

# **Finding the place with best Indian Restaurants in NYC**

Shreshth Jain

September 8, 2019

## **1. Introduction**

### **1.1 Background**

The city of New York, also popularly known as New York City (NYC), is the most populous city in USA. It has an estimated population of 8,398,748 as of 2018 with a land area of 302.6 square miles. Located at the southern tip of the state of New York, the city is the center of the New York metropolitan area, the largest metropolitan area in the world by urban landmass and one of the world's most populous megacities.

Throughout history, New York City has been a major point of entry of immigrants, in fact the term "melting pot" was coined to describe the densely populated immigrant neighborhoods on the Lower East Side. More than 800 languages are spoken in New York, making it the most linguistically diverse city in the world.

Obviously, with diversity of culture comes diversity of food. There are restaurants of all cuisines imaginable in New York City. The City's food culture boasts off an array of international cuisines influenced by city's immigrant history. From European influences like Italian, French and English to Asian countries like India, China, Japan and Korea, all of them have a place in NYC.

With such a background, determining a good location to open a new restaurant can be a challenge of its own.

### **1.2 Problem**

This project aims find out the most optimum place to open a new Indian restaurant and houses the best Indian Restaurants. This would especially be useful for those who are looking for best Indian restaurants and also wish to open a new Indian Restaurant in NYC but are not sure about the location. Also, this project can also be used as template for those who are trying to open a restaurant of any other cuisine as well.

## **2. Data Sources**

Following data will be required for the project:

Data on New York City that contains a list of Boroughs, Neighborhoods along with their latitude and longitude.

- Data Source: [https://cocl.us/new\\_york\\_dataset](https://cocl.us/new_york_dataset)
- Description: It contains all the required information and will be used for exploring neighborhoods of New York City.

Indian Restaurants in each neighbourhood of New York City.

- Data Source: FourSquare API
- Description: We can filter the venues so obtained to get only Indian Restaurants

GeoSpace Data

- Data Source: <https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm>
- Description: It will give us the New York Borough boundaries which will help us in visualizing choropleth map.

### 3. Methodology

#### 3.1 Important Libraries

I began with downloading all the required libraries for this project. This includes **Pandas** and **Numpy** in order carry out basic pandas database functions like `read_csv` etc. I also downloaded **Geopy** so as to use **geocoders** in order to convert an address into its latitude and longitude values. Also, I needed to import **Folium** so that I could visualize maps. Last but not the least, import all the **matplotlib** libraries in order to visualize graphs.

#### 3.2 Data Scrapping

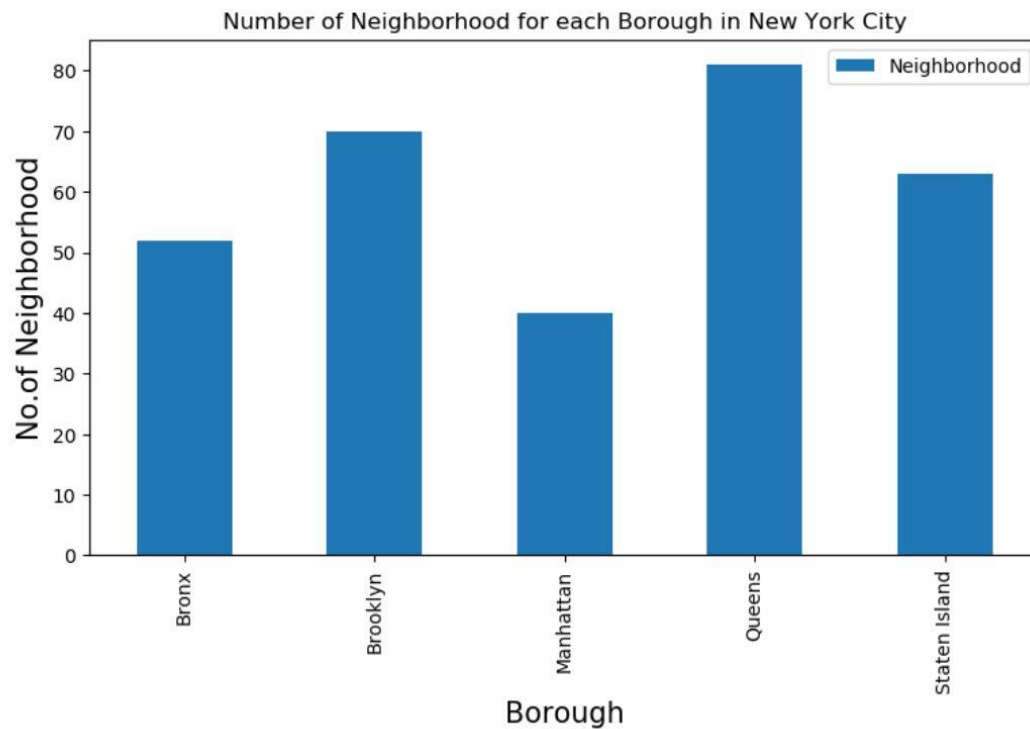
First, I use Geopy to get the latitude and longitude of a given location for which I have defined a function called `geo_location`. Next step was to use FourSquare API to get top 100 venues within a radius of 1000 meters for given latitude and longitude. This data contained Venue ID, Venue Name and Category.

Then I defined a function to get venue details like ratings, tips, likes count for a given venue ID in order to rank all the restaurants. Again, I used FourSquare API to achieve this task.

Finally, I used to the URL "[https://cocl.us/new\\_york\\_dataset](https://cocl.us/new_york_dataset)" to get the New York City data and then iterated through it to create a new dataframe `new_york_data` that would also contain latitude and longitude values.

### 3.3 Data Analysis

First, I visualized a bar graph that displayed the number of neighborhoods for each borough in New York City. It was observed that Queens had the highest number of neighborhoods among all the Boroughs.



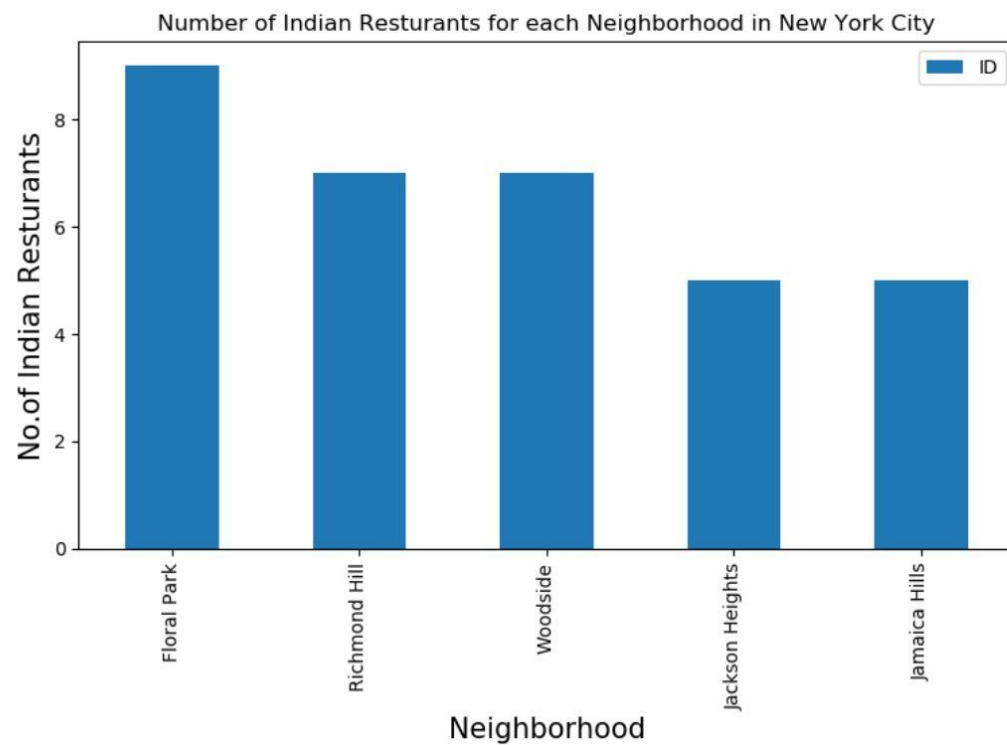
Next, I iterate through the `new_york_data` dataframe to filter out Indian restaurants in each neighborhood. This results in the observation that there are total 153 Indian restaurants in NYC. I also visualize the number of Indian restaurants for each borough in

New York City.



Its easily observed that Queens has the highest number of Indian restaurants.

Lastly, I also visualize a bar graph for number of Indian Restaurants for each neighborhood which leads me to the conclusion that Floral Park has the highest number of Indian restaurants.



```
In [238]: indian_rest_ny[indian_rest_ny['Neighborhood']=='Floral Park']
```

```
Out[238]:
```

	Borough	Neighborhood	ID	Name
103	Queens	Floral Park	527ffc0811d2d329d5e49abd	Jackson Diner
104	Queens	Floral Park	4b647b56f964a520c4b62ae3	Usha Foods & Usha Sweets
105	Queens	Floral Park	4b787c49f964a5209cd12ee3	Santoor Indian Restaurant
106	Queens	Floral Park	4e4e3e22bd4101d0d7a5c2d1	Kerala Kitchen
107	Queens	Floral Park	4c0c01e0bbc676b00d6b4cd5	Mumbai Xpress
108	Queens	Floral Park	4c76ff35a5676dcb72671721	Flavor Of India
109	Queens	Floral Park	4df0f39dd4c04d0392c853ea	Sagar Chinese
110	Queens	Floral Park	571af96a498e9e392d8d3786	Namaste Authenic Indian Cuisine
111	Queens	Floral Park	51d84192498ea979a3c4f13d	Sunshine Grill & Restaurant

As can be seen there are a total of 9 Indian Restaurants in Floral Park alone.

Next step was to rank all the Indian restaurants. In order to achieve that, I first created a dataframe `indian_rest_stats_ny` that contained all the details required for ranking like Ratings, Tips and Likes. I made use of the function `venue_details` which I defined earlier in order to achieve the `Indian_rest_stats_ny` DataFrame.

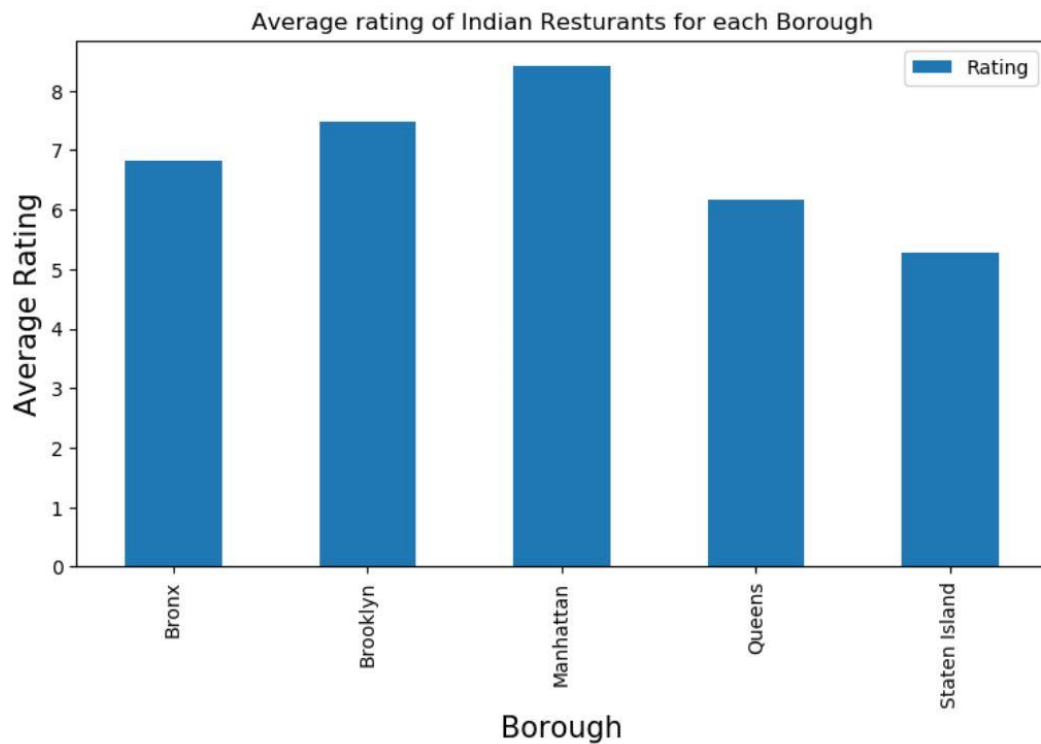
Next logical step was to find those Indian restaurants with the highest count of Likes, Ratings and Tips for which I utilized the `idxmax` method.

Then I find which neighborhood has the highest average rating and the borough that has the highest average rating. For both these tasks, I construct two new DataFrames sorted by average ratings of each restaurant for neighborhoods and borough respectively.

	Neighborhood	Average Rating
0	Astoria	9.200000
71	Sunnyside	9.200000
75	Tribeca	9.200000
5	Blissville	9.200000
11	Civic Center	9.200000
47	Midtown	9.000000
48	Midtown South	9.000000
30	Gramercy	8.866667
66	Roosevelt Island	8.700000
53	North Side	8.700000

	Borough	Average Rating
2	Manhattan	8.414706
1	Brooklyn	7.478571
0	Bronx	6.812500
3	Queens	6.155844
4	Staten Island	5.266667

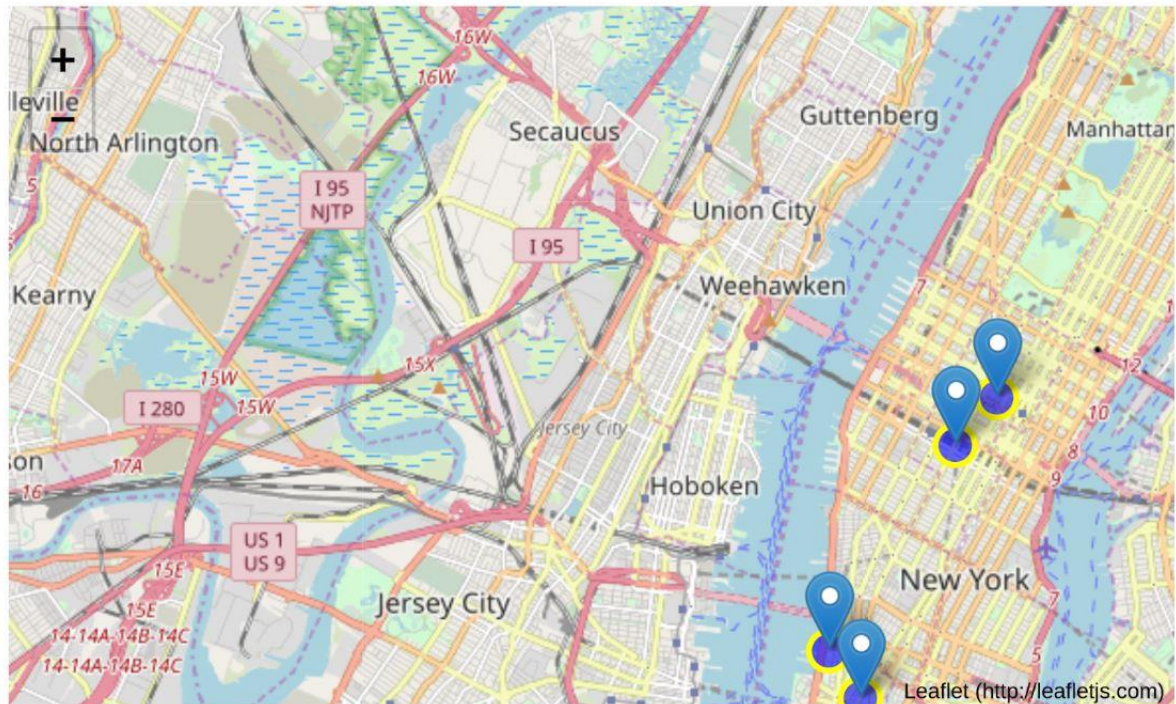
And then a final visualization of average rating of each restaurant for each Borough.



Then I constructed the final dataframe `ny_neighborhood_stats` in order to visualize it on a map. The dataset looked like the picture below.

	Borough	Neighborhood	Latitude	Longitude	Average Rating
0	Queens	Astoria	40.768509	-73.915654	9.2
1	Queens	Blissville	40.737251	-73.932442	9.2
2	Manhattan	Civic Center	40.715229	-74.005415	9.2
3	Manhattan	Midtown	40.754691	-73.981669	9.0
4	Manhattan	Midtown South	40.748510	-73.988713	9.0
5	Queens	Sunnyside	40.740176	-73.926916	9.2
6	Staten Island	Sunnyside	40.612760	-74.097126	9.2
7	Queens	Sunnyside	40.740176	-73.926916	9.2
8	Staten Island	Sunnyside	40.612760	-74.097126	9.2
9	Queens	Sunnyside	40.740176	-73.926916	9.2
10	Staten Island	Sunnyside	40.612760	-74.097126	9.2
11	Queens	Sunnyside	40.740176	-73.926916	9.2
12	Staten Island	Sunnyside	40.612760	-74.097126	9.2
13	Manhattan	Tribeca	40.721522	-74.010683	9.2

Finally, using the Folium library, I visualize the above data on a map.



#### 4. Results

- Astoria (Queens), Blissville (Queens), Civic Center (Manhattan) are some of the best neighborhoods for Indian cuisine.
- Manhattan has good potential for Indian Restaurant market.
- Staten island ranks last in average rating of Indian restaurants

#### 5. Discussion

Using bar graphs, it was easily observed that Queens had the highest number of neighborhoods among all the boroughs which is understandable, given it is one the most vibrant and diverse boroughs and it also houses perhaps, the largest Indian community in the United States. It was also observed that there are 153 Indian restaurants in NYC with the neighborhood of Floral Park having the highest number of Indian restaurants.

After sorting the data by average ratings, it was deduced that Manhattan has the highest average rating, making it the borough with the highest potential for Indian restaurants. Manhattan being the most densely populated borough of all the five boroughs in New York City, its pretty understandable why this is the case.

#### 6. Conclusion

I certainly conclude that Manhattan is the best place to stay if one simply cannot avoid Indian restaurants. It checks all the boxes and would be a delight for anyone who wants to open a good Indian Restaurant in NYC.

