COURSERA CAPSTONE – BATTLE OF THE NEIGHBORHOODS – NEW YORK CITY

Nita Ahuja

JUNE 12, 2020

Coursera Capstone: Battle of the Neighborhoods – New York City

Table of contents

Introduction

Discussion and Description of the Problem:

New York City – there is no place quite like it! Even if you think you know New York City well - the world class museums, amazing food and unforgettable views - there is always something new and exciting to discover. Thrilling shows, fabulous shopping, and first-rate sports draw visitors from around the globe to the world's entertainment capital. With some of the busiest airports in the world it is a global destination. New York City is the largest city in the United States and with a population of 8,550,971 in 2020 it is also the most densely populated major city in the country. New York City is composed of 5 boroughs: The Bronx, Brooklyn, Manhattan, Queens, and Staten Island.

In early 2020, New York City was struck by the novel coronavirus pandemic known as Covid-19. March 1 saw the first confirmed case of COVID-19 in New York State, a 39-year-old woman who lived in Manhattan. As of June 1, New York City's confirmed cases were 200,830 and New York City deaths were 21,607. "Deaths" include test-confirmed cases as well as those with "COVID-19" or equivalent listed on the death certificate. The most brutal toll came among those who were old, poor and in the outer boroughs. The city's deaths are 10 times those of Los Angeles County's. They have surpassed the 16,000 lives lost in Italy's hard-hit Lombardy region. In the U.K., eight times as populous as New York City, about 37,500 have died. With New York's outbreak eclipsing others around the world, it is logical to look into and analyze the population density by neighborhood and borough as well as the Covid-19 cases by neighborhood and borough to determine which ones were the hardest hit. Based on information provided by the New York City Department of Health and Mental Hygiene, the governor's office, The COVID Tracking Project, and the Center for Systems Science and Engineering at Johns Hopkins University, THE CITY is tracking hospital bed and intensive care unit availability in New York City, including efforts to increase capacity. As of March 28, New York State had about 53,000 hospital beds across the state, with 3,000 ICU beds. New York Governor Andrew Cuomo

estimated that the state will need 140,000 beds and asked President Trump to authorize four additional hospital sites. In the context of the population density, the number of cases as well as deaths it also seems logical to examine where the hospitals are located within the five boroughs and to also determine if it would be a good idea to add a new hospital to an area of the city where the cases have been highest and is being underserved at the current time in the event of future pandemics.

Project Interest:

This information and analysis would be beneficial to New York City residents, New York State government agencies including New York State Department of Health, New York State and City Planning authorities as well as Federal agencies like Department of Health and Human Services.

Data

Data Description including Data Sources:

In order to analyze the population density by neighborhood and borough to determine which neighborhoods have the highest population density, we will need a couple of resources:

1. We will need New York City population data for 2020 by borough. This will need to include information by borough, age over 65, school age children. This can be obtained in the form of a csv file from https://data.cityofnewyork.us/City-Government/2020-population/t8c6-3i7b
   This will give us a good idea of the population composition in terms of greatest risk and will allow us to determine which boroughs have the highest population.
2. In addition, we will need New York City population information by neighborhoods. This will include the neighborhood name, borough and population and can be obtained from https://data.cityofnewyork.us/City-Government/New-York-City-Population-By-Neighborhood-Tabulatio/swpk-hqdp/data
   This will give us an insight of which neighborhoods within each borough having the highest populations.

In order to analyze the Covid-19 cases by neighborhood and borough to determine which ones were the hardest hit we will need to get data from a number of sources:

1. We will need New York City Covid-19 case data by neighborhood and boroughs. This will need to include the number of cases, the number who tested positive as well as the

number of Covid-19 deaths. This data also includes the zip codes of the neighborhoods. This data can be obtained in the form of a csv file from github on nychealth-coronavirus-data.
https://github.com/nychealth/coronavirus-data/blob/master/data-by-modzcta.csv
This will give us a breakdown of the Covid-19 cases by neighborhood.

2. We will need the US zip code Latitude and Longitude file in the form of a csv file in order to determine the corresponding latitude and longitude of the various New York City neighborhoods.
https://public.opendatasoft.com/explore/dataset/us-zip-code-latitude-and-longitude/export/?refine.state=NY

In order to analyze and examine where the hospitals are located within the five boroughs and to determine if it would make sense to add another hospital to an area of the city where the cases have been highest, we will need to get data from a number of different sources:

1. We will need data on all the hospitals located in the five boroughs of New York City. This data can be obtained from https://www.newyorkled.com/nyc-hospitals-map-and-list/ in the form of a geoJSON file.
2. We will also need the information on the staffed beds for each hospital in New York City. This data will come from the Individual Hospital Statistics for New York. It includes hospital name, staffed beds and city name for all hospitals in New York State. We will only need the data for New York City hospitals. This data can be obtained from https://www.ahd.com/states/hospital_NY.html It will need to be converted into an excel or csv file.
3. We will need to use the Foursquare API to explore and segment the boroughs to examine the hospital locations that already exist with their bed capacity in order to determine where one might be needed to serve the residents of New York City better.
4. We will use Folium to visualize the results.

Methodology

New York City is comprised of five boroughs and about 306 neighborhoods. In this project we will explore and analyze the data to identify the 25 most populous boroughs and neighborhoods in New York City. Next, we will bring in the data for the COVID-19 Case Counts, Death Counts, and Percent positive by neighborhood. We will determine the 25 neighborhoods with the highest Corornavirus Case counts, 25 neighborhoods with the highest Death Counts from Coronavirus and the 25 neighborhoods with the highest percent testing positive for the Corronavirus also known as COVID-19. These neighborhoods may be different and not necessarily the same as a neighborhood may have high case counts but low death counts or

high positive percent but low death counts. We will then determine which 5 neighborhoods appear on all 3 lists as well as on the highest population list. We will then analyze and compare these neighborhoods based on the hospital information in terms of how many hospitals are located within the neighborhood vicinity of 3000 meters (roughly 1.86 miles). Also, we will analyze how many staffed hospital beds are there in the neighborhood to service the neighborhood residents.

Analysis

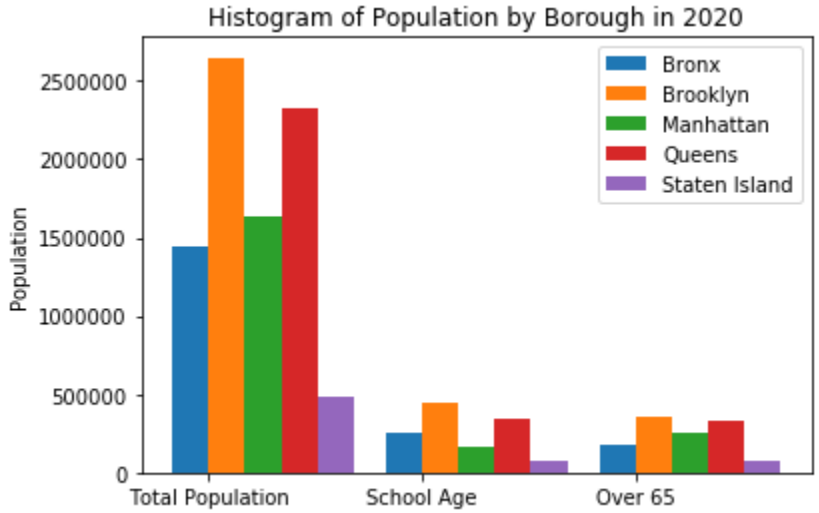
Once the New York City population data for 2020 has been imported into a dataframe, the data is cleaned. Unnecessary columns are dropped, columns are renamed to provide clarity and the index is reset. The data is then split into 3 different dataframes by Borough based on the categories of Population Total, Over 65, and School Age populations. All 3 dataframes are then merged into one dataframe. This is then converted into a list and transposed as follows:

Borough	Bronx	Brooklyn	Manhattan	Queens	Staten Island
Population_Tot	1446788	2648452	1638281	2330295	487155
Population_Sch	259013	441049	162931	341062	78759
Population_O65	171856	351609	250806	325300	77644

The data is then used to plot a histogram:

```
bar_width = 0.17
fig, ax = plt.subplots()
#barNYC = ax.bar(indx - bar_width/2, nyc_list, bar_width, label='New York City')
barbronx = ax.bar(indx - bar_width/2, bronx_list, bar_width, label='Bronx')
barbrooklyn = ax.bar(indx + bar_width/2 , brooklyn_list, bar_width, label='Brooklyn')
barmanhattan = ax.bar(indx + bar_width/2 + bar_width , manhattan_list, bar_width, label='Manhattan')
barqueens = ax.bar(indx + bar_width/2 + bar_width + bar_width , queens_list, bar_width, label='Queens')
barstatenisland = ax.bar(indx + bar_width/2 + bar_width + bar_width + bar_width , staten_list, bar_width, label='Staten Island')

ax.set_xticks(indx)
ax.set_xticklabels(population)
ax.legend()

plt.title('Histogram of Population by Borough in 2020') # add a title to the histogram
plt.ylabel('Population') # add y-label

plt.show()
```

Coursera Capstone: Battle of the Neighborhoods – New York City



From the above histogram we can see that the boroughs of Brooklyn and Queens have the highest Total populations. Also, we can see that there is a population that is Over 65 years of age and composed of School age children that may be considered vulnerable populations that are susceptible to disease in the event of future pandemics. Again, we can see that the boroughs of Brooklyn and Queens seem to have the highest numbers of vulnerable populations.

Next, we import the Population Data by Neighborhood of New York City in order to determine which neighborhoods have the highest populations. These neighborhoods are densely populated. The data imported is cleaned, re-indexed, and sorted in descending order of Total Population to find the top 25 neighborhoods with the highest populations. From the data below we can see that The Upper West Side neighborhood in Manhattan borough of New York City has the largest population of 132,378 compared to the other neighborhoods.

Coursera Capstone: Battle of the Neighborhoods – New York City

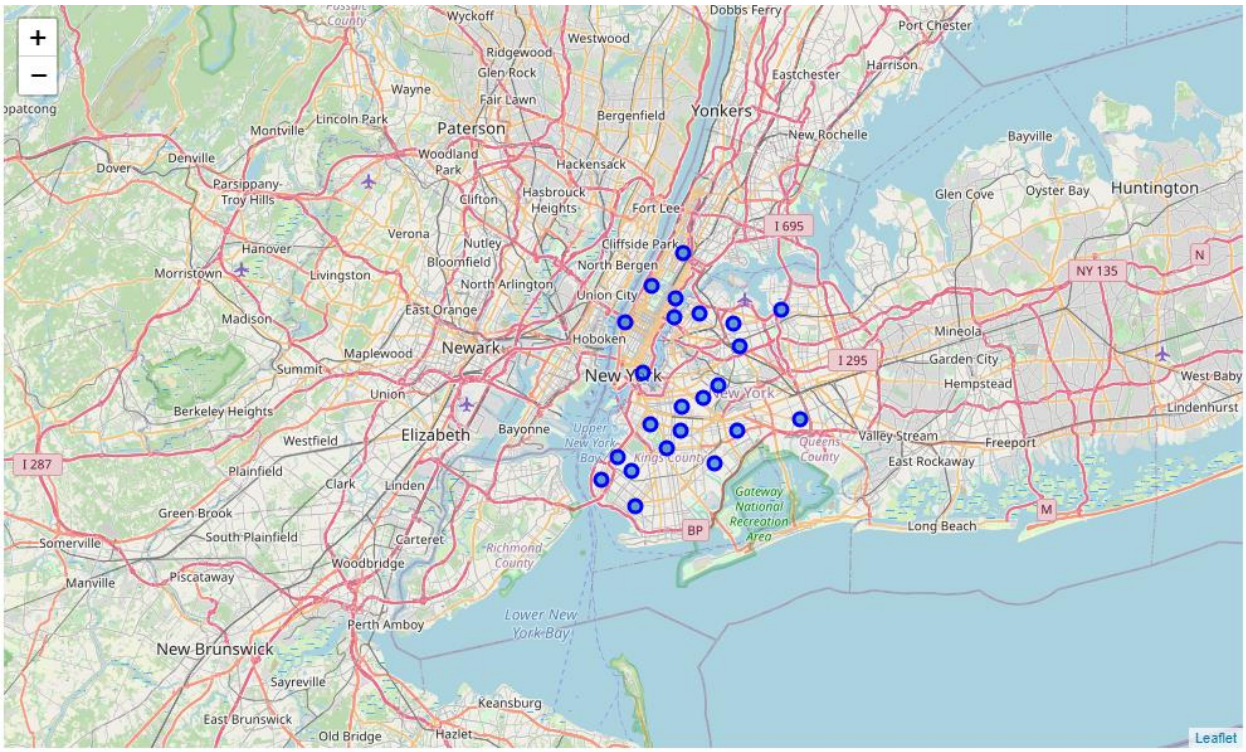
	Borough	County_Code	NTA_Code	Neighborhood	Population	Latitude	Longitude
0	Manhattan	61	MN12	Upper West Side	132378	40.787045	-73.975416
1	Queens	81	QN28	Jackson Heights	108152	40.755656	-73.885775
2	Brooklyn	47	BK88	Borough Park	106357	40.633993	-73.996806
3	Brooklyn	47	BK42	Flatbush	105804	40.652048	-73.959027
4	Brooklyn	47	BK61	Crown Heights North	103169	40.667471	-73.943566
5	Brooklyn	47	BK82	East New York	91958	40.666770	-73.882358
6	Brooklyn	47	BK28	Bensonhurst West	88727	40.604977	-73.993406
7	Queens	81	QN29	Elmhurst	88427	40.736580	-73.878393
8	Manhattan	61	MN36	Washington Heights South	84438	40.883325	-72.515328
9	Queens	81	QN17	Forest Hills	83728	38.215348	-85.585793
10	Brooklyn	47	BK50	Canarsie	83693	40.640232	-73.906059
11	Manhattan	61	MN31	Lenox Hill-Roosevelt Island	80771	40.760500	-73.951000
12	Brooklyn	47	BK31	Bay Ridge	79371	40.626400	-74.029900
13	Queens	81	QN70	Astoria	78793	40.764400	-73.923500
14	Manhattan	61	MN32	Yorkville	77942	40.776200	-73.949200
15	Queens	81	QN55	South Ozone Park	75878	40.676400	-73.812500
16	Manhattan	61	MN03	Central Harlem North-Polo Grounds	75282	40.814300	-73.940100
17	Manhattan	61	MN28	Lower East Side	72957	40.715000	-73.984300
18	Brooklyn	47	BK34	Sunset Park East	72340	40.645531	-74.012383
19	Brooklyn	47	BK78	Bushwick South	72101	40.694270	-73.918748
20	Queens	81	QN22	Flushing	72008	40.767500	-73.833100
21	Brooklyn	47	BK75	Bedford	70713	40.687200	-73.941800
22	Manhattan	61	MN13	Hudson Yards-Chelsea-Flat Iron-Union Square	70150	40.756100	-74.003500
23	Queens	81	QN20	Ridgewood	69317	40.704400	-73.901800
24	Brooklyn	47	BK37	Park Slope-Gowanus	67649	40.672400	-73.977100

We can see from the above data that the top 25 neighborhoods belong to 3 boroughs namely Brooklyn, Manhattan, and Queens.

We will create a Folium map using New York City latitude and longitude as the center points and plot these top 25 most populated neighborhoods using their latitude and longitude values on it.

Coursera Capstone: Battle of the Neighborhoods – New York City

```
# create map of New York using latitude and longitude values
map_newyork = folium.Map(location=[latitude, longitude], zoom_start=10)

# add markers to map
for lat, lng, borough, neighborhood in zip(df_top25['Latitude'], df_top25['Longitude'], df_top25['Borough'], df_top25['Neighborh
ood']):
    label = '{}, {}'.format(neighborhood, borough)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(map_newyork)

map_newyork
```



Next, we import the COVID-19 data by neighborhood and borough. This data contains information on the number of Covid-19 Cases, number of Covid-19 Deaths and the percent positive by neighborhood. This data is cleaned, indexed, stripped for spaces. All string fields are removed and the data is then normalized using sklearn.preprocessing.StandardScalar. This standardizes the features by removing the mean and scaling to unit variance. Standardization of a dataset is a common requirement for many machine learning estimators. Normalization is a statistical method that helps mathematical-based algorithms interpret features with different magnitudes and distributions equally.

Coursera Capstone: Battle of the Neighborhoods – New York City

	ZIP_CODE	COVID_CASE_COUNT	COVID_DEATH_COUNT	PERCENT_POSITIVE
0	10001	366	21	15.38
1	10002	1056	149	20.86
2	10003	450	33	12.18
3	10004	31	1	12.11
4	10005	61	2	10.76

```
from sklearn.preprocessing import StandardScaler

X = df.values[:,1:]
X = np.nan_to_num(X)
cluster_dataset = StandardScaler().fit_transform(X)
cluster_dataset
```

We then use the K-Means clustering algorithm as we are interested in exploring the neighborhood behavior of the case counts, death counts, and percent positive as it relates to the various neighborhoods. We are partitioning the neighborhoods into groups that have similar characteristics in this case numbers of covid cases, number of deaths and percent positive. We run the model and group/segment our neighborhoods into 3 clusters. We do this by assigning a label to each neighborhood.

```
num_clusters = 3

k_means = KMeans(init="k-means++", n_clusters=num_clusters, n_init=12)
k_means.fit(cluster_dataset)
labels = k_means.labels_

print(labels)
```

	ZIP_CODE	COVID_CASE_COUNT	COVID_DEATH_COUNT	PERCENT_POSITIVE
Labels				
0	10946.957746	892.521127	69.690141	26.772394
1	10924.074074	2093.092593	191.981481	27.105926
2	10505.826923	404.326923	33.615385	15.210000

The labels assigned are 0, 1, 2 and we can see that Zip codes with the

label 2 have lowest numbers for COVID_CASE_COUNT (404.326723), COVID_DEATH_COUNT(33.615385) and PERCENT_POSITIVE(15.2100000)

label 0 have medium numbers for COVID_CASE_COUNT (892.521127), COVID_DEATH_COUNT(69.690141) and PERCENT_POSITIVE(26.772394)

label 1 have highest numbers for COVID_CASE_COUNT (2093.092593), COVID_DEATH_COUNT(191.981481) and PERCENT_POSITIVE(27.105926)

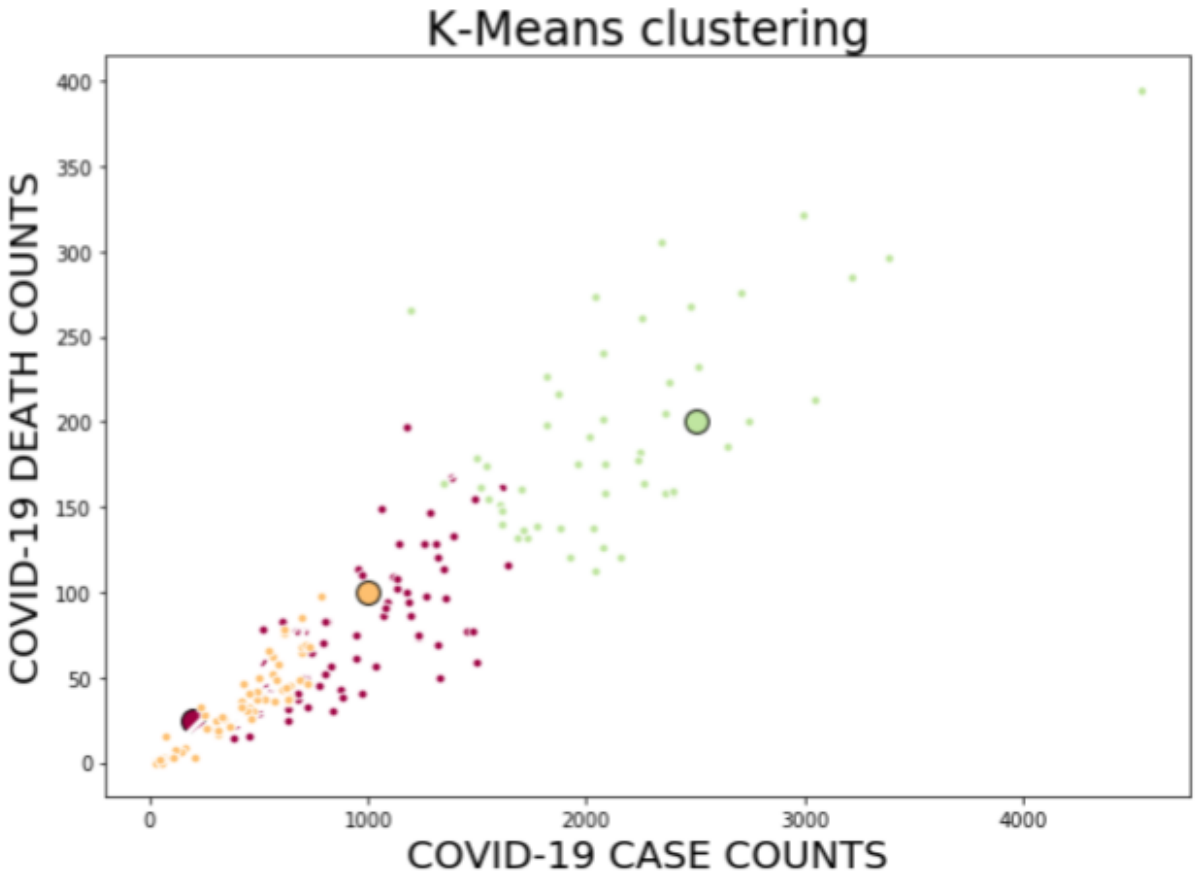
Using this we can create a profile for each neighborhood (High Risk, Medium Risk, Low Risk).

Coursera Capstone: Battle of the Neighborhoods – New York City

We can use the information along with the labels to visualize the 3 segmented clusters. They each appear in different colors.

```
# initialize the plot with the specified dimensions.
fig = plt.figure(figsize=(10, 7))
k_means_cluster_centers = [[200,25],[1000, 100],[2500, 200]]

colors = plt.cm.Spectral(np.linspace(0, 1, 4))
# create a plot
ax = fig.add_subplot(1, 1, 1)

# loop through the data and plot the datapoints and centroids.
# k will range from 0-3, which will match the number of clusters in the dataset.
for k, col in zip(range(len(k_means_cluster_centers)), colors):
#for k, col in zip(range(len([[250,30], [1000, 100], [2000, 175]])), colors):

    # create a list of all datapoints, where the datapoitns that are
    # in the cluster (ex. cluster 0) are labeled as true, else they are
    # labeled as false.
    my_members = (k_means_labels == k)

    # define the centroid, or cluster center.
    cluster_center = k_means_cluster_centers[k]

    # plot the datapoints with color col.
    ax.plot(X[my_members, 0], X[my_members, 1], 'w', markerfacecolor=col, marker='.',markersize=10)

    # plot the centroids with specified color, but with a darker outline
    ax.plot(cluster_center[0], cluster_center[1], 'o', markerfacecolor=col,  markeredgecolor='k', markersize=12)

# title of the plot
ax.set_title('KMeans')

# add labels to axes
ax.set_xlabel('COVID-19 CASE COUNTS', fontsize=20)
ax.set_ylabel('COVID-19 DEATH COUNTS', fontsize=20)

# add title to figure
ax.set_title('K-Means clustering', fontsize=24)

# show the plot
plt.show()
```

Coursera Capstone: Battle of the Neighborhoods – New York City



Next, we determine which are the neighborhoods that have the highest case count, death count, and positive percent. There are some neighborhoods that have high case counts, but not high death counts while others that have high percent positive, but not high death counts. So, we will need to map each of them separately in different maps. We also need to include the latitude and longitude information of these neighborhoods as it is not contained in the data. In order to do that, we import the Zip code file that contains the latitude and longitude information for each New York City neighborhood and merge this with the Covid-19 neighborhood data based on the zip code information that is common to both tables.

We then get sort the dataframe in descending order to get the top 25 neighborhoods with the highest COVID case counts.

Coursera Capstone: Battle of the Neighborhoods – New York City

	ZIP_CODE	NEIGHBORHOOD	BOROUGH	COVID_CASE_COUNT	COVID_DEATH_COUNT	PERCENT_POSITIVE	TOTAL_COVID_TESTS	LATITUDE	LONGITUDE
0	11368	Corona/North Corona	Queens	4544	395	34.72	13088	40.75000	-73.8750
1	10467	Allerton/Norwood/Pelham Parkway/Williamsbridge	Bronx	3383	296	30.57	11066	40.87500	-73.8750
2	11373	Elmhurst	Queens	3221	285	30.64	10513	40.75000	-73.8750
3	11219	Borough Park	Brooklyn	3051	213	21.95	13898	40.62500	-74.0000
4	10469	Allerton/Baychester/Pelham Gardens/Williamsbridge	Bronx	2993	322	30.74	9735	40.87500	-73.8750
5	10468	Fordham/Kingsbridge/University Heights	Bronx	2747	200	28.44	9658	40.87500	-73.8750
6	11236	Canarsie	Brooklyn	2711	276	29.90	9068	40.62500	-73.8750
7	10314	Bloomfield/Freshkills Park	Staten Island	2645	186	26.05	10152	40.59375	-74.1250
8	11372	Jackson Heights	Queens	2517	232	31.56	7976	40.75000	-73.8750
9	10456	Claremont/Morrisania	Bronx	2480	268	27.35	9067	40.84375	-73.9375
10	10452	Concourse/Highbridge	Bronx	2400	158	29.12	8241	40.84375	-73.9375
11	10466	Edenwald/Wakefield	Bronx	2393	159	31.17	7678	40.90625	-73.8750
12	11208	Cypress Hills/East New York	Brooklyn	2382	223	30.04	7930	40.68750	-73.8750
13	11230	Midwood	Brooklyn	2358	205	24.19	9747	40.62500	-73.9375
14	10453	Morris Heights/Mount Hope/University Heights	Bronx	2357	158	30.12	7825	40.84375	-73.9375
15	11691	Edgemere/Far Rockaway	Queens	2340	306	29.14	8030	40.59375	-73.7500
16	10462	Parkchester/Pelham Parkway/Van Nest/Westcheste...	Bronx	2267	164	27.26	8315	40.84375	-73.8750
17	11226	Flatbush/Prospect Lefferts Gardens	Brooklyn	2255	261	27.28	8267	40.65625	-73.9375
18	11234	Bergen Beach/Flatlands/Marine Park/Mill Basin	Brooklyn	2242	182	27.48	8159	40.62500	-73.9375
19	11377	Woodside	Queens	2238	177	25.99	8611	40.75000	-73.8750
20	11385	Glendale/Ridgewood	Queens	2157	120	27.30	7902	40.68750	-73.8750
22	10457	Belmont/Claremont/Mount Hope/Tremont	Bronx	2087	175	27.84	7496	40.84375	-73.8750
21	11434	Airport/South Jamaica/Springfield Gardens/St. ...	Queens	2087	158	30.24	6902	40.68750	-73.7500
23	11211	East Williamsburg/Williamsburg (North)/William...	Brooklyn	2080	126	19.48	10678	40.71875	-73.9375
24	11207	Cypress Hills/East New York	Brooklyn	2077	201	27.27	7617	40.65625	-73.8750

Using the geographical coordinates of New York City, we create a Folium map and plot the top 25 neighborhoods based on their latitude and longitude information.

Coursera Capstone: Battle of the Neighborhoods – New York City

```
address = 'New York City, NY'

geolocator = Nominatim(user_agent="ny_explorer")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of New York City are {}, {}.'.format(latitude, longitude))
```

```
The geograpical coordinate of New York City are 40.7127281, -74.0060152.
```

```
# create map of NYC using latitude and longitude values
map_nyc = folium.Map(location=[latitude, longitude], zoom_start=10)

title_html = '''
             <h3 align="center" style="font-size:20px"><b>NEIGHBORHOODS WITH HIGHEST COVID-19 CASE COUNTS</b></h3>
             '''
map_nyc.get_root().html.add_child(folium.Element(title_html))
# add markers to map
for lat, lng, borough, neighborhood, covidcounts, coviddeaths, pcpositive in zip(top25_df['LATITUDE'], top25_df['LONGITUDE'], to
p25_df['BOROUGH'], top25_df['NEIGHBORHOOD'], top25_df['COVID_CASE_COUNT'], top25_df['COVID_DEATH_COUNT'], top25_df['PERCENT_POSI
TIVE']):
    label = '{}, {}, Covid Case Count: {}, Covid Death Count: {}, Percent Positive: {}%'.format(neighborhood, borough, covidcoun
ts, coviddeaths, pcpositive)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='red',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(map_nyc)

map_nyc
```

We now determine the neighborhoods with the highest COVID death counts. The top 25 neighborhoods having the highest COVID-19 deaths:

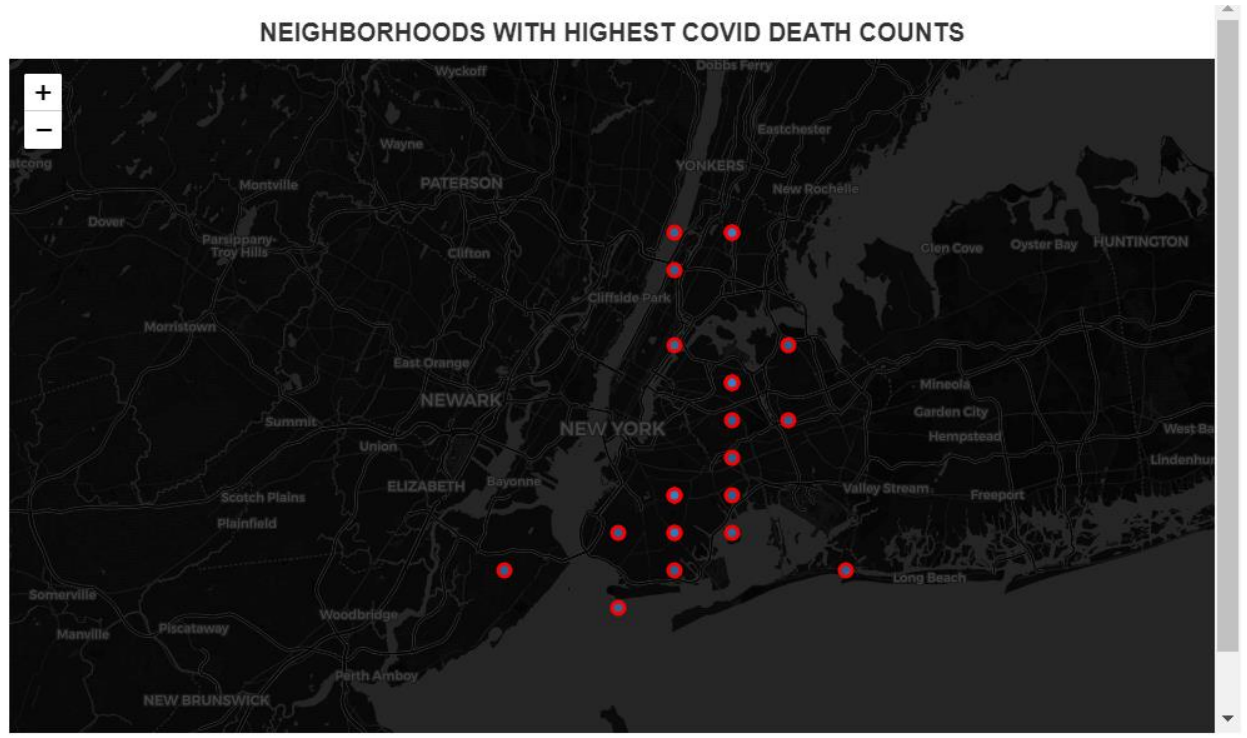
Coursera Capstone: Battle of the Neighborhoods – New York City

	ZIP_CODE	NEIGHBORHOOD	BOROUGH	COVID_CASE_COUNT	COVID_DEATH_COUNT	PERCENT_POSITIVE	TOTAL_COVID_TESTS	LATITUDE	LONGITUDE
0	11368	Corona/North Corona	Queens	4544	395	34.72	13088	40.75000	-73.8750
4	10469	Allerton/Baychester/Pelham Gardens/Williamsbridge	Bronx	2993	322	30.74	9735	40.87500	-73.8750
15	11691	Edgemere/Far Rockaway	Queens	2340	306	29.14	8030	40.59375	-73.7500
1	10467	Allerton/Norwood/Pelham Parkway/Williamsbridge	Bronx	3383	296	30.57	11066	40.87500	-73.8750
2	11373	Elmhurst	Queens	3221	285	30.64	10513	40.75000	-73.8750
6	11236	Canarsie	Brooklyn	2711	276	29.90	9068	40.62500	-73.8750
27	11235	Brighton Beach/Manhattan Beach/Sheepshead Bay	Brooklyn	2042	274	24.76	8246	40.59375	-73.9375
9	10456	Claremont/Morrisania	Bronx	2480	268	27.35	9067	40.84375	-73.9375
69	11354	Flushing/Murray Hill	Queens	1191	265	26.66	4467	40.78125	-73.8125
17	11226	Flatbush/Prospect Lefferts Gardens	Brooklyn	2255	261	27.28	8267	40.65625	-73.9375
25	11203	East Flatbush (North)/East Flatbush (South)	Brooklyn	2073	240	28.29	7328	40.65625	-73.9375
8	11372	Jackson Heights	Queens	2517	232	31.56	7976	40.75000	-73.8750
34	11432	Hillcrest/Jamaica Estates/Jamaica Hills	Queens	1816	227	29.88	6078	40.71875	-73.8125
12	11208	Cypress Hills/East New York	Brooklyn	2382	223	30.04	7930	40.68750	-73.8750
33	11212	Ocean Hill-Brownsville	Brooklyn	1873	216	26.65	7029	40.65625	-73.9375
3	11219	Borough Park	Brooklyn	3051	213	21.95	13898	40.62500	-74.0000
13	11230	Midwood	Brooklyn	2358	205	24.19	9747	40.62500	-73.9375
24	11207	Cypress Hills/East New York	Brooklyn	2077	201	27.27	7617	40.65625	-73.8750
5	10468	Fordham/Kingsbridge/University Heights	Bronx	2747	200	28.44	9658	40.87500	-73.8750
35	10029	East Harlem	Manhattan	1816	198	24.23	7494	40.78125	-73.9375
72	11224	Brighton Beach/Coney Island/Seagate	Brooklyn	1172	197	26.72	4386	40.56250	-74.0000
29	10463	Kingsbridge/Marble Hill/Riverdale/Spuyten Duyvil	Bronx	2016	191	24.10	8365	40.87500	-73.9375
7	10314	Bloomfield/Freshkills Park	Staten Island	2645	186	26.05	10152	40.59375	-74.1250
18	11234	Bergen Beach/Flatlands/Marine Park/Mill Basin	Brooklyn	2242	182	27.48	8159	40.62500	-73.9375
49	11375	Forest Hills	Queens	1500	179	24.92	6019	40.71875	-73.8750

Using the geographical coordinates of New York City, we create a Folium map and plot the top 25 neighborhoods based on their latitude and longitude information.

```
# create map of NYC using latitude and longitude values
map_nyc = folium.Map(location=[latitude, longitude], tiles='CartoDB dark_matter', zoom_start=10)
title_html = '''
             <h3 align="center" style="font-size:20px"><b>NEIGHBORHOODS WITH HIGHEST COVID DEATH COUNTS</b></h3>
             '''
map_nyc.get_root().html.add_child(folium.Element(title_html))

# add markers to map
for lat, lng, borough, neighborhood, covidcounts, coviddeaths, pcpositive in zip(top25_deathdf['LATITUDE'], top25_deathdf['LONGI
TUDE'], top25_deathdf['BOROUGH'], top25_deathdf['NEIGHBORHOOD'], top25_deathdf['COVID_CASE_COUNT'], top25_deathdf['COVID_DEATH_C
OUNT'], top25_deathdf['PERCENT_POSITIVE']):
    label = '{}, {}, Covid Case Count: {}, Covid Death Count: {}, Percent Positive: {}%'.format(neighborhood, borough, covidcoun
ts, coviddeaths, pcpositive)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='red',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(map_nyc)

map_nyc
```

Coursera Capstone: Battle of the Neighborhoods – New York City

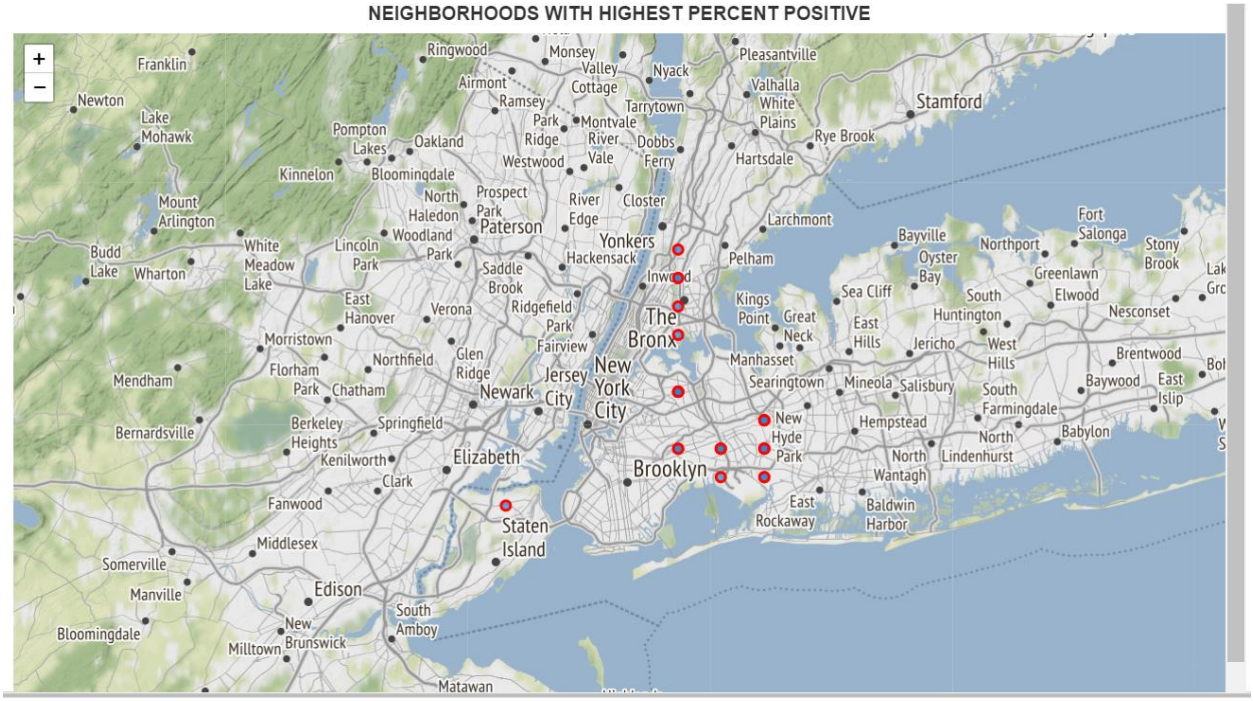


Next, we determine the neighborhoods with the highest COVID-19 positive percentages. The top 25 neighborhoods having the highest positive percentages:

Coursera Capstone: Battle of the Neighborhoods – New York City

	ZIP_CODE	NEIGHBORHOOD	BOROUGH	COVID_CASE_COUNT	COVID_DEATH_COUNT	PERCENT_POSITIVE	TOTAL_COVID_TESTS	LATITUDE	LONGITUDE
0	11368	Corona/North Corona	Queens	4544	395	34.72	13088	40.75000	-73.8750
43	11369	Airport/East Elmhurst	Queens	1610	162	34.56	4658	40.75000	-73.8750
110	11428	Queens Village	Queens	674	40	33.63	2004	40.71875	-73.7500
59	11370	Jackson Heights/Rikers Island	Queens	1331	50	33.16	4014	40.75000	-73.8750
30	10472	Soundview	Bronx	1958	175	32.57	6011	40.84375	-73.8750
109	11411	Cambria Heights	Queens	681	37	32.17	2117	40.68750	-73.7500
61	11413	Laurelton/Rosedale	Queens	1315	69	32.13	4093	40.65625	-73.7500
78	11423	Hollis/Holliswood	Queens	1075	91	32.04	3355	40.71875	-73.7500
114	10302	Elm Park	Staten Island	633	31	31.86	1987	40.62500	-74.1250
66	11412	St. Albans	Queens	1234	75	31.67	3897	40.68750	-73.7500
8	11372	Jackson Heights	Queens	2517	232	31.56	7976	40.75000	-73.8750
94	11414	Hamilton Beach/Howard Beach/Lindenwood	Queens	799	52	31.51	2536	40.65625	-73.8125
56	11420	South Ozone Park	Queens	1350	96	31.21	4326	40.68750	-73.8125
11	10466	Edenwald/Wakefield	Bronx	2393	159	31.17	7678	40.90625	-73.8750
102	11416	Ozone Park	Queens	709	50	31.16	2275	40.68750	-73.8750
87	11429	Queens Village	Queens	943	61	31.01	3041	40.71875	-73.7500
70	11421	Woodhaven	Queens	1188	94	30.78	3860	40.68750	-73.8750
77	11418	Richmond Hill	Queens	1091	94	30.78	3544	40.68750	-73.8125
4	10469	Allerton/Baychester/Pelham Gardens/Williamsbridge	Bronx	2993	322	30.74	9735	40.87500	-73.8750
51	10459	Charlotte Gardens/Hunts Point	Bronx	1487	155	30.69	4845	40.81250	-73.8750
2	11373	Elmhurst	Queens	3221	285	30.64	10513	40.75000	-73.8750
1	10467	Allerton/Norwood/Pelham Parkway/Williamsbridge	Bronx	3383	296	30.57	11066	40.87500	-73.8750
68	11419	Richmond Hill/South Ozone Park	Queens	1195	86	30.53	3914	40.68750	-73.8125
84	11422	Rosedale	Queens	971	41	30.34	3200	40.65625	-73.7500
21	11434	Airport/South Jamaica/Springfield Gardens/St. ...	Queens	2087	158	30.24	6902	40.68750	-73.7500

Using the geographical coordinates of New York City, we create a Folium map and plot the top 25 neighborhoods based on their latitude and longitude information.



Coursera Capstone: Battle of the Neighborhoods – New York City

At this point, we have established the 25 most populous neighborhoods as well as neighborhoods with the highest number of Covid-19 cases, highest death counts, and highest positive percent. When comparing these lists of neighborhoods, we can identify neighborhoods that belong on all these lists. We determined that the following five neighborhoods were one of the top 25 most populous as well as were in the top 25 list of the highest Covid-19 Case Counts, Death Counts and Positive Percent. We identify them as "**High Risk**" and will explore them in greater detail to determine if they are being serviced adequately by nearby hospitals.

NEIGHBORHOOD	BOROUGH	COVID_CASE_COUNT	COVID_DEATH_COUNT	PERCENT_POSITIVE	LATITUDE	LONGITUDE
Corona/North Corona	Queens	4534	395	34.57	40.75	-73.875
Allerton/Norwood/Pelham Parkway/Williamsbridge	Bronx	3379	295	30.48	40.875	-73.875
Elmhurst	Queens	3218	285	30.55	40.75	-73.875
Allerton/Baychester/Pelham Gardens/Williamsbridge	Bronx	2991	322	30.69	40.875	-73.875
Jackson Heights	Queens	2513	231	31.42	40.75	-73.875

Get hospital data

We download the hospital data in the form of a json file. This data contains the name of the hospital, its address location along with the latitude and longitude information that will be used for plotting. We perform data wrangling to get it in an appropriate format. A small sample is shown below.

	Hospital	Borough	Address	Latitude	Longitude
0	Albert Einstein Hospital	Bronx	Albert Einstein Hospital- Hospital, Eastcheste...	40.848885	-73.845555
1	Bronx Lebanon Hospital Center	Bronx	Bronx-Lebanon Hospital Center, Grand Concourse...	40.843481	-73.911050
2	Bronx Lebanon Hospital Center	Bronx	Bronx-Lebanon Hospital Center, Fulton Avenue, ...	40.831386	-73.903063
3	Calvary Hospital	Bronx	Calvary Hospital, Eastchester Road, Bronx, NY,...	40.848126	-73.843523
4	Jacobi Medical Center	Bronx	Jacobi Medical Center, Pelham Parkway South, N...	40.855673	-73.847908

```
neighborhoods.shape
```

```
(63, 5)
```

```
neighborhoods['Borough'].value_counts()
```

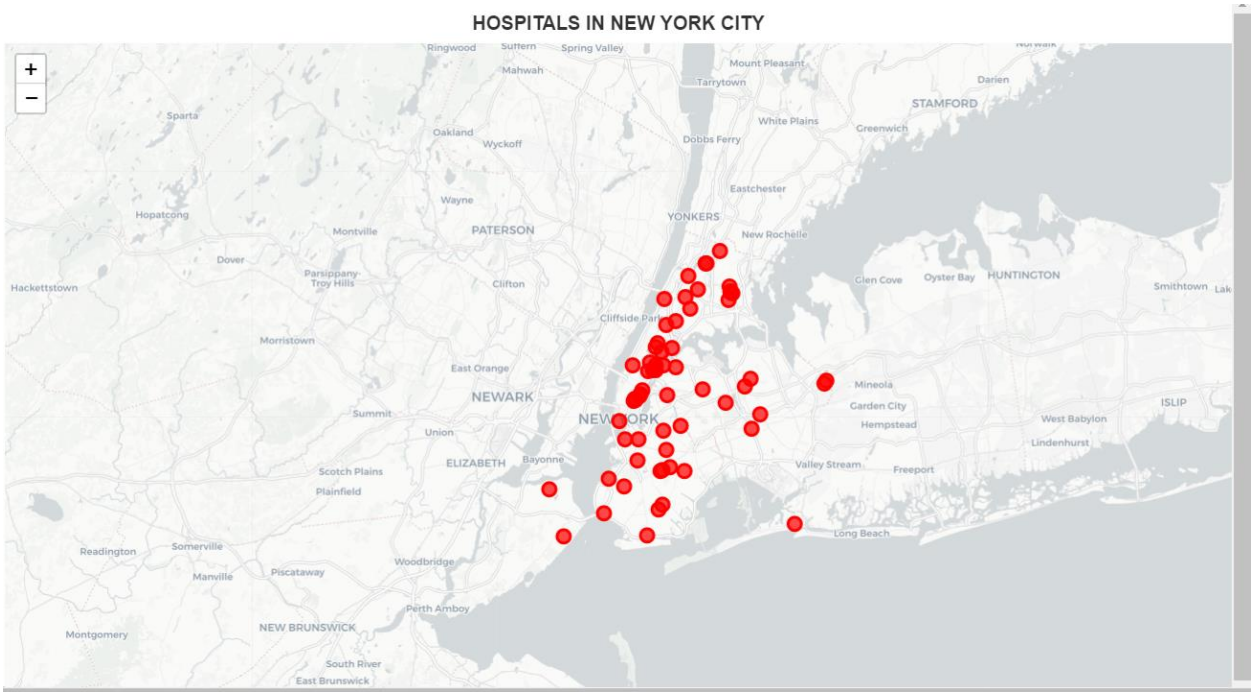
```
Manhattan        20
Brooklyn         16
Bronx            13
Queens           11
Staten Island     2
Manhatttan        1
Name: Borough, dtype: int64
```

Coursera Capstone: Battle of the Neighborhoods – New York City

We see that the dimensions of the dataframe are 63 rows by 5 columns. We determine the unique values of hospitals by Borough and can see that there are 21 hospitals servicing Manhattan, 16 hospitals servicing Brooklyn, 13 for Bronx, 11 for Queens, and 2 for Staten Island.

Using the latitude and longitude values of New York City as well as those of the hospitals we plot them on a Folium map of New York City with the hospitals superimposed on top using tiles = ‘cartodbpositron’.

```
# create map of NYC using Hospital latitude and longitude values
map_nyc = folium.Map(location=[latitude, longitude], tiles = 'cartodbpositron', zoom_start=10)
title_html = '''
             <h3 align="center" style="font-size:20px"><b>HOSPITALS IN NEW YORK CITY</b></h3>
             '''
map_nyc.get_root().html.add_child(folium.Element(title_html))

# add markers to map
for lat, lng, borough, hospital, address in zip(neighborhoods['Latitude'], neighborhoods['Longitude'], neighborhoods['Borough'], neighborhoods['Hospital'], neighborhoods['Address']):
    label = '{}, {}, {}'.format(hospital, borough, address)
    label = folium.Popup(label, parse_html=True)
    folium. CircleMarker(
        [lat, lng],
        radius=7,
        popup=label,
        color='red',
        fill=True,
        fill_color='red',
        fill_opacity=0.7,
        parse_html=False).add_to(map_nyc)

map_nyc
```



Coursera Capstone: Battle of the Neighborhoods – New York City

Using Foursquare API to Explore the Five Neighborhoods Identified as "High Risk":

Let us explore the five neighborhoods identified as "High Risk" in our dataframe.

Explore Corona/North Corona Neighborhood in Queens

Latitude and longitude values of Corona/North Corona are 40.75, -73.875.
We construct a URL to send a request to the Foursquare API to search for a specific type of Venue, in this case "Hospital". We set a LIMIT of 30 and a radius of 3000 metres (approx. 1.86 miles)

```
search_query = 'Hospital'
LIMIT = 30
radius = 3000
url = 'https://api.foursquare.com/v2/venues/search?&client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(
    CLIENT_ID,CLIENT_SECRET,neighborhood_latitude,neighborhood_longitude,VERSION,search_query,radius,LIMIT)
url
```

The results are received in the json format and this is stored in results.

```
results = requests.get(url).json()
results
```

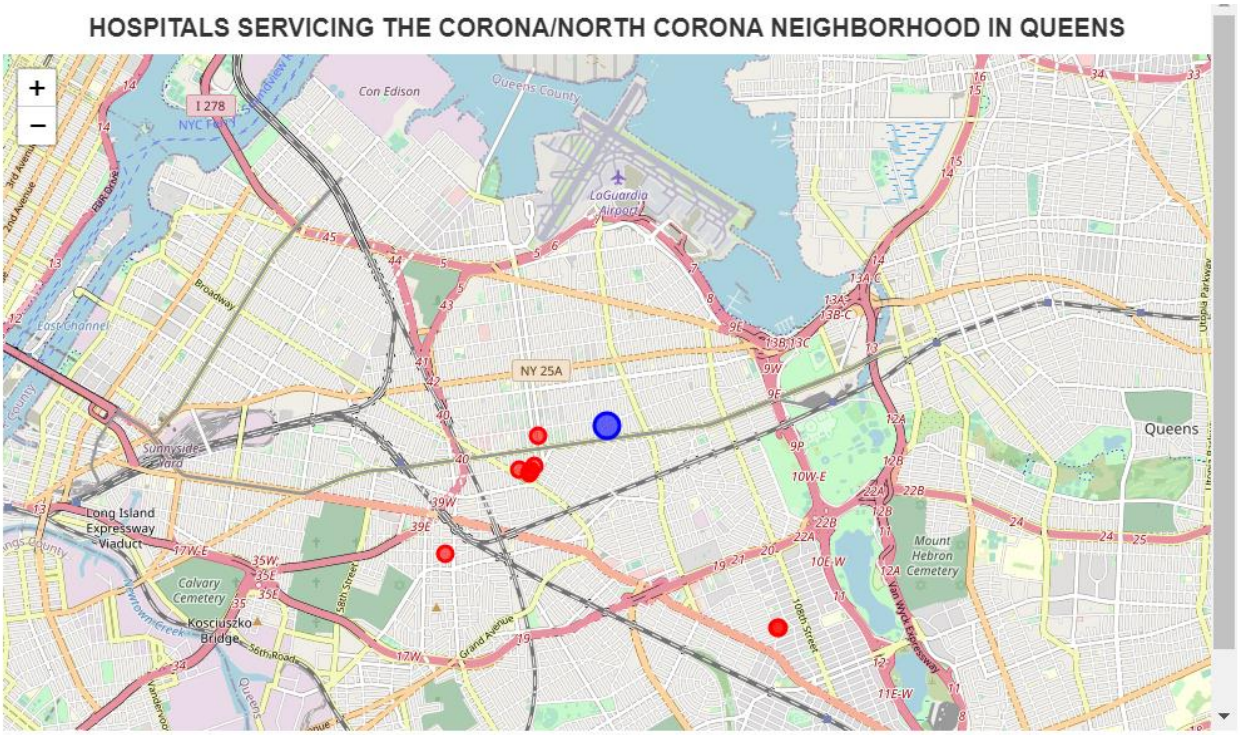
The venues are transformed into a pandas dataframe. The data is comprised of any venue with the term "Hospital" in it. This data is cleaned as it includes Hospitals, Veterinarian, Hospital Library, Café in Hospital, Bookstore, Hospital Gym/Fitness center etc. The categories are filtered on the keyword "Hospital". The index is reset. The result is as follows:

	name	categories	address	crossStreet	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	id
0	Elmhurst Hospital Center	Hospital	79-01 Broadway	at 78th St.	40.745244	-73.885536	[{'label': 'display', 'lat': 40.74524449809677...	1034	11373	US	Elmhurst	NY	United States	[79-01 Broadway (at 78th St.), Elmhurst, NY 11...	4be0708998f2a593de40c25a
1	Elmhurst Hospital Senior Care	Hospital	41 Ave	NaN	40.745266	-73.885393	[{'label': 'display', 'lat': 40.74526551814158...	1022	11373	US	Elmhurst	NY	United States	[41 Ave, Elmhurst, NY 11373, United States]	4f392dc9e4b0653d4c81946b
2	Long Island Jewish Forest Hills Hospital	Hospital	102-01 66th Rd	North Shore-LIJ Health System	40.729062	-73.851625	[{'label': 'display', 'lat': 40.7290618, 'lng'...	3052	11375	US	Forest Hills	NY	United States	[102-01 66th Rd (North Shore-LIJ Health System...	4b7d6f3df964a52009be2fe3
3	Elmhurst Hospital Center Mortuary Department	Hospital	79-01 Broadway	Baxter Av	40.745110	-73.885597	[{'label': 'display', 'lat': 40.74510955810547...	1046	11373	US	Elmhurst	NY	United States	[79-01 Broadway (Baxter Av), Elmhurst, NY 1137...	4db8388ecda1051541235683
4	Elmhurst Hospital Center Clinics	Hospital	79-01 Broadway	NaN	40.745869	-73.884859	[{'label': 'display', 'lat': 40.74586850915786...	950	11373	US	Elmhurst	NY	United States	[79-01 Broadway, Elmhurst, NY 11373, United St...	4ec6c16c722eeaa24d2f719f
5	Forest Hills Hospital - Nursery	Hospital	102-01 66th Rd	NaN	40.729031	-73.851509	[{'label': 'display', 'lat': 40.72903060913086...	3061	11375	US	Forest Hills	NY	United States	[102-01 66th Rd, Forest Hills, NY 11375, Unite...	4dd18a75d22d67839c47f966
6	Mount Sinai Hospital Jackson Heights	Hospital	37-22 82nd St	NaN	40.748975	-73.884433	[{'label': 'display', 'lat': 40.74897456715504...	803	11372	US	Jackson Heights	NY	United States	[37-22 82nd St, Jackson Heights, NY 11372, Uni...	5c11779f625a66002ceafe1a
7	Elmhurst Hospital Womens Health Center	Hospital	78-20 41st Avenue	NaN	40.745499	-73.887046	[{'label': 'display', 'lat': 40.74549881751365...	1132	11374	US	Elmhurst	NY	United States	[78-20 41st Avenue, Elmhurst, NY 11374, United...	5142fe07e4b03cf303b19298
8	Woodside Anumal Hospital	Hospital	50-10 69th St	NaN	40.736684	-73.897215	[{'label': 'display', 'lat': 40.73668367920025...	2389	11377	US	Woodside	NY	United States	[50-10 69th St, Woodside, NY 11377, United Sta...	4d72ac4ed145a1cdfd8cb8f3

The list contains 9 Hospitals in the Corona/North Corona neighborhood. There is some redundancy in the list and this needs to be cleared. On examining the address location, we can

see that Elmhurst Hospital Center, Elmhurst Hospital Center Mortuary Department and Elmhurst Hospital Center Clinics are one and the same as they have the same address location. We drop the redundant records and use only the Elmhurst Hospital Center. This leaves us with only 7 names with a category of Hospital.

We plot these 7 hospitals on a map.



On doing further research into the number of staffed beds at the hospitals located in the Corona/North Corona neighborhood using various websites as this information is not readily available, we determine that the total number of staffed beds is 1092 beds.

Explore Allerton/Norwood/Pelham Parkway/Williamsbridge Neighborhood in Bronx

Latitude and longitude values of Allerton/Norwood/Pelham Parkway/Williamsbridge are 40.875, -73.875

We construct a URL to send a request to the Foursquare API to search for a specific type of Venue, in this case “Hospital”. We set a LIMIT of 30 and a radius of 3000 metres (approx. 1.86 miles).

```
search_query = 'Hospital'
LIMIT = 30
radius = 3000
url = 'https://api.foursquare.com/v2/venues/search?&client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(
    CLIENT_ID,CLIENT_SECRET,neighborhood_latitude,neighborhood_longitude,VERSION,search_query,radius,LIMIT)
url
```

Coursera Capstone: Battle of the Neighborhoods – New York City

The results are received in the json format and this is stored in results.

```
results = requests.get(url).json()
results
```

The venues are transformed into a pandas dataframe. The data is comprised of any venue with the term “Hospital” in it. This data is cleaned as it includes Hospitals, Dog and Cat Hospital, Hospital Farmer’s Market, Emergency Room, Optical Shop etc. The categories are further filtered on the keyword “Hospital”. The index is reset. The result is as follows:

	name	categories	address	crossStreet	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	id
0	The Children's Hospital at Montefiore	Hospital	3415 Bainbridge Ave	at E Gun Hill Rd	40.880640	-73.879153	[{'label': 'display', 'lat': 40.88063954722417...	718	10467	US	Bronx	NY	United States	[3415 Bainbridge Ave (at E Gun Hill Rd), Bronx...	4c3e307edb3b1b8d56106595
1	North Central Bronx Hospital	Hospital	3224 Kossuth Ave	at E 208th St	40.880441	-73.881526	[{'label': 'display', 'lat': 40.88044133267162...	817	10467	US	Bronx	NY	United States	[3224 Kossuth Ave (at E 208th St), Bronx, NY 1...	4b6de6a8f964a520439a2ce3
2	St. Barnabas Hospital	Hospital	4422 3rd Ave	at E 183th St	40.852741	-73.891168	[{'label': 'display', 'lat': 40.85274125306508...	2827	10457	US	Bronx	NY	United States	[4422 3rd Ave (at E 183th St), Bronx, NY 10457...	4a88980cf964a5200d0720e3
3	Montefiore Hospital Adult medicine practice	Hospital	3444 Kossuth Ave	E. Gunhill Rd.	40.881245	-73.881121	[{'label': 'display', 'lat': 40.881245, 'lng':...	865	10467	US	Bronx	NY	United States	[3444 Kossuth Ave (E. Gunhill Rd.), Bronx, NY ...	50630a42e4b0530cb8d3e6f5
4	NYC Health + Hospitals/Jacobi	Hospital	1400 Pelham Pkwy S	at Eastchester Rd	40.856809	-73.846739	[{'label': 'display', 'lat': 40.85680869123934...	3124	10461	US	Bronx	NY	United States	[1400 Pelham Pkwy S (at Eastchester Rd), Bronx...	4ae0c9e6f964a5207f8221e3
5	Montefiore Medical Center	Hospital	111 E 210th St	NaN	40.879938	-73.880868	[{'label': 'display', 'lat': 40.87993788345708...	738	10467	US	Bronx	NY	United States	[111 E 210th St, Bronx, NY 10467, United States]	4ae06d2df964a520347f21e3
6	St Barnabas Hospital Human Resource Dept	Hospital	NaN	NaN	40.853150	-73.891749	[{'label': 'display', 'lat': 40.85314963490056...	2811	10457	US	Bronx	NY	United States	[Bronx, NY 10457, United States]	4f8c2ce2e4b0e6f34b0afe8c
7	Montefiore Medical Center - Wakefield Campus	Hospital	600 E 233rd St	Carpenter Ave.	40.894207	-73.860790	[{'label': 'display', 'lat': 40.89420660561763...	2449	10466	US	Bronx	NY	United States	[600 E 233rd St (Carpenter Ave.), Bronx, NY 10...	4addedeaf964a520646621e3
8	Bronx VA Medical Center	Hospital	130 W Kingsbridge Rd	Webb Ave	40.867172	-73.905960	[{'label': 'display', 'lat': 40.86717211737055...	2747	10468	US	Bronx	NY	United States	[130 W Kingsbridge Rd (Webb Ave), Bronx, NY 10...	4b0a9041f964a520212523e3

The list contains 9 hospital venues. On examining them, there appears to be some redundancy with similar addresses. These are removed and we calculate the number of staffed beds cumulatively for all the above hospitals by researching the number of staffed beds using various websites as this information is not readily available for each of them. The total number of staffed beds for the Allerton/Norwood/Pelham Parkway/Williamsbridge Neighborhood is 3399 beds.

Using the latitude and longitude values of the neighborhood Allerton/Norwood/Pelham Parkway/Williamsbridge Neighborhood (latitude = 40.875, longitude = -73.875), and the hospitals that service them, we plot the hospitals on a map.

Coursera Capstone: Battle of the Neighborhoods – New York City

Explore Elmhurst Neighborhood in Queens

Latitude and longitude values of Elmhurst neighborhood are
40.75, -73.875
We construct a URL to send a request to the Foursquare API to search for a specific type of Venue, in this case “Hospital”. We set a LIMIT of 30 and a radius of 3000 metres (approx. 1.86 miles).

```
search_query = 'Hospital'
LIMIT = 30
radius = 3000
url = 'https://api.foursquare.com/v2/venues/search?&client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(
    CLIENT_ID,CLIENT_SECRET,neighborhood_latitude,neighborhood_longitude,VERSION,search_query,radius,LIMIT)
url
```

The results are received in the json format and this is stored in results.

```
results = requests.get(url).json()
results
```

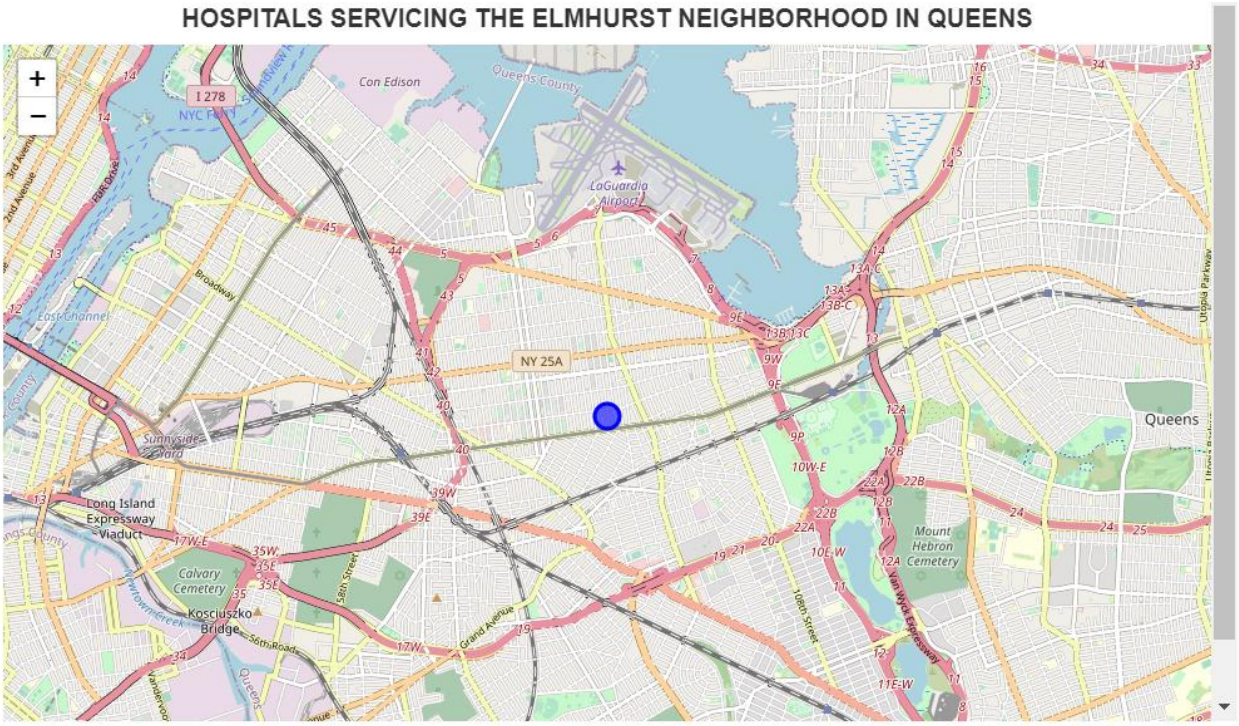
The venues are then transformed into a pandas dataframe. The data is comprised of any venue with the term “Hospital” in it. This data is cleaned as it includes Hospitals, Doctor’s Office, Office etc. The categories are further filtered on the keyword “Hospital”. The index is reset. The result is as follows:

Coursera Capstone: Battle of the Neighborhoods – New York City

	name	categories	address	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	crossStreet	id
0	St. Mary Hospital for Children	Hospital	29-01 216th St	40.776777	-73.768746	[{'label': 'display', 'lat': 40.77677744162408...	3373	11360	US	Bayside	NY	United States	[29-01 216th St, Bayside, NY 11360, United Sta...	NaN	4bb0f5bbf964a520fc6d3ce3
1	Long island Jewish Hospital	Hospital	NaN	40.743799	-73.717957	[{'label': 'display', 'lat': 40.7437994934359,...	2789	NaN	US	Glen Oaks	NY	United States	[Glen Oaks, NY, United States]	NaN	4db6e6b31e7248d135d2417c
2	John T Mather Hospital	Hospital	75 N Country Rd	40.938637	-73.054260	[{'label': 'display', 'lat': 40.93863677978515...	62238	11777	US	Port Jefferson	NY	United States	[75 N Country Rd, Port Jefferson, NY 11777, Un...	NaN	4fd121b2e4b0af969d016a0d
3	LIJ- Oncology	Hospital	NaN	40.737518	-73.715942	[{'label': 'display', 'lat': 40.73751799123268...	3190	11004	US	Glen Oaks	NY	United States	[Glen Oaks, NY 11004, United States]	NaN	50478d7ae4b0a21d2f8b134b

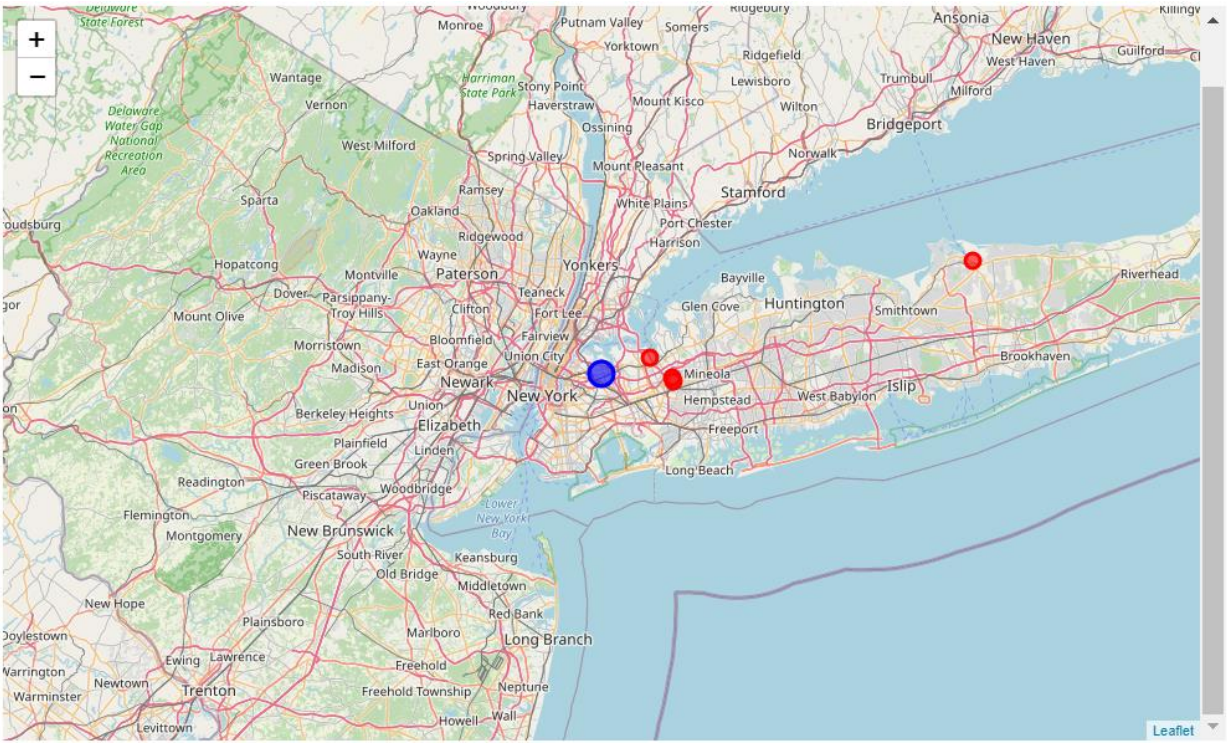
As we can see only 4 hospitals were returned for the Elmhurst neighborhood. Also, only 1 hospital is in the 3000 meter vicinity. The Long Island Jewish Hospital with a bed capacity of 312 beds. The distance for the John T Mather Hospital is 62338 metres from the latitude, longitude location of Elmhurst specified. Using various internet sources, it is determined that the total number of staffed beds for the Elmhurst neighborhood is 707 beds excluding John T Mather Hospital and 955 beds including John T Mather Hospital.

The hospitals in the Elmhurst neighborhood are plotted based on their latitude and longitude values as follows:



As we can see from the above map, none of the 4 hospitals can be seen on the map at the current zoom level of 13. If we zoom out further, we can see that these hospitals are located at a considerable distance from the Elmhurst neighborhood.

Coursera Capstone: Battle of the Neighborhoods – New York City



Explore Allerton/Baychester/Pelham Gardens/Williamsbridge Neighborhood in the Bronx

The latitude of Allerton/Baychester/Pelham Gardens/Williamsbridge is 40.875

The longitude of Allerton/Baychester/Pelham Gardens/Williamsbridge is -73.875

We construct a URL to send a request to the Foursquare API to search for a specific type of Venue, in this case “Hospital”. We set a LIMIT of 30 and a radius of 3000 metres (approx. 1.86 miles).

```
search_query = 'Hospital'
LIMIT = 30
radius = 3000
url = 'https://api.foursquare.com/v2/venues/search?&client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(
    CLIENT_ID,CLIENT_SECRET,neighborhood_latitude,neighborhood_longitude,VERSION,search_query,radius,LIMIT)
url
```

The results are received in the json format and this is stored in results.

```
results = requests.get(url).json()
results
```

The venues are then transformed into a pandas dataframe. The data is comprised of any venue with the term “Hospital” in it. This data is cleaned as it includes Hospitals, Doctor’s Office,

Medical Supply Store, Optical Shop etc. The categories are further filtered on the keyword “Hospital”. The index is reset. The result is as follows:

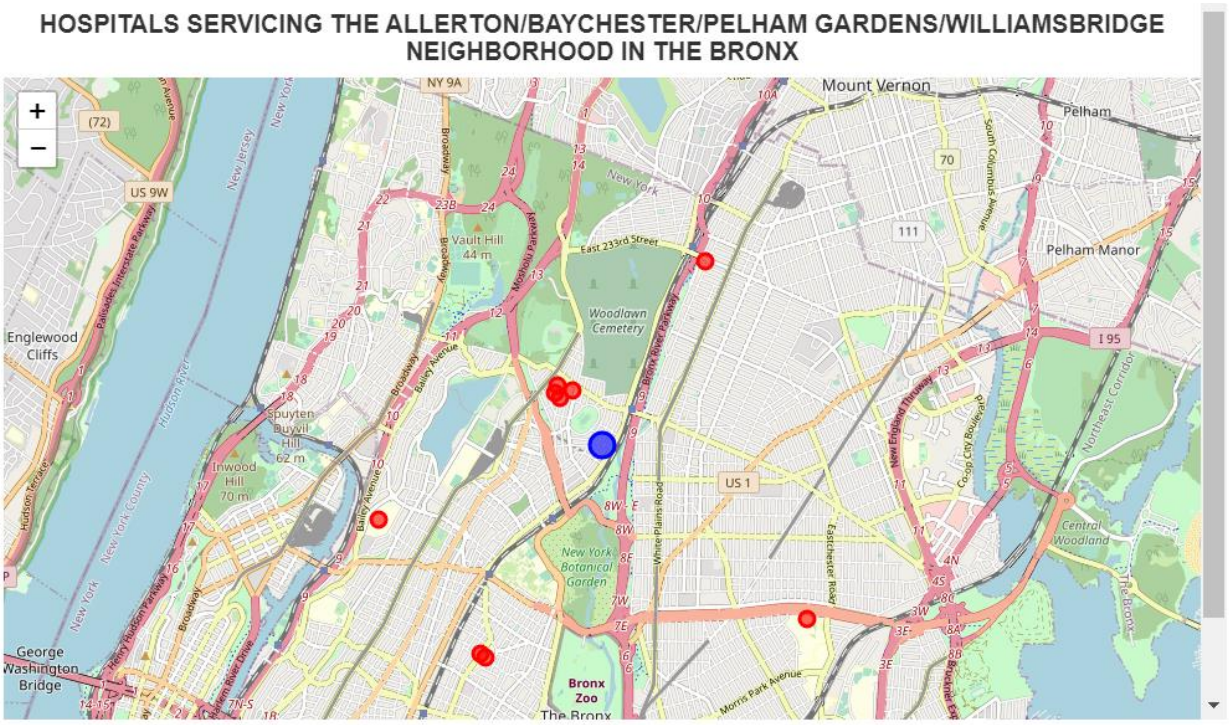
	name	categories	address	crossStreet	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	id
0	The Children's Hospital at Montefiore	Hospital	3415 Bainbridge Ave	at E Gun Hill Rd	40.880640	-73.879153	[{'label': 'display', 'lat': 40.88063954722417...	718	10467	US	Bronx	NY	United States	[3415 Bainbridge Ave (at E Gun Hill Rd), Bronx...	4c3e307edb3b1b8d56106595
1	North Central Bronx Hospital	Hospital	3224 Kossuth Ave	at E 208th St	40.880441	-73.881526	[{'label': 'display', 'lat': 40.88044133267162...	817	10467	US	Bronx	NY	United States	[3224 Kossuth Ave (at E 208th St), Bronx, NY 1...	4b6de6a8f964a520439a2ce3
2	St. Barnabas Hospital	Hospital	4422 3rd Ave	at E 183th St	40.852741	-73.891168	[{'label': 'display', 'lat': 40.85274125306508...	2827	10457	US	Bronx	NY	United States	[4422 3rd Ave (at E 183th St), Bronx, NY 10457...	4a88980cf964a5200d0720e3
3	Adolescent AIDS program At The Children's Hosp...	Hospital	NaN	NaN	40.881056	-73.877279	[{'label': 'display', 'lat': 40.88105598218058...	700	10467	US	Bronx	NY	United States	[Bronx, NY 10467, United States]	4f4d00dfe4b082dc5b340456
4	Children's Hospital At Montefiore	Hospital	NaN	NaN	40.880391	-73.878600	[{'label': 'display', 'lat': 40.88039137785859...	672	10467	US	Bronx	NY	United States	[Bronx, NY 10467, United States]	50e6d8ffe4b075f63dd58289
5	Montefiore Hospital Adult medicine practice	Hospital	3444 Kossuth Ave	E. Gunhill Rd.	40.881245	-73.881121	[{'label': 'display', 'lat': 40.881245, 'lng':...	865	10467	US	Bronx	NY	United States	[3444 Kossuth Ave (E. Gunhill Rd.), Bronx, NY ...	50630a42e4b0530cb8d3e6f5
6	NYC Health + Hospitals/Jacobi	Hospital	1400 Pelham Pkwy S	at Eastchester Rd	40.856809	-73.846739	[{'label': 'display', 'lat': 40.85680869123934...	3124	10461	US	Bronx	NY	United States	[1400 Pelham Pkwy S (at Eastchester Rd), Bronx...	4ae0c9e6f964a5207f8221e3
7	Montefiore Medical Center	Hospital	111 E 210th St	NaN	40.879938	-73.880868	[{'label': 'display', 'lat': 40.87993788345708...	738	10467	US	Bronx	NY	United States	[111 E 210th St, Bronx, NY 10467, United States]	4ae06d2df964a520347f21e3
8	St Barnabas Hospital Human Resource Dept	Hospital	NaN	NaN	40.853150	-73.891749	[{'label': 'display', 'lat': 40.85314963490056...	2811	10457	US	Bronx	NY	United States	[Bronx, NY 10457, United States]	4f8c2ce2e4b0e6f34b0afe8c
9	Montefiore Medical Center - Wakefield Campus	Hospital	600 E 233rd St	Carpenter Ave.	40.894207	-73.860790	[{'label': 'display', 'lat': 40.89420660561763...	2449	10466	US	Bronx	NY	United States	[600 E 233rd St (Carpenter Ave.), Bronx, NY 10...	4addedeaf964a520646621e3
10	Bronx VA Medical Center	Hospital	130 W Kingsbridge Rd	Webb Ave	40.867172	-73.905960	[{'label': 'display', 'lat': 40.86717211737055...	2747	10468	US	Bronx	NY	United States	[130 W Kingsbridge Rd (Webb Ave), Bronx, NY 10...	4b0a9041f964a520212523e3

We can see that 11 hospital venues were found. There are some redundancies. These rows are dropped.

Using various sources, we determine that the total number of staffed beds for the Allerton/Baychester/Pelham Gardens/Williamsbridge neighborhood is 3399 beds. This neighborhood is the same as Allerton/Norwood/Pelham Parkway/Williamsbridge neighborhood. It is also being serviced by the same hospitals as Allerton/Norwood/Pelham Parkway/Williamsbridge neighborhood.

The hospitals are plotted based on their latitude and longitude values in relation to the Allerton/Baychester/Pelham Gardens/Williamsbridge neighborhood latitude and longitude values.

Coursera Capstone: Battle of the Neighborhoods – New York City



Explore Jackson Heights Neighborhood in Queens

Latitude and longitude values of Jackson Heights are 40.75, -73.875.
We construct a URL to send a request to the Foursquare API to search for a specific type of Venue, in this case “Hospital”. We set a LIMIT of 30 and a radius of 3000 metres (approx. 1.86 miles).

```
search_query = 'Hospital'
LIMIT = 30
radius = 3000
url = 'https://api.foursquare.com/v2/venues/search?&client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(
    CLIENT_ID,CLIENT_SECRET,neighborhood_latitude,neighborhood_longitude,VERSION,search_query,radius,LIMIT)
url
```

The results are received in the json format and this is stored in results.

```
results = requests.get(url).json()
results
```

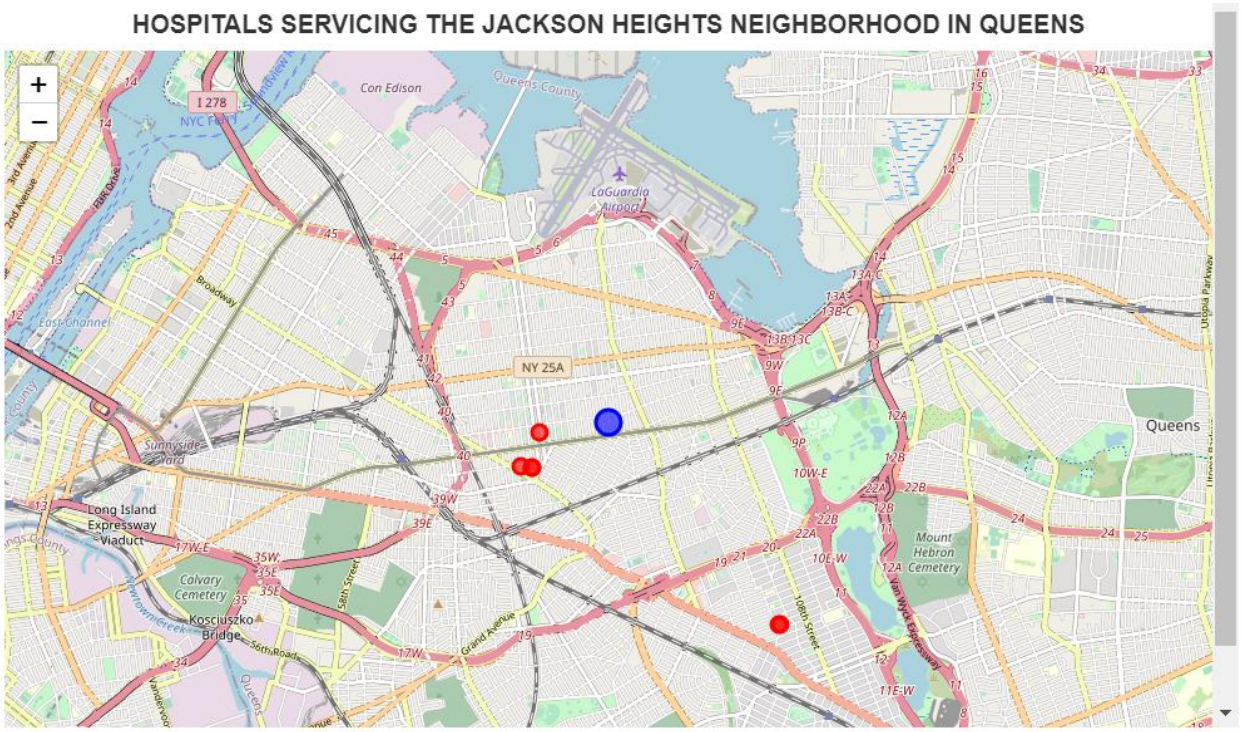
The venues are then transformed into a pandas dataframe. The data is comprised of any venue with the term “Hospital” in it. This data is cleaned as it includes Hospitals, Emergency Room, Gift Shop, Food Truck etc. The categories are further filtered on the keyword “Hospital”. The index is reset. The result is as follows:

Coursera Capstone: Battle of the Neighborhoods – New York City

	name	categories	address	crossStreet	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	id
0	Elmhurst Hospital Center	Hospital	79-01 Broadway	at 78th St.	40.745244	-73.885536	[{'label': 'display', 'lat': 40.74524449809677...	1034	11373	US	Elmhurst	NY	United States	[79-01 Broadway (at 78th St.), Elmhurst, NY 11...	4be0708998f2a593de40c25a
1	Elmhurst Hospital Senior Care	Hospital	41 Ave	NaN	40.745266	-73.885393	[{'label': 'display', 'lat': 40.74526551814158...	1022	11373	US	Elmhurst	NY	United States	[41 Ave, Elmhurst, NY 11373, United States]	4f392dc9e4b0653d4c81946b
2	Long Island Jewish Forest Hills Hospital	Hospital	102-01 66th Rd	North Shore-LIJ Health System	40.729062	-73.851625	[{'label': 'display', 'lat': 40.7290618, 'lng'...	3052	11375	US	Forest Hills	NY	United States	[102-01 66th Rd (North Shore-LIJ Health System...	4b7d6f3df964a52009be2fe3
3	Forest Hills Hospital - Nursery	Hospital	102-01 66th Rd	NaN	40.729031	-73.851509	[{'label': 'display', 'lat': 40.72903060913086...	3061	11375	US	Forest Hills	NY	United States	[102-01 66th Rd, Forest Hills, NY 11375, Unite...	4dd18a75d22d67839c47f966
4	Mount Sinai Hospital Jackson Heights	Hospital	37-22 82nd St	NaN	40.748975	-73.884433	[{'label': 'display', 'lat': 40.74897456715504...	803	11372	US	Jackson Heights	NY	United States	[37-22 82nd St, Jackson Heights, NY 11372, Uni...	5c11779f625a66002ceafe1a
5	Elmhurst Hospital Womens Health Center	Hospital	78-20 41st Avenue	NaN	40.745499	-73.887046	[{'label': 'display', 'lat': 40.7454988175136 5...	1132	11374	US	Elmhurst	NY	United States	[78-20 41st Avenue, Elmhurst, NY 11374, United...	5142fe07e4b03cf303b19298

There are a few redundant venues based on looking at the address. We eliminate them. Based on research from various internet sources of the various hospitals in the above list, we determine that the number of staffed beds for the Jackson Heights neighborhood is 1092 beds.

We plot the hospitals on a map relative to their location from the neighborhood using the latitude and longitude values as shown below.



Coursera Capstone: Battle of the Neighborhoods – New York City

Results and Discussion

We analyzed the New York City population data by boroughs and neighborhoods and found that the Queens borough is the second most populous borough after Manhattan. It comprises a total of 2,330,295 residents with an at-risk population of 325,300 residents over the age of 65 and 341,062 school age children.

We then looked at the populations of the neighborhoods to determine which neighborhoods were the most populous. We found that the top 25 neighborhoods were in the 3 boroughs of Manhattan, Queens, and Brooklyn.

We used K-Means clustering to cluster the neighborhoods based on Covid-19 Case Counts, Death Counts, and Percent Positive into 3 different clusters and categorized them as High risk, Medium Risk, and Low risk clusters.

We then analyzed the neighborhood tabulations of Covid-19 Case Counts, Death Counts, and Percent Positive cases separately. We found that some neighborhoods were on all 3 Top 25 lists with high Case Counts, high Death counts, and high Percent Positive while other neighborhoods were on one or two lists but not on all 3. We narrowed down our focus on 5 neighborhoods that appeared in the top 25 on all three lists. The neighborhood of Corona/North Corona appeared in the number 1 spot on all three lists. We can see that 3 of these neighborhoods are in Queens and 2 in the Bronx.

NEIGHBORHOOD	BOROUGH	COVID_CASE_COUNT	COVID_DEATH_COUNT	PERCENT_POSITIVE
Corona/North Corona	Queens	4534	395	34.57
Allerton/Norwood/Pelham Parkway/Williamsbridge	Bronx	3379	295	30.48
Elmhurst	Queens	3218	285	30.55
Allerton/Baychester/Pelham Gardens/Williamsbridge	Bronx	2991	322	30.69
Jackson Heights	Queens	2513	231	31.42

Next, we analyzed the locations of the hospitals within the boroughs and found that Queens has far fewer hospitals only 11 hospitals when compared with other boroughs - Manhattan which has 20 hospitals, Brooklyn which has 16 hospitals, and the Bronx which has 13 hospitals.

Using Foursquare API, we analyzed each of these 5 neighborhoods to see how many hospitals were in their vicinity. We also tabulated the number of staffed hospital beds. We found that Allerton/Northwood/Pelham Parkway/Williamsbridge and Allerton/Baychester/Pelham Gardens/Williamsbridge were in close proximity and were serviced by the same hospitals with 3399 staffed beds. The Jackson Heights neighborhood in Queens was being serviced mainly by 3 hospitals with a total number of 1092 hospital staffed beds. The Corona/North Corona neighborhood in Queens was being serviced by 7 hospitals with a total staffed bed capacity of 1032. The Elmhurst neighborhood was being serviced by only 1 hospital in the 3000 meter

radius. The Long Island Jewish Hospital with a total staffed bed capacity of 583 beds. If we include three other hospitals that showed up in the venue search, the total bed capacity for Elmhurst is 707 beds excluding John T Mather Hospital and 955 beds including John T Mather Hospital.

Based on our analysis of the above data points we can tell that the Elmhurst neighborhood in the Queens borough of New York City is being underserviced in terms of hospitals and staffed hospital beds. The Elmhurst neighborhood appears at:

- number 8 on the list of neighborhoods with the highest population with a population of 88,437
- number 3 on the list of highest Covid Case counts with 3221 cases,
- number 5 on the list with the highest Death counts with 285 deaths
- number 21 on the highest percent positive list with 30.64% testing positive,

When compared with the other 4 similar neighborhoods, it does seem like there is a need for another hospital in the Elmhurst neighborhood.

Conclusion

The purpose of this project was to propose a new location in New York City to add a hospital based on the current COVID-19 cases and to prepare better for future pandemics. After examining all the data points explained above, we can conclude that the Queens borough of New York city is not being adequately serviced in terms of staffed hospital beds. Also, specifically the Elmhurst neighborhood in the Queens borough despite having a high population, a high Covid-19 case count, and a high number of deaths due to Covid-19 is lacking in the number of hospitals nearby as well as the number of staffed hospital beds. This neighborhood and its residents will be well-served if a new hospital was added to serve and care for its residents. In addition to the data and analysis above, it may be a good idea to explore other factors like size, affordability and cost, availability of space, location, accessibility, safety and security etc. while determining where to add a new hospital in New York City.

Coursera Capstone: Battle of the Neighborhoods – New York City

References

The following websites were used to gather information and data for this project:

https://www.iloveny.com/places-to-go/new-york-city/

https://www.usapopulation.org/new-york-city-population/

https://data.cityofnewyork.us/City-Government/2020-population/t8c6-3i7b

https://projects.thecity.nyc/2020_03_covid-19-tracker/

https://www.usnews.com/news/health-news/articles/2020-03-27/new-york-is-establishing-more-temporary-hospitals-to-handle-surge-of-coronavirus-patients

https://en.wikipedia.org/wiki/COVID-19_pandemic_in_New_York_(state)

https://www.bloomberg.com/news/articles/2020-05-28/why-was-new-york-hit-so-badly-with-covid-19

https://www.newyorkled.com/maps/export/geojson/10/

Population data

https://data.cityofnewyork.us/City-Government/2020-population/t8c6-3i7b

https://data.cityofnewyork.us/City-Government/New-York-City-Population-By-Neighborhood-Tabulatio/swpk-hqdp/data

https://www.ahd.com/states/hospital_NY.html

https://www.bloomberg.com/graphics/2020-new-york-coronavirus-outbreak-how-many-hospital-beds/

https://www.google.com/search

The following is a address on github to the Jupyter notebook containing the programming logic and code for this project:

https://github.com/nitarahuja/Coursera_Capstone/blob/master/NYC%20Population%20Breakdown%20Borough.ipynb