Finding Potential Neighbourhoods in New York city for Indian Food Restaurants

Guneet Deol
February 13, 2020

1.Introduction

New York (NY), is the most populous city in the United States. With an estimated 2018 population of 8,398,748 distributed over a land area of about 302.6 square miles. New York metropolitan area is the largest metropolitan area in the world and is one of the world's most populous mega cities. People have travelled from various parts of the world and inhabited New York.

As different immigrants come from different nations in New York, they bring with them different and unique cultures and food cuisines. Different cuisines are not only liked by foreign born residents of New York but also by the natives in the city who like trying different food items. Opening a foreign-based food cuisine restaurant in New York would thus be a good business idea. Indian cuisine reflects an 8,000-year history of various groups and cultures interacting with the Indian subcontinent, leading to diversity of flavours and regional cuisines found in India today. British and Portuguese influence has also added to the already diverse Indian cuisine.

In this project we look at different neighbourhoods in New York city and conclude which would be a good neighborhood to open an Indian food restaurant. The project would be helpful for audience who have moved from India or South Asia to New York and are looking to set up restaurant business in the city. For example: Let us say a person has moved from India to permanently live in New York and now he wishes to set up an Indian Food restaurant in New York to earn his living. The Analysis would provide the audience insights about various neighbourhoods in New York city and help them in making decision about which neighbourhood would be most suitable to start their business. For Example: The person who has moved from India to New York can know from the analysis that :

- which neighbourhoods have maximum number of Indian Restaurants,

- which neighbourhood has minimum number of restaurants,

- which Neighbourhood is best for setting up new business,

- how much competition is there in the neighborhood from other Indian Restaurants(this can be seen by viewing the rating).

2. Data acquisition and cleaning

To analyse our data and reach a conclusion, we first need to prepare our data. Data preparation includes: cleaning the data and getting the data to a stage where it may be easier to work with the data. Here we will convert addresses into their equivalent latitude and longitude values and **use Foursquare API** to explore neighborhoods in New York City. We will use the Location intelligence provided by Foursquare API to provide a solution to our client. We import the following required libraries as prerequisites before working with our data:

1. import numpy as np

2. import pandas as pd

3. from geopy.geocoders import Nominatim

4. import json

5. import requests

6. from pandas.io.json import json_normalize

7. import matplotlib.pyplot as plt

8. import matplotlib.cm as cm

9. import matplotlib.colors as colors

10. !conda install -c conda-forge folium=0.5.0 –yes (to install Folium)

11. import folium

2.1 Data Sources

In order to segement the neighborhoods and explore them, we will essentially need a dataset that contains all the boroughs and the neighborhoods that exist in each borough as well as the the latitude and logitude coordinates of each neighborhood. We can load the data from https://cocl.us/new_york_dataset which is a json file available on the server. We load the data from the json file and then loop through the file one row at a time to transfer our data into a dataframe that has structured format. We will use geopy library to get the latitude and longitude values of New York City. we can explore the restaurants in the neighborhood, after the required data is ready i.e:

- we have the latitudes and longitudes of all the neighborhoods

- we have structured data ready in the form of a table having columns: Boroughs, Neighborhoods, Latitude, Longitude

- we have generated our Foursquare credentials i.e client_id and client_secret. Foursquare credentials can be generated by creating a developer Foursquare account.

2.2 Feature Selection

Once our data is prepared, we can extract various features from our data using the Foursquare API like:

- getting the venues(restaurants) within a specified radius.

- filtering Indian restaurants within that radius.

- getting venue details like tip count, likes count for a given venue id.

- finding the ratings of different Indian cuisine restaurants in the area which can furthus help the client who is trying to establish a new business in New York

- Find tips for each Indian restaurant. All this information can help our client to get an idea about the present market of Indian cuisine restaurants and prepare better in launching his or her own restaurant.
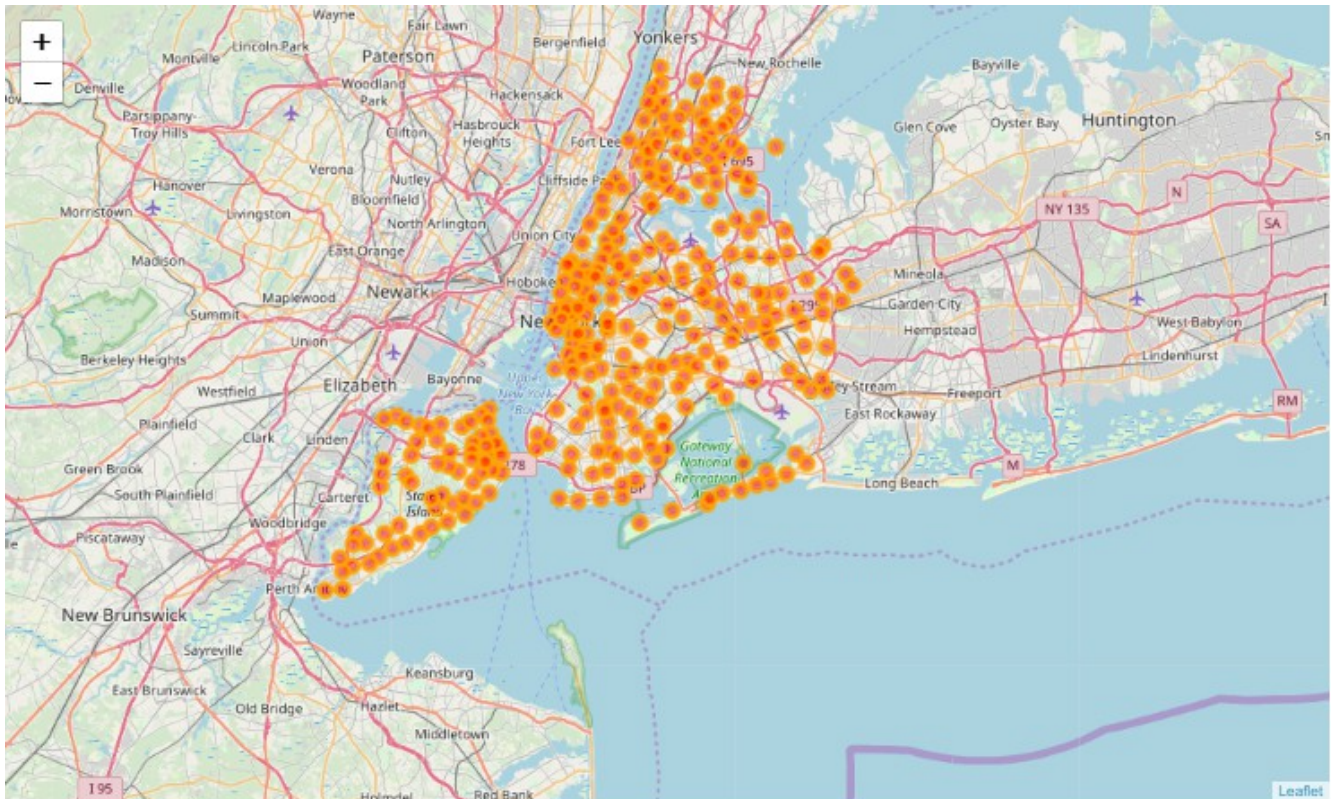
3. Methodology:

3.1 Exploratory Data Analysis:

Exploratory data analysis allows us to:
• summarize main characteristics of the data.
• uncover relationships between the data.
• gain better understanding of the data.
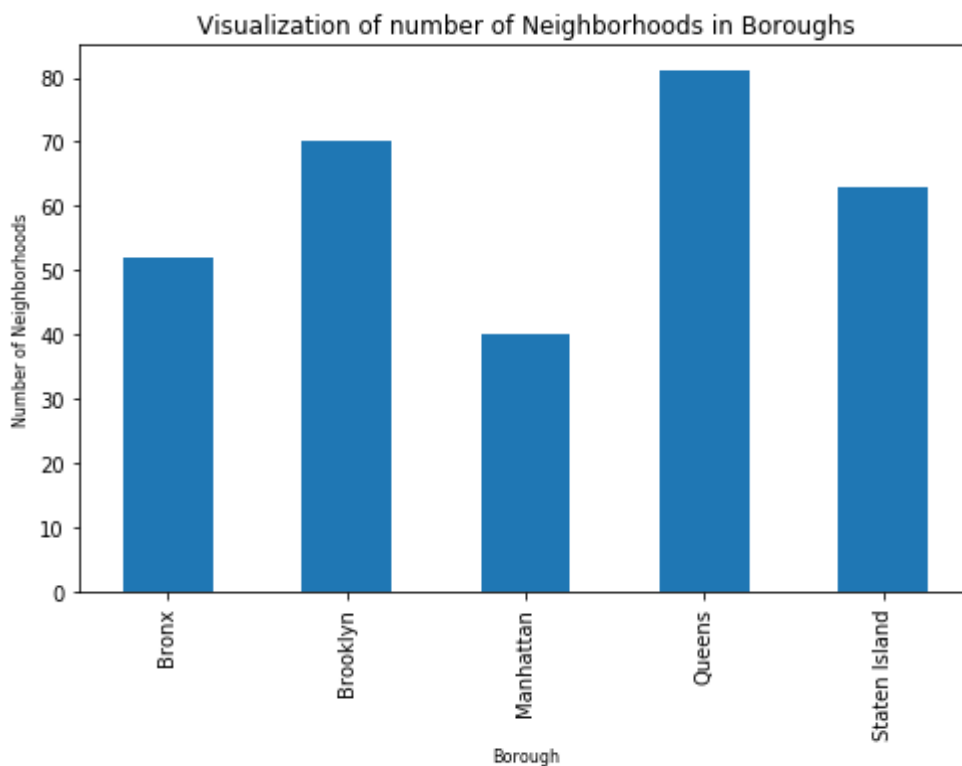• Extract important variables from the data

To understand our data we use exploratory data analysis to generate a map of New York city with neighborhoods superimposed on the top of the map. The Neighborhoods are superimposed with the help of latitude and longitude values.

3.2. Relationship plots

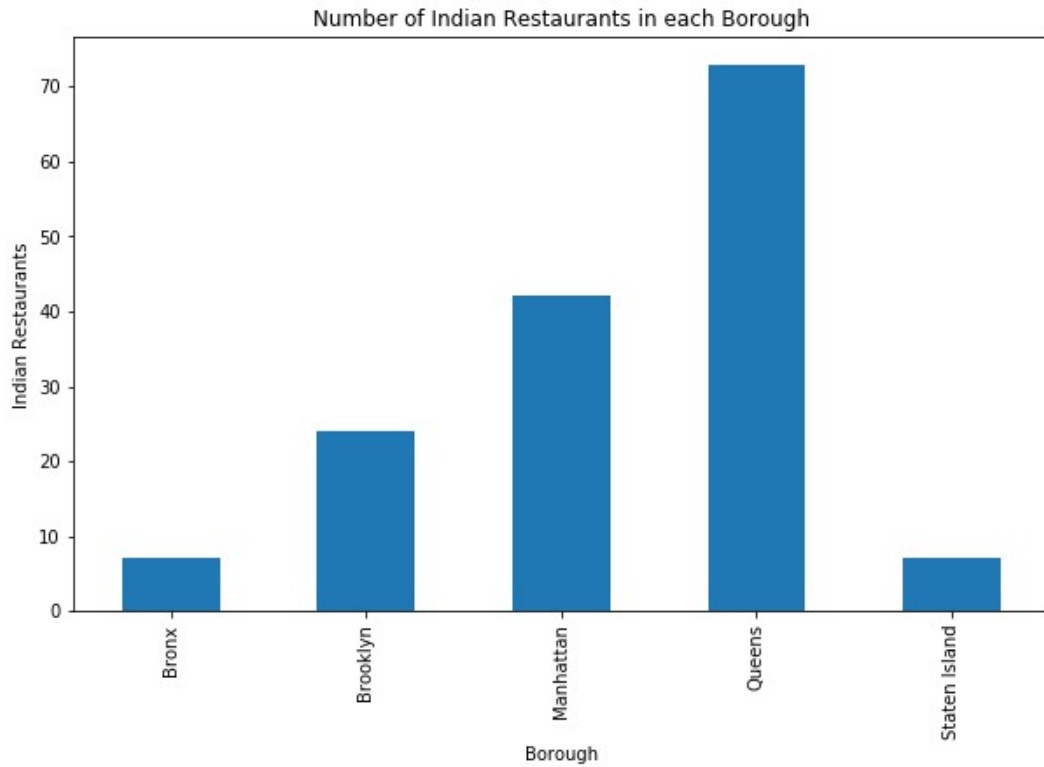### 3.2.1. Number of Neighborhoods in each borough:

New York City has five administrative divisions called boroughs. These are: The Bronx, Brooklyn, Manhattan, Queens, and Staten Island. To understand our data better, we visualize the number of Neighborhoods in each borough. this tells us, which Borough has the highest number of neighborhoods.

We can see that Queens has the highest number of Neighborhoods, followed by Brooklyn and Staten Island. where as Bronx and Manhattan have least number of Neighborhoods.

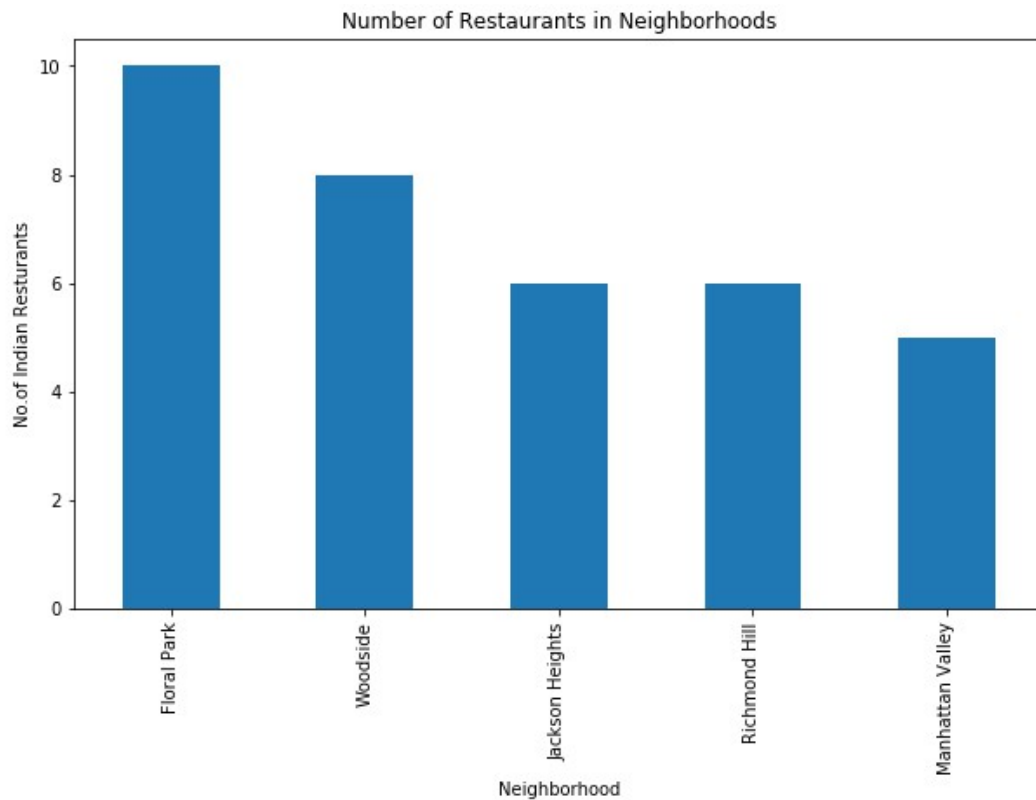### 3.2.2. Number of Indian Restaurants in each Borough:

To get an idea about the Indian Restaurant market in New York we find out the number of Indian Food Restaurant in each Borough. This tells us which Borough has the highest number of Indian restaurants and thus has a good established market for Indian food.

We can see that Queen's has the highest number of Indian Restaurants followed by Manhatten and Brooklyn whereas Bronx and Staten Island have the least number of Indian Restaurants.

3.2.3 Neighborhoods with most number of Indian restaurants:

We already know that Borough: Queens has the highest number of Indian Restaurants so now we find the Neighborhoods with most number of Indian Food restaurants. The Bar chart shows that Floral Park has the highest number of Restaurants followed by Woodside, Jackson Heights, Richmond Hill and Manhattan Valley.
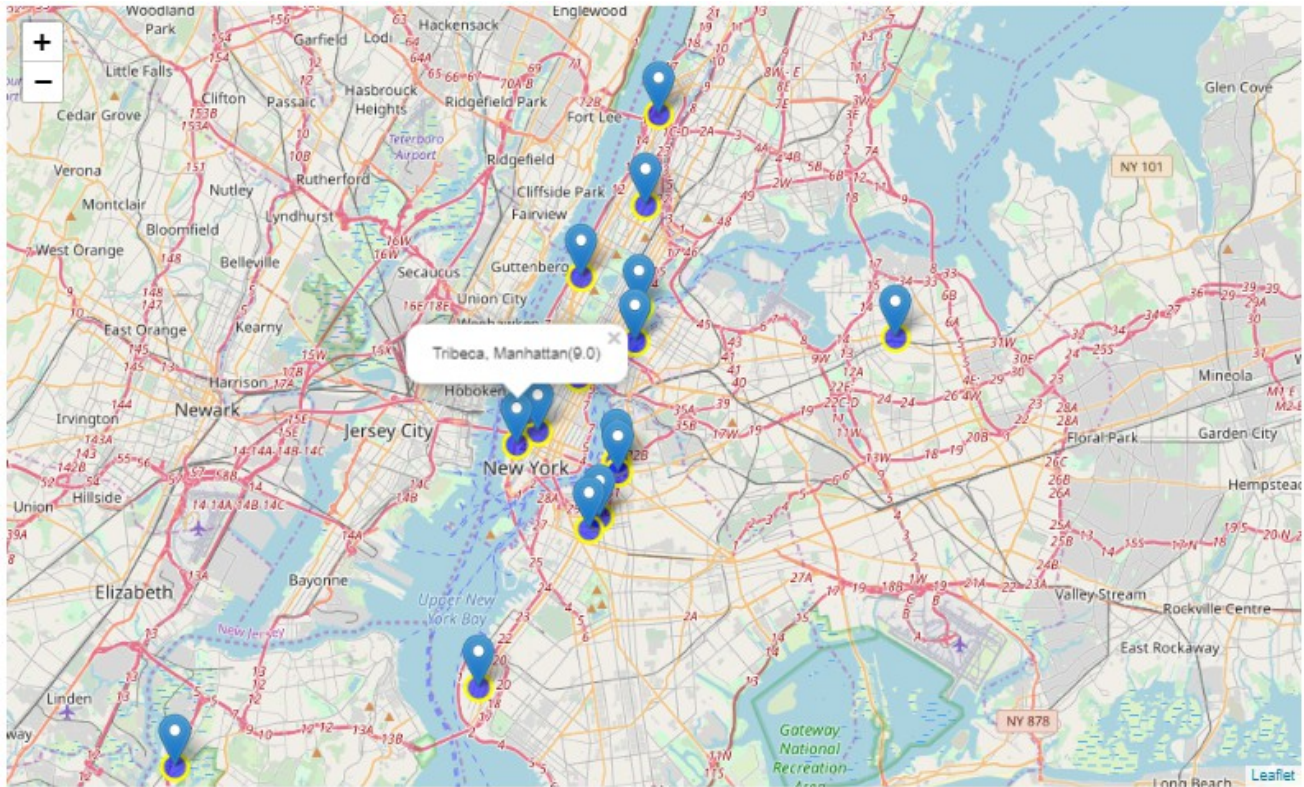
4. Results

Once we know the numeric values about the number of restaurants in each Borough and Neighborhoods with high number of restaurants we find the Average Rating for the restaurants in the Neighborhoods to know about the already running market of Indian cuisine Restaurants and get an estimate about the competition in the Indian Restaurant market. Foursquare API was used to get the Likes , Ratings and number of tips for the restaurants. The below table shows the average rating for Indian food restaurants in the Neighborhoods in descending order. Only the first 20 rows are shows in the table.

	Borough	Neighborhood	Latitude	Longitude	Average Rating
12	Manhattan	Tribeca	40.721522	-74.010683	9.000000
6	Manhattan	Greenwich Village	40.726933	-73.999914	8.900000
15	Manhattan	West Village	40.734434	-74.006180	8.850000
8	Queens	Murray Hill	40.764126	-73.812763	8.800000
7	Manhattan	Murray Hill	40.748303	-73.978332	8.800000
2	Manhattan	Chelsea	40.744035	-74.003116	8.700000
3	Staten Island	Chelsea	40.594726	-74.189560	8.700000
5	Brooklyn	Fort Greene	40.688527	-73.972906	8.700000
11	Brooklyn	South Side	40.710861	-73.958001	8.500000
9	Brooklyn	North Side	40.714823	-73.958809	8.500000
0	Brooklyn	Bay Ridge	40.625801	-74.030621	8.400000
10	Manhattan	Roosevelt Island	40.762160	-73.949168	8.300000
16	Manhattan	Yorkville	40.775930	-73.947118	8.250000
4	Brooklyn	Clinton Hill	40.693229	-73.967843	8.150000
13	Manhattan	Upper West Side	40.787658	-73.977059	8.133333
1	Manhattan	Central Harlem	40.815976	-73.943211	8.100000
14	Manhattan	Washington Heights	40.851903	-73.936900	8.000000

The Neighborhoods having restaurants with an Average Rating greater than 8 are marked on a map which provides us an insight about the potential neighborhoods to start an Indian cuisine restaurant.

5. Discussion:

From the above analysis by looking at the Bar plots and Ratings we can conclude that:

- Queens has the Maximum number of Indian Restaurants
- Tribeca, Manhattan has the Highest Average Rating for Indian Restaurants.
- Manhattan has some of the top rated Indian restaurants
- Brooklyn also has many high rated Indian Restaurants
- Manhattan and Brooklyn would be good neighborhoods to stay if you want to try Indian cuisine in New York.
- Tribeca, Manhattan Would be a Good choice to start a New Indian Restaurant.
- North and South Side in Brooklyn would also be good option as they have Rating ~ 8.5 which provides a less competitive market to open a new Business.

6. Conclusion:

The report visualizes all the neighborhoods in the city of New York on a Map and uses bar plots to show us:

- which Borough has the highest number of neighborhoods
- which Borough has the highest number of Indian restaurants
- which neighborhoods have the most number of Indian restaurants

The report also :

- sorts the Indian food restaurants according to their rating
- visualizes the restaurants with rating higher than 8.0 on a map
- provides insights about:
    - which neighborhood has the maximum number of Indian restaurants
    - which neighborhood is best to stay if you like Indian cuisine
    - which neighborhoods have some of the highest rating restaurants in the city
    - which are the potential neighborhoods to start a new Indian restaurant.