**My Capstone project-Analysis Indian restaurants in New York City**

Shakambari Swain

August 12,2019

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. Over the last decade the city has been growing faster than the region. The New York region continues to be by far the leading metropolitan gateway for legal immigrants admitted into the United States.

As many as 800 languages are spoken in New York, making it the most linguistically diverse city in the world.

With its diverse culture, comes diverse food items. There are many restaurants in New York City, each belonging to different categories like Chinese, Indian and French etc.

The problem statement is to identify and visualize all major parts of New York City that contains Indian restaurants. (With limited set of test data, as we can access only limited data from foursquare per day)

Following Questions Answered

What is best location in New York City for Indian Cuisine?

Which areas have potential Indian Restaurant Market?

Which all areas lack Indian Restaurants?

Which is the best place to stay if I prefer Indian Cuisine?

1. **Data acquisition and cleaning**

For this project we need the following data:

New York City data that contains list Boroughs, Neighborhoods along with their latitude and longitude.

Data source: **https://cocl.us/new\_york\_dataset**

Description: This data set contains the required information. And we will use this data set to explore various neighborhoods of New York City.

Indian restaurants in each neighborhood of New York City.

Data source: **Foursquare API**

Description: By using this API we will get all the venues in each neighborhood. We can filter these venues to get only Indian restaurants.

GeoSpace data

Data source: [**https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm**](https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm)

Description: By using this geo space data we will get the New York Borough boundaries that will help us visualize choropleth map.

1. **Exploratory Data Analysis**

Collect the New York City data from https://cocl.us/new\_york\_dataset Using Foursquare API we will find all venues for each neighborhood.

Filter out all venues that are Indian Restaurants. Find rating, tips and like count for each Indian Restaurants using Foursquare API.

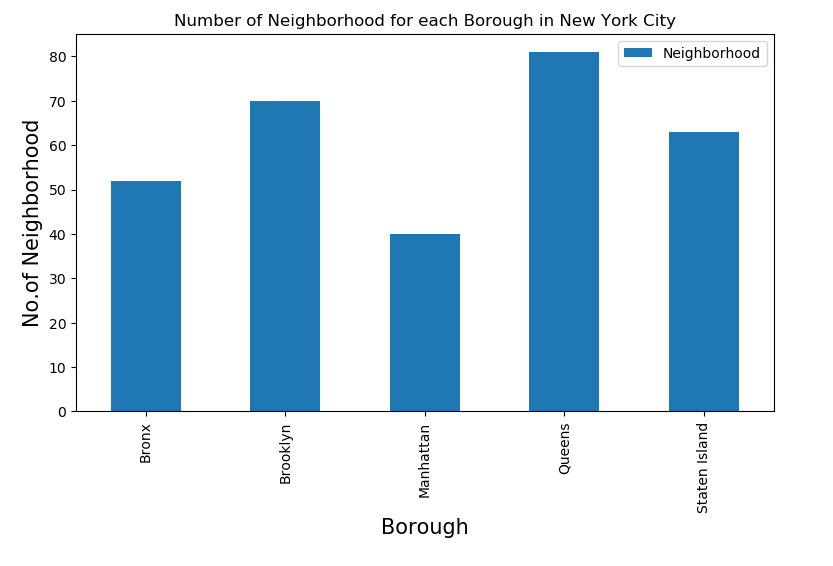
Using rating for each restaurant, we will sort that data.

Visualize the Ranking of neighborhoods using folium library (python)

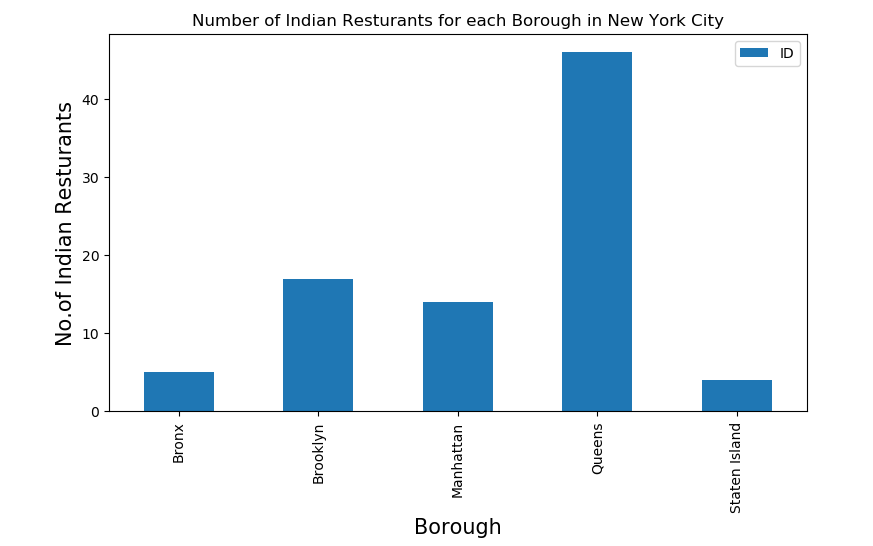
The Analysis was performed based on available per day Foursquare data Limit

**The main steps followed**

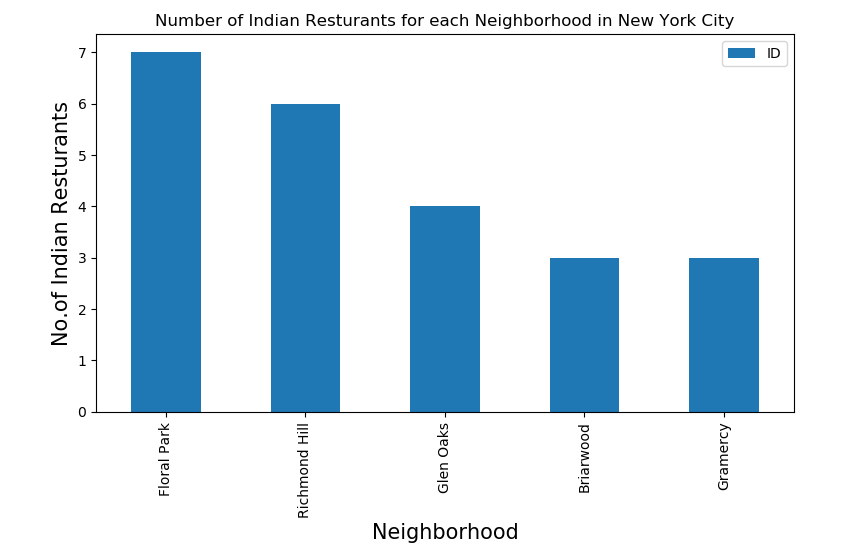
1. Calculated Number of Neighborhood for each Borough in New York City



1. Calculated Number of Indian Restaurants for each Borough in New York City



1. Calculated Number of Indian Restaurants for each Neighborhood in New York City



1. **Predictive Modeling**

Used Bar chat to predict the result.

* Solution to the problems

From the figure we concluded the following information.

1. Sunnyside (Staten Island), Sunnyside (Queens), Tribeca (Manhattan) are some of the best neighborhoods for Indian cuisine.
2. Manhattan have potential Indian Restaurant Market/
3. Manhattan is the best place to stay if you prefer Indian Cuisine.
4. It seems “Staten Island” ranks last in average rating of Indian Restaurants

* Limitations

The ranking is purely on basis of rating of restaurants

The accuracy of data depends purely depends on the data provided by Foursquare