Applied Data Science Capstone by IBM/Coursera

Avani Bhatnagar

August 11, 2021

Business Problem

- Location is very important when it comes to opening a restaurant. If a family wants to open up a restaurant in a city, they need to know different things about the location to be able to strategically select the correct neighborhood which will attract the most crowd, is affordable to open up a restaurant at, has lesser competition, population etc.

Data Acquisition and Cleaning

- New York Data: We use this dataset: https://en.wikipedia.org/wiki/New_York_City to get information about the various New York neighborhoods which will help us explore and decide the best location for an Indian restaurant.

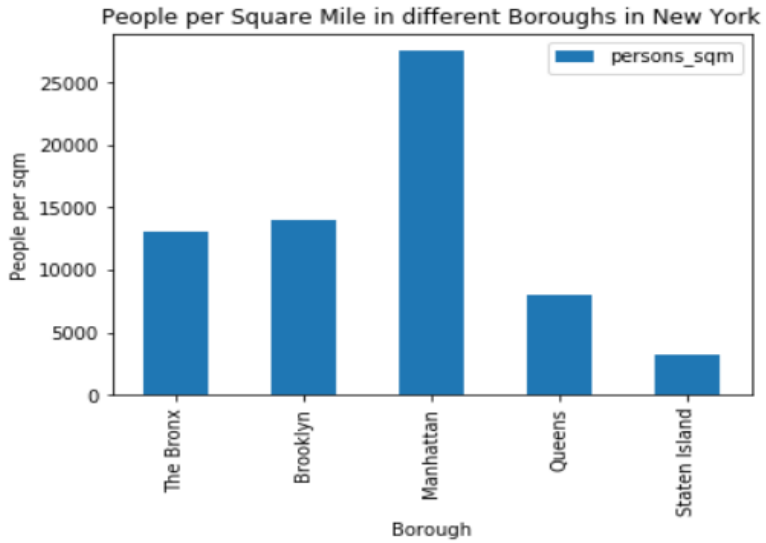
- Foursquare API: We use Foursquare API to get various restaurants in New York such as Indian Cuisine. This will help the business get an idea of what's the best location to open their restaurant in terms on relative competition in the same business.

- Geospatial data - We use https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm to get the location details of New York - it gives us the Latitude and Longitude values for each Borough in New York. This will be greatly helpful in generating maps and visualizing the overall results.

- Demographic data - We use https://en.wikipedia.org/wiki/Demographics_of_New_York_City to get the demographic data for NYC's jurisdiction.

Finding the best restaurant

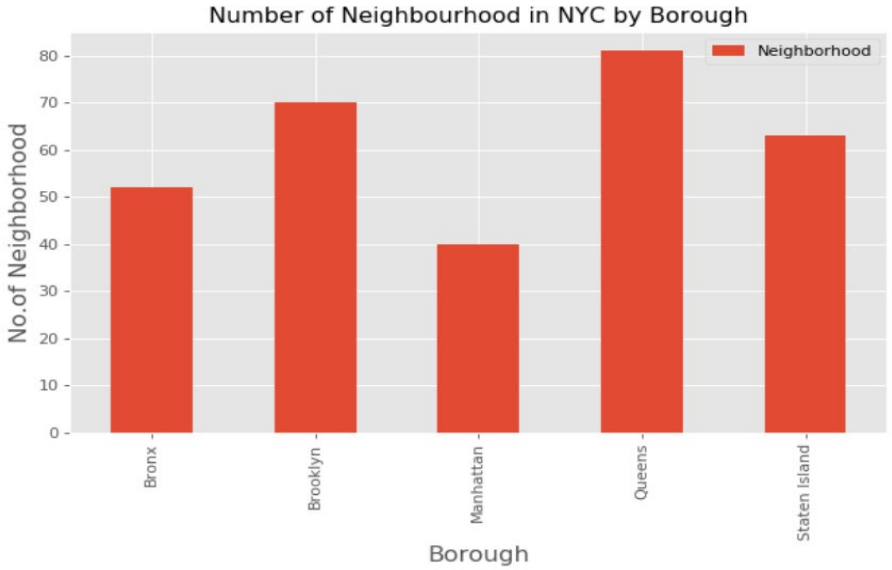
	borough	county	population	gdp_per_capita	persons_sqm
0	The Bronx	Bronx	1418207.0	42.10	13006.0
1	Brooklyn	Kings	2559903.0	70.82	13957.0
2	Manhattan	New York	1628706.0	22.83	27544.0
3	Queens	Queens	2253858.0	108.53	8018.0
4	Staten Island	Richmond	476143.0	58.37	3150.0

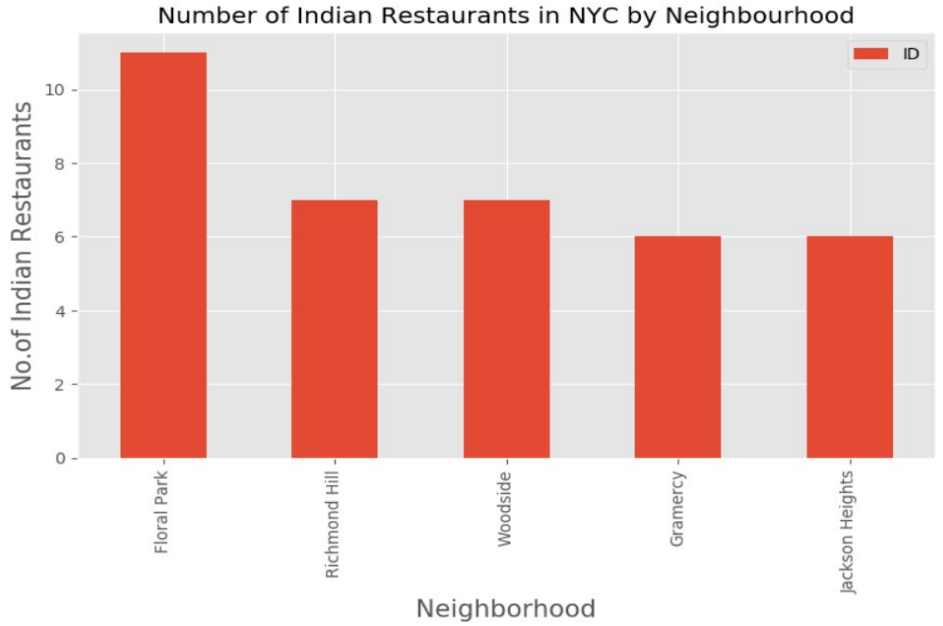


	jurisdiction	population_census	White	African_Amercian	Asian	Other	Mixed_Race	Hispanic_Latino_of_other_race
0	Queens	2,229,379	44.1	20.0	17.6	12.3	6.1	NaN
1	Manhattan	1,537,195	54.4	17.4	9.4	14.7	4.1	NaN
2	Bronx	1,332,650	29.9	35.6	3.0	25.7	5.8	NaN
3	Staten Island	443,728	77.6	9.7	5.7	4.3	2.7	NaN
4	NYC Total	8,008,278	44.7	26.6	9.8	14.0	4.9	NaN

We get all the boroughs of New York City with their population, gdp per capita and person sqm.

Finding the best restaurant





We use the Foursquare API to get the number of neighborhoods in NYC by Borough and Number of Indian restaurants in NYC by neighborhood

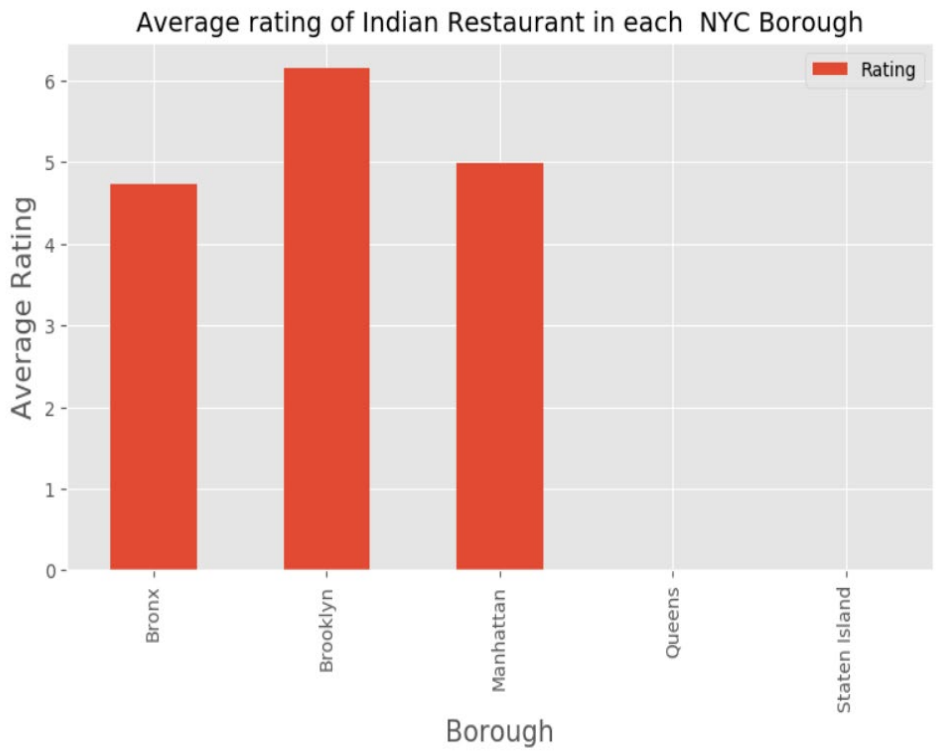
Finding the best restaurant

We found neighborhoods with the highest ratings

```
[81]: ny_neighborhood_stats.sort_values(['Average Rating'],ascending=False).head(10)
```

Out[81]:

	Neighborhood	Average Rating
69	Tribeca	9.10
52	Prospect Heights	9.00
31	Greenwich Village	9.00
74	West Village	8.85
20	East Village	8.70
26	Fort Greene	8.70
11	Chelsea	8.70
44	Midtown	8.70
13	Clinton Hill	8.70
53	Prospect Lefferts Gardens	8.70

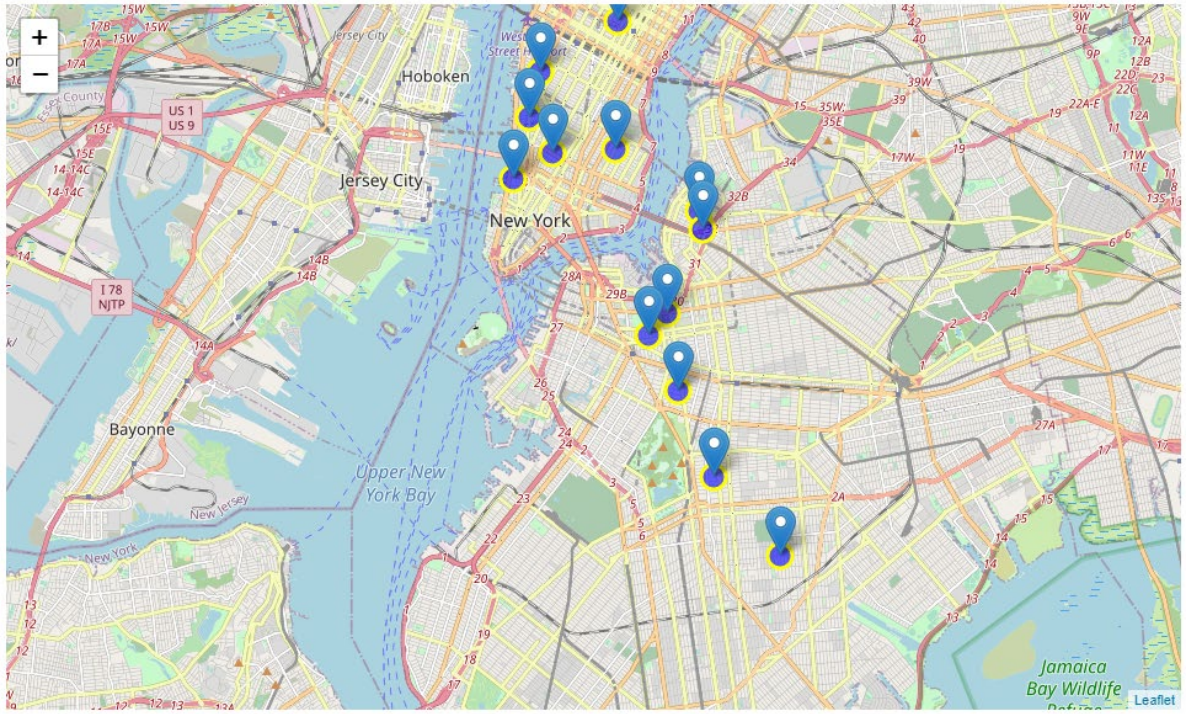


Finding in the restaurant

Neighborhoods with highest ratings

In [91]:
```
# add pop-up text to each marker on the map
for lat, lng, label in ny_neighborhood_stats[['Latitude','Longitude','Label']].values:
    folium.Marker([lat, lng], popup=label).add_to(ny_map)
# add incidents to map
ny_map.add_child(incidents)
```

Out[91]:



Neighborhoods with highest ratings in NYC visualized.