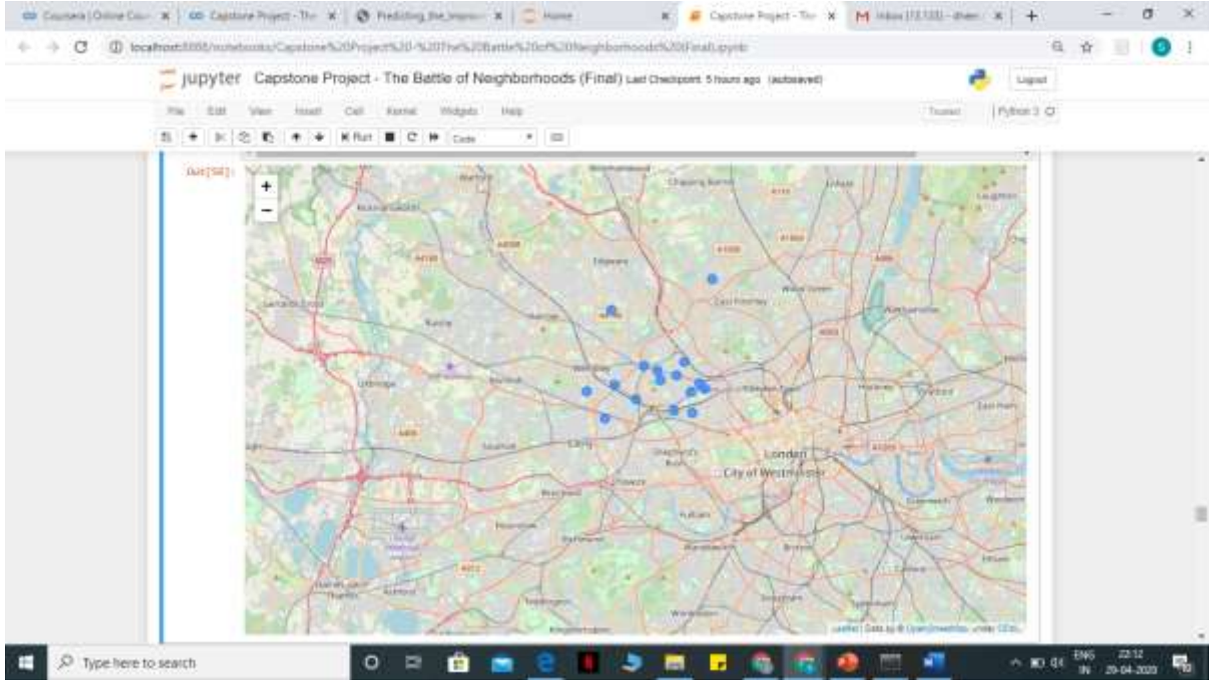
```
Brent_merged[Brent_merged['Cluster Labels'] == 0]
```

	Neighborhood	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
1	Brondesbury	Brent	51.545049	-0.202596	0.0	Hospital	Hospital	Hospital	Hospital	Hospital	Hospital	Hospital	Hospital	Hospital
8	Kilburn	Brent	51.541882	-0.197936	0.0	Hospital	Hospital	Hospital	Hospital	Hospital	Hospital	Hospital	Hospital	Hospital
14	Willesden	Brent	51.546622	-0.235866	0.0	Hospital	Hospital	Hospital	Hospital	Hospital	Hospital	Hospital	Hospital	Hospital

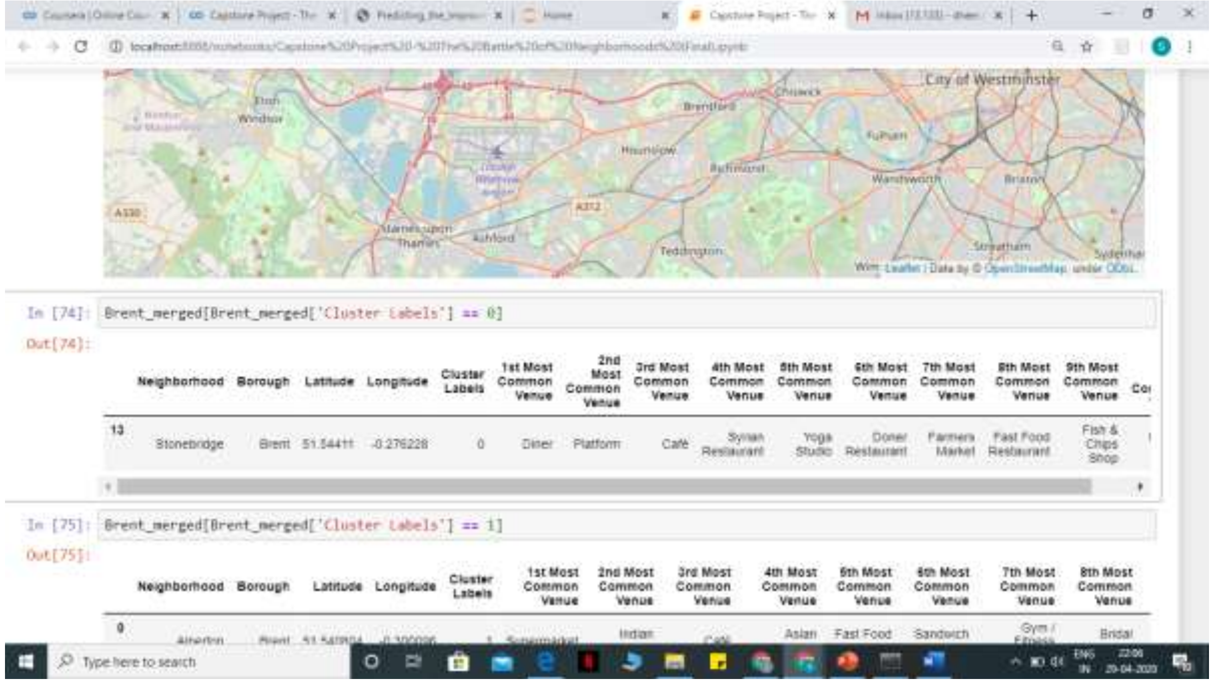
```
Brent_merged[Brent_merged['Cluster Labels'] == 1]
```

Neighborhood	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	Co

Visualizing the clustered neighbourhoods on a map using folium library



First Cluster with other Venus



- Second Cluster with other venues, we will choose this cluster due to maximum number of venues available and also neighbourhoods with hospitals will be situated nearer i.e in Brondesbury and Kilburn which will help the doctors to traverse from hospitals to the health care facilities.

```
Brent_merged[Brent_merged['Cluster Labels'] == 1]
```

	Neighborhood	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue
0	Alperton	Brent	51.540804	-0.300096	1	Supermarket	Indian Restaurant	Café	Asian Restaurant	Fast Food Restaurant	Sandwich Place	Gym / Fitness Center
1	Brondesbury	Brent	51.545049	-0.202596	1	Coffee Shop	Pub	Café	Bus Stop	Grocery Store	Middle Eastern Restaurant	Pizza Place
2	Brondesbury Park	Brent	51.540514	-0.210336	1	Park	Pub	Japanese Restaurant	Gym / Fitness Center	Train Station	Coffee Shop	Grocery Store
3	Church End	Brent	51.601117	-0.191890	1	Supermarket	Indian Restaurant	Café	Turkish Restaurant	Coffee Shop	Pub	Japanese Restaurant
4	Cricklewood	Brent	51.556699	-0.215751	1	Coffee Shop	Bus Stop	Hotel	Grocery Store	Fast Food Restaurant	Gym / Fitness Center	Pub
7	Kensal Green	Brent	51.530606	-0.224445	1	Pub	Wine Shop	Portuguese Restaurant	Indian Restaurant	Train Station	Bakery	Restaurant
8	Kilburn	Brent	51.541882	-0.197936	1	Café	Indian Restaurant	Pub	Supermarket	Brazilian Restaurant	Italian Restaurant	Discount Store
9	Kingsbury	Brent	51.584317	-0.278714	1	Grocery Store	Pub	Fruit & Vegetable Store	Coffee Shop	Sandwich Place	Snack Place	Supermarket

- Third Cluster with Venues

9	Kingsbury	Brent	51.584317	-0.278714	1	Grocery Store	Pub	Fruit & Vegetable Store	Coffee Shop	Sandwich Place	Snack Place	Supermarket	Portuguese Restaurant
10	Neasden	Brent	51.554336	-0.250749	1	Portuguese Restaurant	Sandwich Place	Train Station	Supermarket	Bus Stop	Fish & Chips Shop	Diner	Discount Store
11	Park Royal	Brent	51.526434	-0.283935	1	Hotel	Movie Theater	Gym / Fitness Center	Fast Food Restaurant	Pizza Place	Portuguese Restaurant	Hookah Bar	Sandwich Place
15	Willesden Green	Brent	51.549352	-0.222223	1	Coffee Shop	Italian Restaurant	Burger Joint	Pub	Pizza Place	Café	Kebab Restaurant	Metro Station

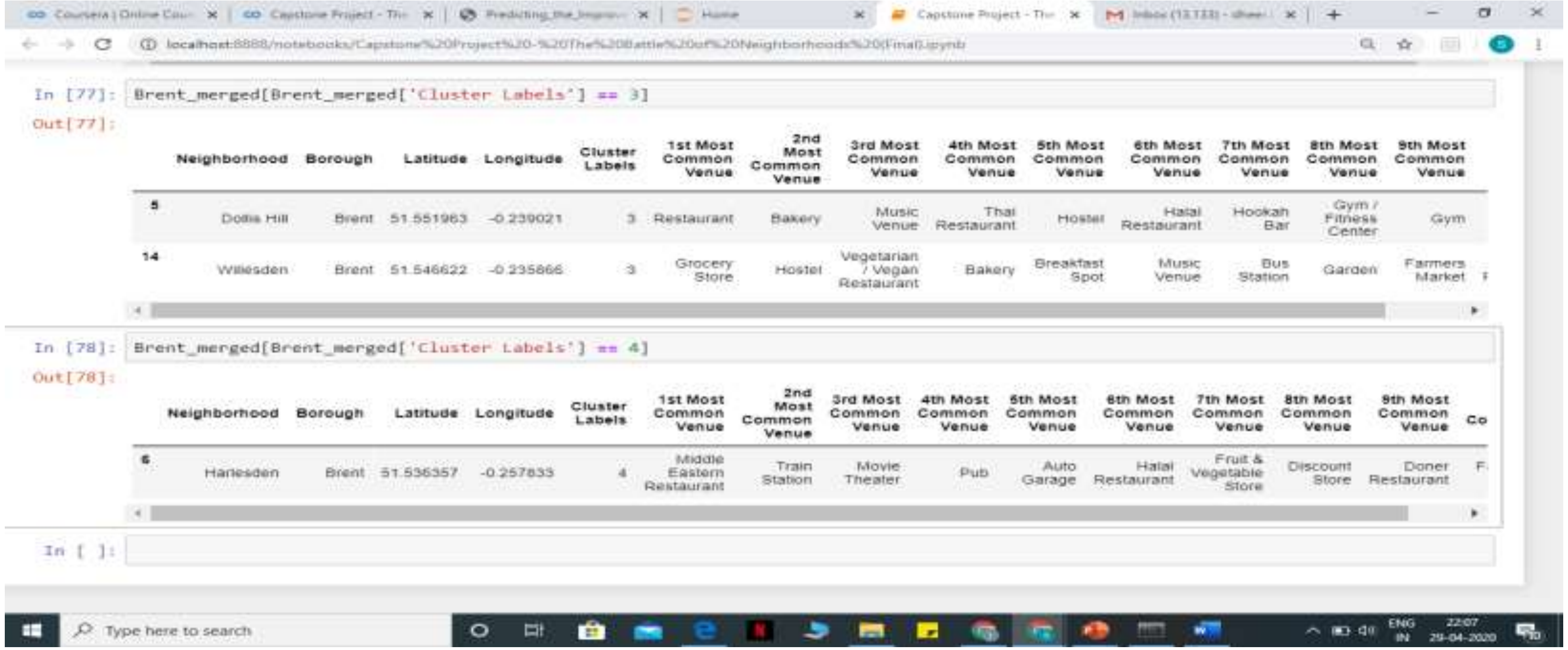
In [76]:
```
Brent_merged[Brent_merged['Cluster Labels'] == 2]
```
Out[76]:

	Neighborhood	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
12	Queens Park	Brent	51.529149	-0.208813	2	Yoga Studio	Gym	Town Hall	Café	Garden	Italian Restaurant	Train Station	Flower Shop	Diner

In [77]:
```
Brent_merged[Brent_merged['Cluster Labels'] == 3]
```
Out[77]:

	Neighborhood	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
6	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Music [illegible]	Thai [illegible]	[illegible]	Halal [illegible]	Hookah [illegible]	Gym / [illegible]	[illegible]

- Fourth and Fifth Cluster with Venues, we can consider this fourth cluster also since it has the neighbourhood i.e Willesden with hospital nearby.



```
In [77]: Brent_merged[Brent_merged['Cluster Labels'] == 3]
```

Out[77]:

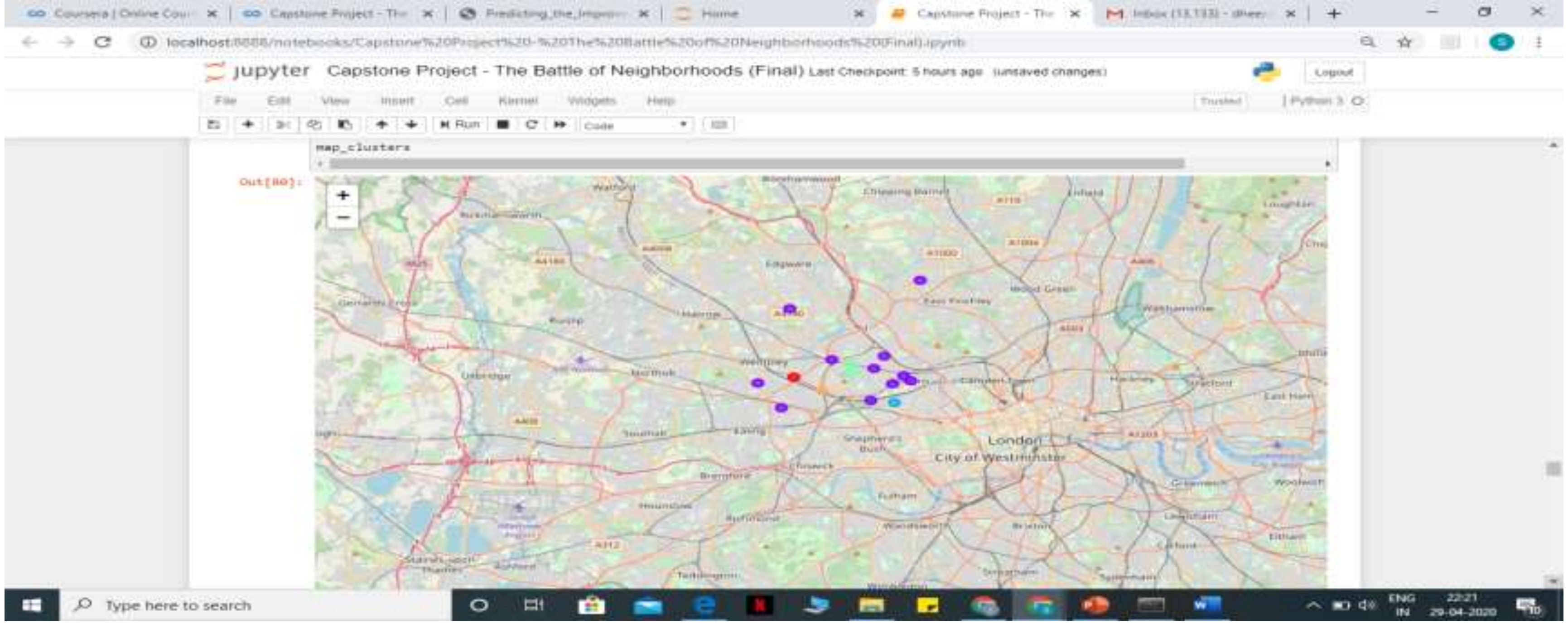
	Neighborhood	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
5	Dollis Hill	Brent	51.551963	-0.239021	3	Restaurant	Bakery	Music Venue	Thai Restaurant	Hostel	Halal Restaurant	Hookah Bar	Gym / Fitness Center	Gym
14	Willesden	Brent	51.546622	-0.235866	3	Grocery Store	Hostel	Vegetarian / Vegan Restaurant	Bakery	Breakfast Spot	Music Venue	Bus Station	Garden	Farmers Market

```
In [78]: Brent_merged[Brent_merged['Cluster Labels'] == 4]
```

Out[78]:

	Neighborhood	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	Co
6	Harlesden	Brent	51.536357	-0.257833	4	Middle Eastern Restaurant	Train Station	Movie Theater	Pub	Auto Garage	Halal Restaurant	Fruit & Vegetable Store	Discount Store	Doner Restaurant	F

- Visualizing a clustered map using folium library
- Each cluster is color coded for the ease of presentation. We can see that majority of neighbourhood falls in purple which is second cluster. First, Third and fifth cluster are in Red, Yellow and blue. Whereas fourth cluster is lemon green in color.



Discussion

- The project is created by looking at the current global scenario of coronavirus pandemic. No wonder how long it prevails in the global market and how worst effect it can have to the economy of the countries round the world. The aim of the project is to assist the present government of London as well the health care ministry in tackling the spread of coronavirus in the worst effected areas(Boroughs) of London.
- The availability of hospitals and health care facilities should be the prior need of all people living there. Hence considering that the 3 neighbourhoods have been figured out for the worst affected region there. Hence cluster 1 should be preferred first.
- For creating institutional quarantine facilities, second cluster and fourth cluster with maximum number of venues as well close proximity of the hospitals should be chosen. Apart from that other facilities like voluntarily contribution of restaurant owners for providing online food deliveries at this time of crisis should be considered as one major criteria as a preferred cluster. Apart form that there are certain clusters like 3rd & 4th which has townhall & hostel facilities and are closed for now and hence can be converted to Institutional quarantine facilities.
- First and fifth cluster should be considered for there containment of food and vegetable supplies to the places which are strictly under lockdown.
- Similarly, third cluster has a good transportation facility which is suitable for movement of the required emergency medicines as well as medical equipment to meet the enormous demand.

Conclusion

This project provides a better understanding of the clustering of the neighbourhoods in terms of common venues or specific places based on certain categories mentioned like hospitals, gyms, restaurant, grocery stores, etc. Hence, we can efficiently use data science to solve the problem of coronavirus spread in certain hotspots of the country where the infections have gone out of control and risking the life of people in those areas. The project can be further enhanced to building of more and more hospitals and transfer of medicines and medical equipment for the treatment of Covid-19 critically ill patients.