Capstone Project

Analysis of Indian Restaurants in New York City

Introduction:

We can see from the New York City's demographic data that it is very ethnically diverse. It is also one of the largest city in the US. The city's population in 2010 was 44% white (33.3% non-Hispanic white), 25.5% black (23% non-Hispanic black), 0.7% Native American, and 12.7% Asian. Hispanics of any race represented 28.6% of the population, while Asians constituted the fastest-growing segment of the city's population between 2000 and 2010. Throughout its history, New York has been a major port of entry for immigrants into the United States. More than 12 million European immigrants were received at Ellis Island between 1892 and 1924. The term "melting pot" was first coined to describe densely populated immigrant neighbourhoods on the Lower East Side. By 1900, Germans constituted the largest immigrant group, followed by the Irish, Jews, and Italians. In 1940, whites represented 92% of the city's population.

With this huge diversity in population there is also is a huge diversity in food. There are many different restaurants which provide many different types of cuisine like Chinese, Indian, Italian, etc. Since, Indian cuisine is one of my favourite I have decided to use data regarding Indian restaurants using data obtained from foursquare.com APIs and form the capstone project around that idea.

Problem:

We will try to deal with the following problems in this project:

1. Find areas which have restaurants which serve Indian cuisine
2. Find best locations for Indian cuisine
3. Areas which lack good Indian restaurants
4. Visualize best locations for Indian cuisine on a map

Data Section:

- New York data containing Boroughs, Neighbourhoods along with latitude and longitude

  Source: https://cocl.us/new_york_dataset

  Description: This data contains various information about the neighbourhoods of New York City which is needed for this project.

- Indian restaurants in each neighbourhood of New York City

  Source: Foursquare API

  Description: Using this API we can get information regarding Indian restaurants (venues) in each neighbourhood along with data like ratings, tips, etc.

- GeoSpace data

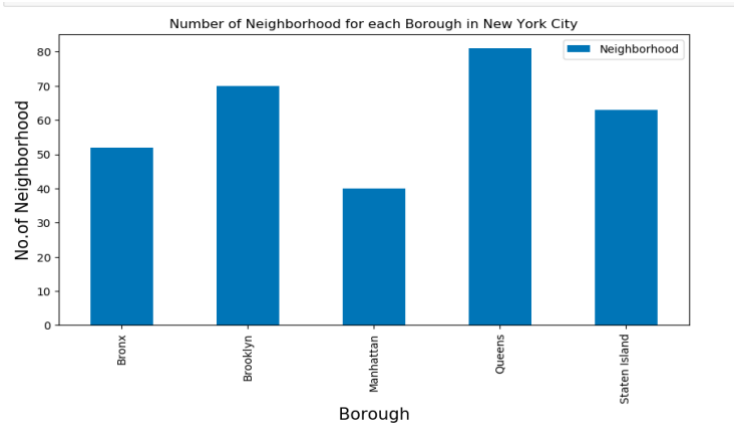
  Source: : https://data.cityofnewyork.us/City-Government/BoroughBoundaries/tqmj-j8zm

  Description: We can use this geospacer data to help visualize the various boroughs boundaries on the map.

Methodology:

1. The first step is obviously to collect the New York City data, which we collected through the link 'https://cocl.us/new_york_dataset '
2. We find all the neighbourhoods and boroughs along with their longitude and latitude from the dataset.

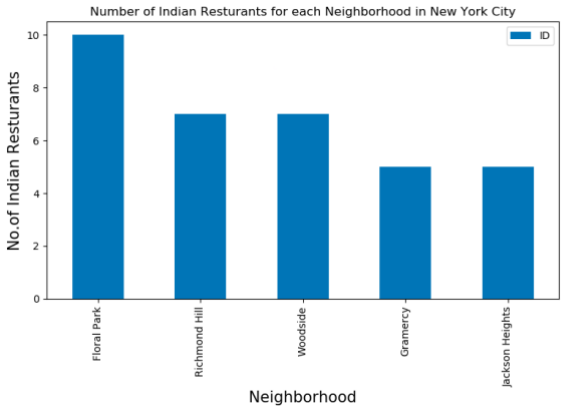
	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585



3. We then fetch data regarding the Indian restaurants from the Foursquare API for the neighbourhoods we have identified.

```
( 1 / 306 ) Indian Resturants in Wakefield, Bronx:0
( 2 / 306 ) Indian Resturants in Co-op City, Bronx:0
( 3 / 306 ) Indian Resturants in Eastchester, Bronx:0
( 4 / 306 ) Indian Resturants in Fieldston, Bronx:0
( 5 / 306 ) Indian Resturants in Riverdale, Bronx:0
( 6 / 306 ) Indian Resturants in Kingsbridge, Bronx:0
( 7 / 306 ) Indian Resturants in Marble Hill, Manhattan:0
( 8 / 306 ) Indian Resturants in Woodlawn, Bronx:1
( 9 / 306 ) Indian Resturants in Norwood, Bronx:0
( 10 / 306 ) Indian Resturants in Williamsbridge, Bronx:0
( 11 / 306 ) Indian Resturants in Baychester, Bronx:0
( 12 / 306 ) Indian Resturants in Pelham Parkway, Bronx:0
( 13 / 306 ) Indian Resturants in City Island, Bronx:0
( 14 / 306 ) Indian Resturants in Bedford Park, Bronx:0
( 15 / 306 ) Indian Resturants in University Heights, Bronx:0
( 16 / 306 ) Indian Resturants in Morris Heights, Bronx:0
( 17 / 306 ) Indian Resturants in Fordham, Bronx:0
( 18 / 306 ) Indian Resturants in East Tremont, Bronx:0
( 19 / 306 ) Indian Resturants in West Farms, Bronx:0
```

4. Next we find the number of Indian restaurants in each borough and neighbourhood.



5. Again using the Foursquare API we will find all the relevant data regarding each of the restaurants like Ratings, Tips, etc.

	Borough	Neighborhood	ID	Name	Likes	Rating	Tips
0	Bronx	Woodlawn	4c0448d9310fc9b6bf1dc761	Curry Spot	5.0	7.8	10.0
1	Bronx	Parkchester	4c194631838020a13e78e561	Melanies Roti Bar And Grill	3.0	6.0	2.0
2	Bronx	Spuyten Duyvil	4c04544df423a593ac83d116	Cumin Indian Cuisine	13.0	6.0	9.0
3	Bronx	Concourse	551b7f75498e86c00a0ed2e1	Hungry Bird	8.0	6.8	3.0
4	Bronx	Unionport	4c194631838020a13e78e561	Melanies Roti Bar And Grill	3.0	6.0	2.0

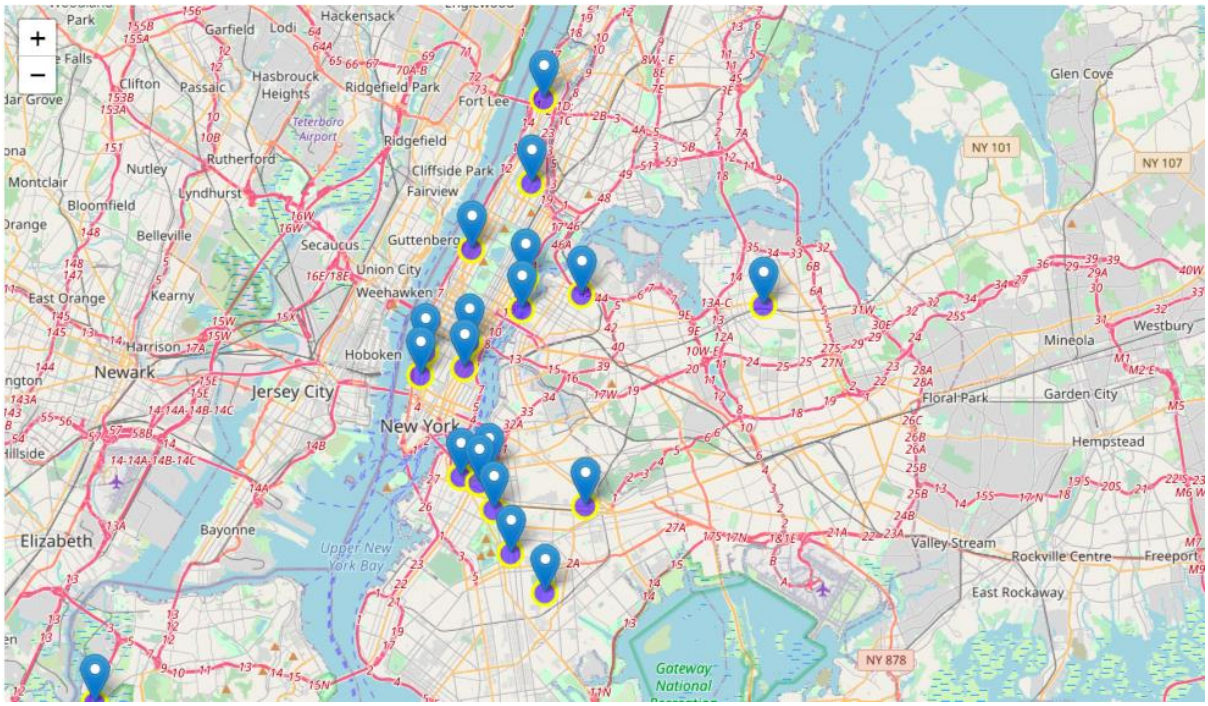
6. Using average ratings as a constraint, we can sort neighbourhoods and boroughs.

	Neighborhood	Average Rating
0	Astoria	9.000000
4	Chelsea	8.800000
5	Clinton Hill	8.800000
11	Fort Greene	8.800000
32	West Village	8.700000
25	Prospect Lefferts Gardens	8.700000
24	Prospect Heights	8.700000
21	Murray Hill	8.600000
30	Upper West Side	8.500000
27	Roosevelt Island	8.500000
14	Gramercy	8.400000

	Borough	Average Rating
3	Queens	9.000000
2	Manhattan	8.116000
1	Brooklyn	7.836842
0	Bronx	6.520000

7. We join this dataset with the latitude and longitude database, and visualize it on the map using python's Folium library.

	Borough	Neighborhood	Latitude	Longitude	Average Rating
0	Queens	Astoria	40.768509	-73.915654	9.0
1	Manhattan	Central Harlem	40.815976	-73.943211	8.0
2	Manhattan	Chelsea	40.744035	-74.003116	8.8
3	Staten Island	Chelsea	40.594726	-74.189560	8.8
4	Brooklyn	Clinton Hill	40.693229	-73.967843	8.8
5	Brooklyn	Downtown	40.690844	-73.983463	8.0
6	Brooklyn	East Flatbush	40.641718	-73.936103	8.3
7	Brooklyn	Fort Greene	40.688527	-73.972906	8.8
8	Manhattan	Gramercy	40.737210	-73.981376	8.4
9	Manhattan	Murray Hill	40.748303	-73.978332	8.6
10	Queens	Murray Hill	40.764126	-73.812763	8.6
11	Brooklyn	Ocean Hill	40.678403	-73.913068	8.0
12	Brooklyn	Prospect Heights	40.676822	-73.964859	8.7
13	Brooklyn	Prospect Lefferts Gardens	40.658420	-73.954899	8.7
14	Manhattan	Roosevelt Island	40.762160	-73.949168	8.5
15	Manhattan	Upper West Side	40.787658	-73.977059	8.5
16	Manhattan	Washington Heights	40.851903	-73.936900	8.0
17	Manhattan	West Village	40.734434	-74.006180	8.7
18	Manhattan	Yorkville	40.775930	-73.947118	8.2



Results:

From our analysis the answers to the above questions are:

1. We found areas which have Indian restaurants. It is detailed above in the notebook

	Borough	Neighborhood	ID	Name
0	Bronx	Woodlawn	4c0448d9310fc9b6bf1dc761	Curry Spot
1	Bronx	Parkchester	4c194631838020a13e78e561	Melanies Roti Bar And Grill
2	Bronx	Spuyten Duyvil	4c04544df423a593ac83d116	Cumin Indian Cuisine
3	Bronx	Concourse	551b7f75498e86c00a0ed2e1	Hungry Bird
4	Bronx	Unionport	4c194631838020a13e78e561	Melanies Roti Bar And Grill

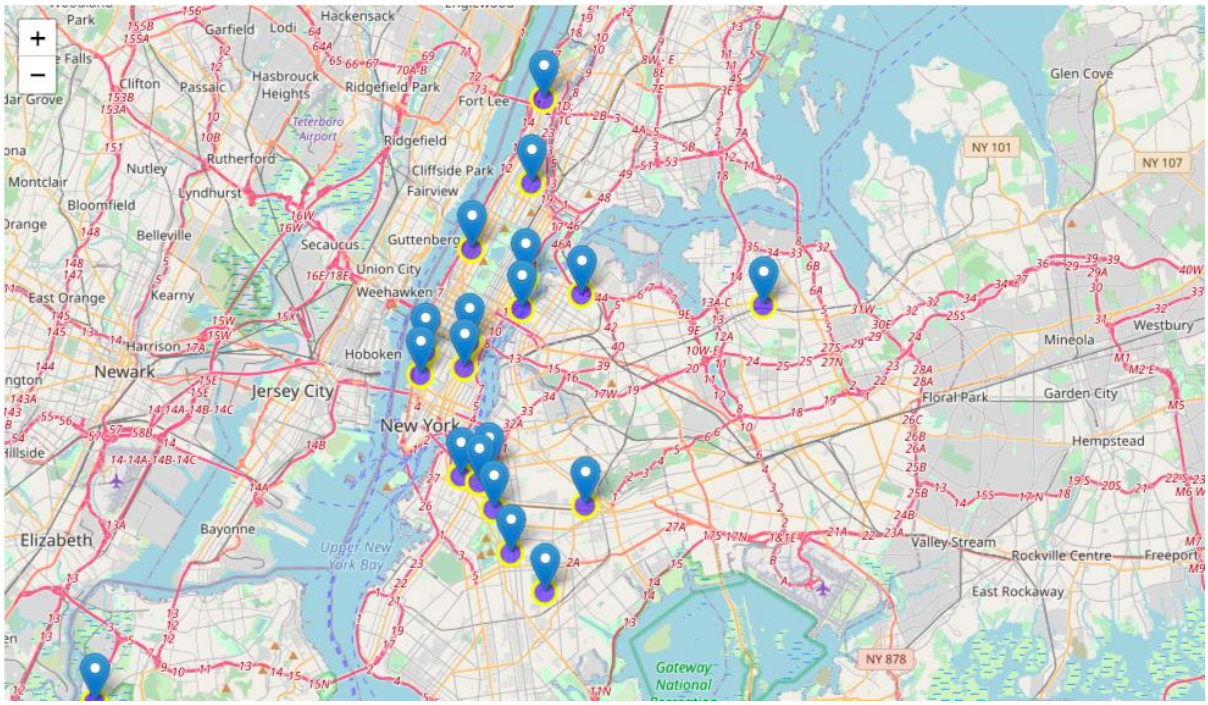
2. Astoria(Queens), Fort Greene(Brooklyn), Clinton Hill(Brooklyn) are some of the best neighbourhoods for Indian cuisine.

	Neighborhood	Average Rating
0	Astoria	9.000000
4	Chelsea	8.800000
5	Clinton Hill	8.800000
11	Fort Greene	8.800000
32	West Village	8.700000
25	Prospect Lefferts Gardens	8.700000
24	Prospect Heights	8.700000
21	Murray Hill	8.600000

3. Bronx has least average rating of any boroughs. Also in our data Union port has the least average rating.

	Borough	Average Rating
3	Queens	9.000000
2	Manhattan	8.116000
1	Brooklyn	7.836842
0	Bronx	6.520000

4. As you can see in the previous cell we have visualized the best locations to get Indian cuisine.



Discussions:

There are various conclusions which we can draw from the analysis of the data we have obtained after carrying out this study.

As we can see, Queens has the most number of restaurants with the highest average rating which can lead to the conclusion that the market is already saturated with good restaurants and a large number of them to boot.

But Manhattan, has quite few restaurants when compared to Queens with nearly a same rating which might suggest that there can be room to grow for a new business as not as many restaurants are present. Also the Bronx, has the lowest overall rating so it might suggest that a good quality restaurant might be able to succeed in this area.

Ultimately, the aim of the data gathered and the results obtained is to make the decision making process more data driven. The results obtained here can help with this task.

Conclusion:

There is always room for improvement in whatever we try to achieve. This project has compiled data regarding Indian restaurants in the New York City area, and we have provided the information in a clear, concise and visualized manner so that it can be used for any application that requires this data, from a potential owner of a new restaurant looking for the best spot to open a new restaurant or a restaurant chain trying to open a new franchise. This data can be utilized to help make the decision of the businesses more information driven which is always a benefit.