Coursera Capstone Project
IBM Applied Data Science Capstone
Analysing to open a Indian Restaurant in New York City

Data Science Presentation

By Vasanthan Veerasingham
April 2020

Introduction

Background
This data science report explores the best locations for Indian restaurants throughout the city of New York. Potentially the owner of the new Indian restaurant can have great success and consistent profit.

Problem
we have to find the solution for a real business problem. Which neighborhood place is best to open a Indian restaurant in New York? And Which Indian restaurant their favorite now?

Background
This data science report is particularly useful to developers and investors looking to open or invest in a Indian restaurant in the city of New York. In consequence, as a reader of this data science report, common people of New York can find where they have Indian restaurants to eat healthy foods for better life.

2. Data Acquisition and Cleaning

Data Acquisition

The data acquired for this project is a combination of data from three sources.

- ❖ Open source dataset of New York: https ://cocl.us/new_york_dataset

- ❖ Foursquare API for get the venue's information

- ❖ Geocoder package for get the latitude and longitude of venues

Full Dataset get from sources

Out[5]:

|	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

Initial Dataset get from sources

Dimension of Data
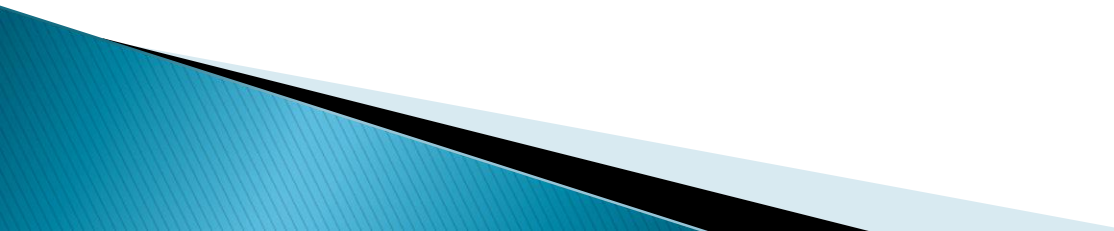
In [6]: `new_york_data.shape`

Out[6]: (306, 4)

There are total of 306 different Neighborhoods in New York.

Neighborhoods per Borough: NYC
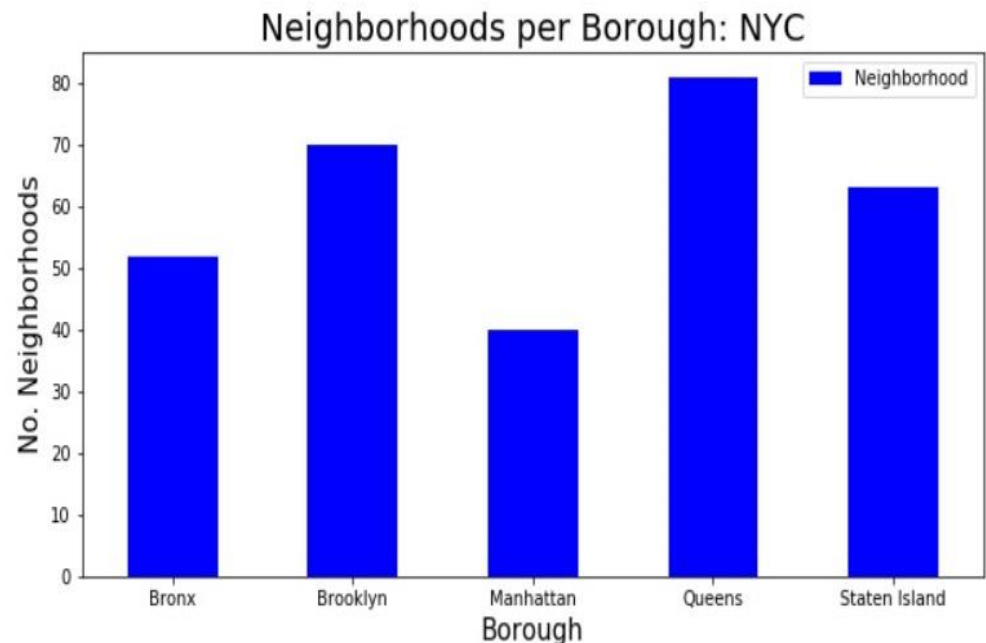
3. Methodology / Data wrangling

❖ Collect the data from https://cocl.us/new_york_dataset. clean and process the data into a dataframe.

❖ Using FourSquare to locate all venues and then filter by Indian restaurants. Ratings, tips, and likes given by users will be counted and added to the dataframe

❖ Sort the data based on rankings, likes and tips

❖ Visually analysing the datasets using the dataframe tables get by using Python libraries.

Results and Findings

Out[14]:

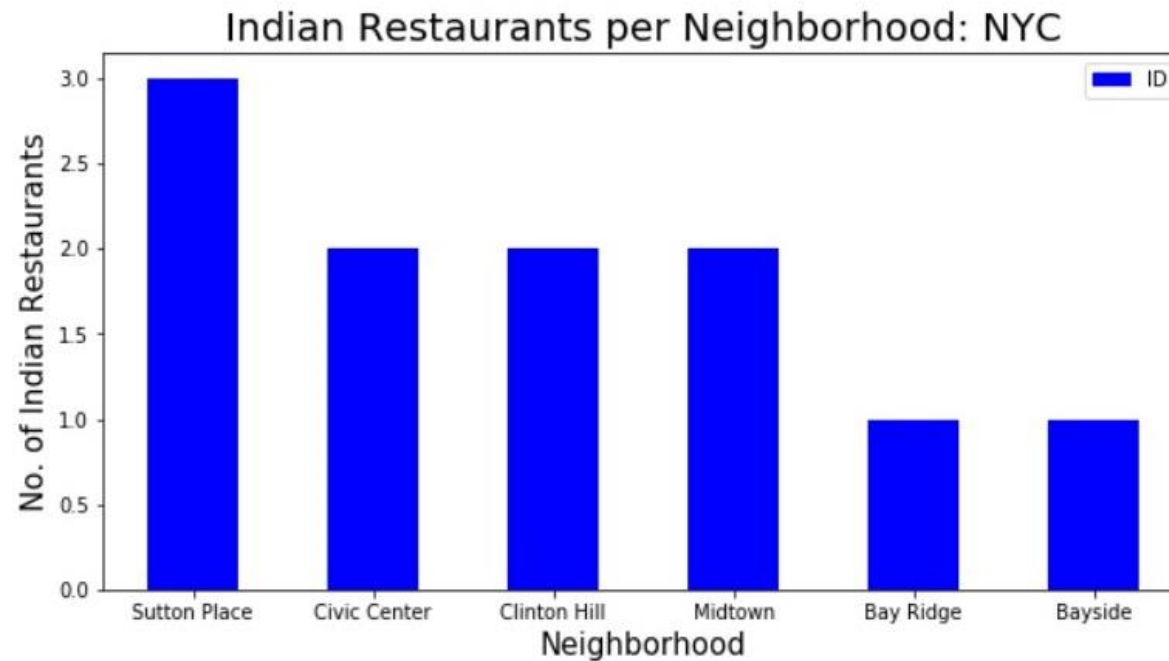|	Unnamed: 0	Borough	Neighborhood	ID	Name
0	0	Bronx	Woodlawn	4c0448d9310fc9b6bf1dc761	Curry Spot
1	1	Bronx	Unionport	4c194631838020a13e78e561	Melanies Roti Bar And Grill
2	2	Brooklyn	Bay Ridge	545835a1498e820edc6f636f	Bombay Grill
3	3	Brooklyn	Greenpoint	51a5445c498ee0f182370cb2	Agra Taj Mahal
4	4	Brooklyn	Bushwick	5169f5c4e4b0c7fcb77a0f3c	Agra Heights

In [15]: indian_ny.shape

Out[15]: (42, 5)

There are 42 Indian Restaurants in the New York City

Indian Restaurants per Borough: NYC
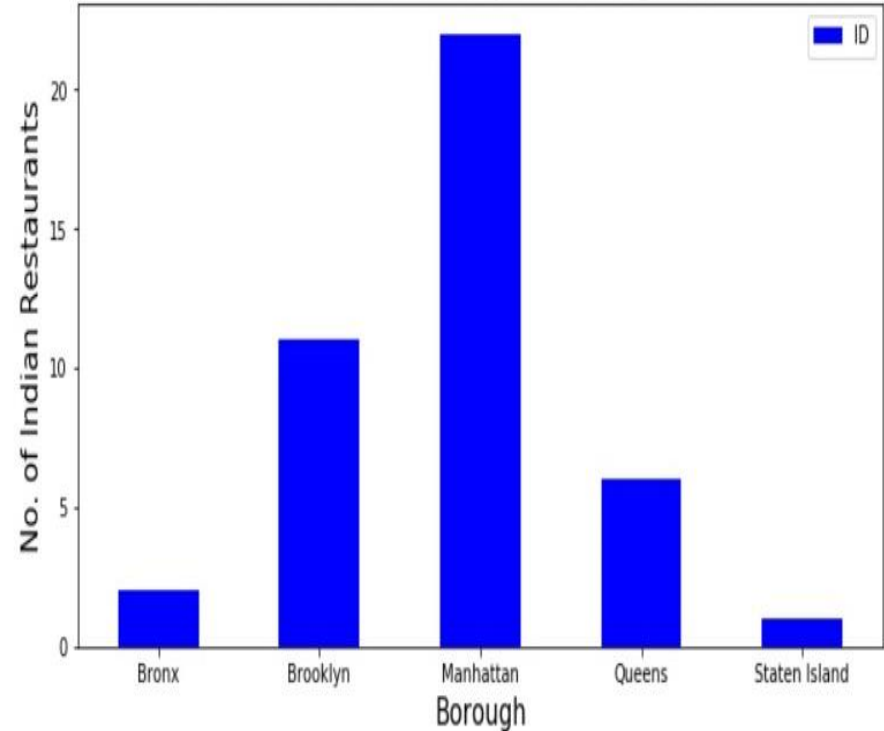
Manhattan has the highest number of Indian restaurants comparing to the other boroughs.

Indian Restaurants per Neighborhood: NYC


```
In [19]:  indian_ny[indian_ny['Neighborhood']=='Sutton Place']
```

Out[19]:

	Unnamed: 0	Borough	Neighborhood	ID	Name
37	37	Manhattan	Sutton Place	4a63bfb4f964a520b3c51fe3	Chola Eclectic Indian Cuisine
38	38	Manhattan	Sutton Place	3fd66200f964a52099e91ee3	Dawat
39	39	Manhattan	Sutton Place	4fe4fb50c2eee335e4fea69d	Moti Mahal Delux

Sutton Place neighborhood in Manhattan Borough have the highest no of Indian Restaurants = 3

Out[26]:

	Likes	Rating	Tips
count	42.000000	42.000000	42.000000
mean	95.404762	7.785714	30.785714
std	178.039193	0.819748	47.582809
min	3.000000	5.700000	1.000000
25%	17.250000	7.525000	8.250000
50%	34.500000	7.850000	16.500000
75%	81.000000	8.400000	28.500000
max	834.000000	9.100000	259.000000

Now we can find which restaurant have maximum likes

In [27]:
```
# Resturant with maximum Likes
indian_rest_stats_ny.iloc[indian_rest_stats_ny['
```

Out[27]:
```
Borough                          Manhattan
Neighborhood                       Midtown
ID            49d91c12f964a520015e1fe3
Name            The Kati Roll Company
Likes                                834
Rating                               8.7
Tips                                 259
Name: 18, dtype: object
```

In [28]:
```
# Resturant with maximum Ratings
indian_rest_stats_ny.iloc[indian_rest_stats_ny['Rating'].idxmax()]
```

Out[28]:
```
Borough                          Manhattan
Neighborhood                       Tribeca
ID            4bbb9dbded7776b0e1ad3e51
Name            Tamarind TriBeCa
Likes                                590
Rating                               9.1
Tips                                 148
Name: 23, dtype: object
```

In [29]:
```
# Resturant with maximum Tips
indian_rest_stats_ny.iloc[indian_rest_stats_ny['Tips'].idxmax()]
```

Out[29]:
```
Borough                          Manhattan
Neighborhood                       Midtown
ID            49d91c12f964a520015e1fe3
Name            The Kati Roll Company
Likes                                834
Rating                               8.7
Tips                                 259
Name: 18, dtype: object
```

Here you can see which Indian restaurant have maximum likes, ratings, and tips. We find these by using python library function to dataset and extract the findings.
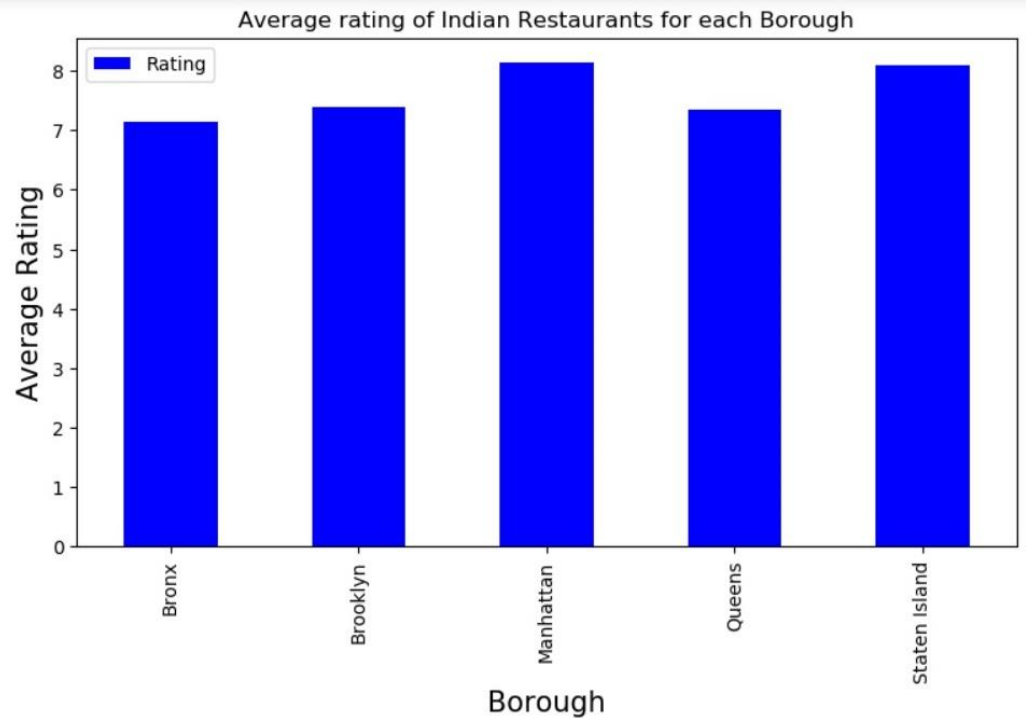
Out[30]:

	Neighborhood	Average Rating
30	Tribeca	9.10
24	Noho	8.80
6	Chelsea	8.70
12	Fort Greene	8.70
26	Prospect Lefferts Gardens	8.60
22	Midtown	8.55
25	North Side	8.50
34	West Village	8.50
15	Greenwich Village	8.40
11	Flatiron	8.40

Above are the top neighborhoods with the highest average ratings of Indian restaurants.

Average rating of Indian Restaurants for each Borough

Out[31]:

	Borough	Average Rating
2	Manhattan	8.140909
4	Staten Island	8.100000
1	Brooklyn	7.400000
3	Queens	7.350000
0	Bronx	7.150000

Out[33]:

	Neighborhood	Average Rating
1	Bayside	8.20
6	Chelsea	8.70
9	East Village	8.40
11	Flatiron	8.40
12	Fort Greene	8.70
15	Greenwich Village	8.40
22	Midtown	8.55
23	New Dorp	8.10
24	Noho	8.80
25	North Side	8.50
26	Prospect Lefferts Gardens	8.60
29	Sutton Place	8.20
30	Tribeca	9.10
34	West Village	8.50
35	Woodlawn	8.00

Now we join this Dataset to original
New York data to get longitude and latitude
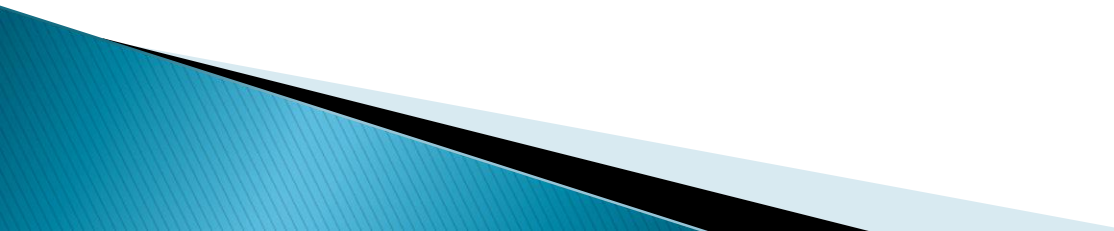
Joined dataset

Out[35]:

	Borough	Neighborhood	Latitude	Longitude	Average Rating
0	Queens	Bayside	40.766041	-73.774274	8.20
1	Manhattan	Chelsea	40.744035	-74.003116	8.70
2	Staten Island	Chelsea	40.594726	-74.189560	8.70
3	Manhattan	East Village	40.727847	-73.982226	8.40
4	Manhattan	Flatiron	40.739673	-73.990947	8.40
5	Brooklyn	Fort Greene	40.688527	-73.972906	8.70
6	Manhattan	Greenwich Village	40.726933	-73.999914	8.40
7	Manhattan	Midtown	40.754691	-73.981669	8.55
8	Staten Island	New Dorp	40.572572	-74.116479	8.10
9	Manhattan	Noho	40.723259	-73.988434	8.80
10	Brooklyn	North Side	40.714823	-73.958809	8.50
11	Brooklyn	Prospect Lefferts Gardens	40.658420	-73.954899	8.60
12	Manhattan	Sutton Place	40.760280	-73.963556	8.20
13	Manhattan	Tribeca	40.721522	-74.010683	9.10
14	Manhattan	West Village	40.734434	-74.006180	8.50
15	Bronx	Woodlawn	40.898273	-73.867315	8.00

Results

❖ Manhattan has the highest number of Indian restaurants comparing to the other boroughs.

❖ Manhattan and Staten Island have the best rated Indian restaurants on average.

❖ Tribeca neighborhoods with the highest average rating for Indian Restaurants
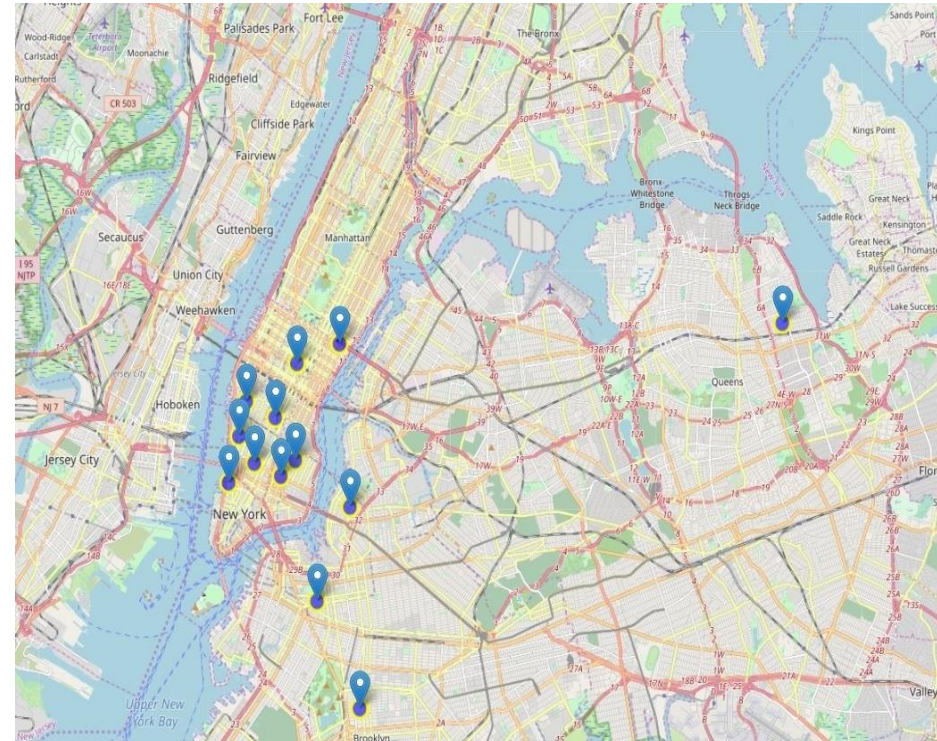
Discussion Section

❖ Based on the results of our analysis, Manhattan and Staten Island are the best locations for Indian cuisine in New York City.

❖ Staten Island is a more promising place to open a new Indian restaurant.

❖ Neighborhoods Chelsea is the best place in Staten Island to open a new Indian restaurant.

Conclusion

Why did we decide to focus on Indian cuisine in our project? Now when the idea of a healthy lifestyle conquered the minds of people all over the World, Indian restaurants became extremely popular, as they offer a healthy alternative to regular American eating habits. In consequence, as a reader of this data science report, common people of New York can find where they have Indian restaurants to have healthy foods for better life.