Battle of the Neighborhoods

Radhika Prasad

1.Introduction

1.1 Background

Restaurants are a notoriously difficult business to own or operate. Not only is it a relatively capital and employee-intensive business, restaurants are also highly regulated, low margin and in most cases have a plethora of competition to deal with. There are more dining establishments and more diners now than during any other time in history; the National Restaurant Association reports over 1 million restaurant locations in the U.S. alone, a particular boon considering more than half the American population visits at least one of them each week. To fill a niche and remain relevant, prospective, as well as established, restaurants have to hedge their bets with well-rounded and well-directed dataset. Data science provides valuable insights regarding market trends and evolving consumer lifestyles so that restaurateurs can better address and meet public demand.

1.2 Business Problem

Restaurant is a place where people come to have food and drinks for a cost, People love to do many things and try something new or stick with their own routines, it depends on the individual and there are so many of them with different cultural and various aspects in New York city. There are so many cuisines, which is based on the style of cooking, the

ingredients, dishes and techniques. For our problem let's stick with Indian cuisine.

My client, a successful Indian restaurant chain in Hyderabad-India is looking to expand operation into North America through New York (NYC), so here are all the factors we have to take into account such as follows:

- Market Places
- Competition in particular location
- Aiding places that make people come to restaurants like Gym, Entertaining Public places
- Population
- Menu from competitors

And so on... So, our solution needs to be data driven for avoiding or considering low risk criteria and high success rate and thus apply our overall knowledge in the techniques and the tools gained so far in this course.

2. Data Acquisition and Cleaning

2.1 Data Sources

Since we will be focusing on New York for opening our restaurant, we will be gathering the data for the same and the three main data we are going to use are the New York Neighborhood data which we already used in the Neighborhoods in New York and the data about the cuisines in New York and the Geospatial data of boundaries of the Neighborhoods in New York.

2.1.1 Data 1

New York Neighborhood Data which will be used to get to know about the various neighborhoods which are going to take into consideration for our Project.

	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

Data source: https://cocl.us/new_york_dataset

2.1.2 Data 2:

NYC Population & Demographic characteristics

Datasource:https://en.wikipedia.org/wiki/New_York_City ; https://en.wikipedia.org/wiki/Demographics_of_New_York_City.

Web scraping techniques was used to get NYC's population density and demographics data from Wikipedia. Preliminary finding indicates that Queens being the second most populous urban area in New York City (NYC), behind Brooklyn; and the most ethnically diverse urban area in NYC with the highest Asian ethnic minority population.

2.1.3 Data 3

Finally, the data going to be collected/acquired from the Foursquare API about the various restaurants in each neighborhood of New York city especially Indian cuisine, which will be used for acquiring the information regarding same for all the venues of each neighborhood.

Data source: Foursquare

2.2 Data Cleaning and approach

• Acquire the data from the various sources as mentioned below,

a. New York neighborhood data fromhttps://cocl.us/new_york_dataset
b. NYC Population & Demographic characteristics from https://en.wikipedia.org/wiki/New_York_City ; https://en.wikipedia.org/wiki/Demographics_of_New_York_City.
c. Indian Cuisine related data from Foursquare

• Filter the data for the required data acquired from Foursquare which is only about Indian Cuisine.

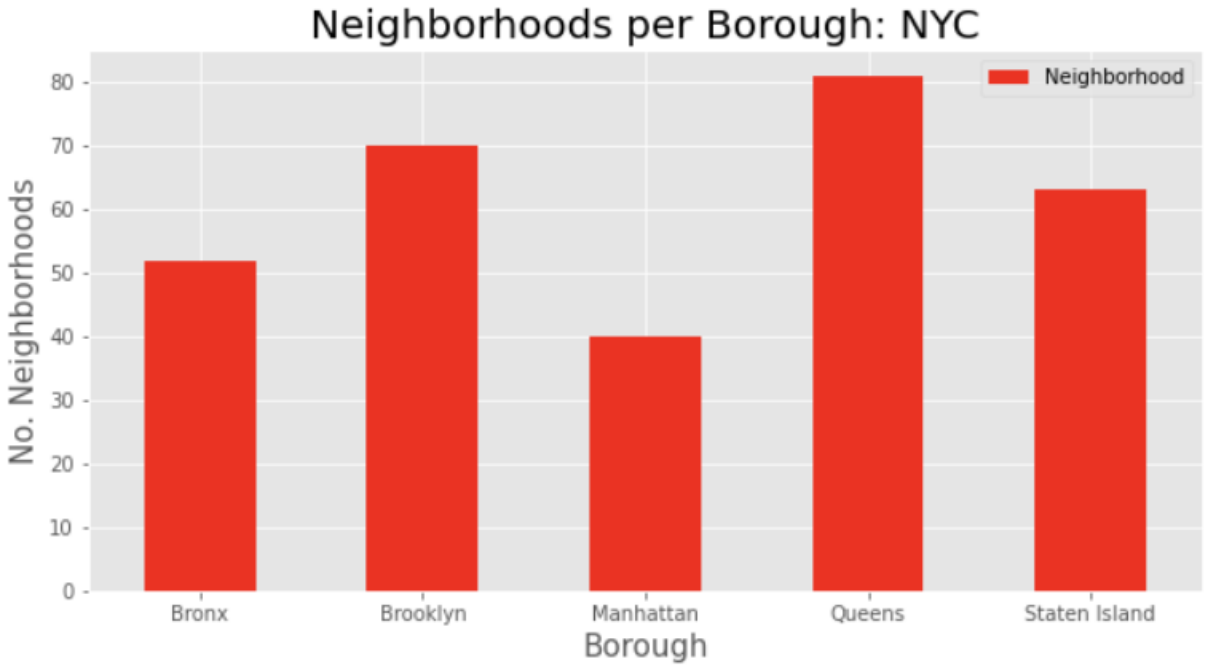
• Acquire the tips and from that get the rating of every restaurant in the neighborhood of New York.

• Visualize the findings using Matplotlib and Folium for better understanding.

3. Exploratory Data Analysis

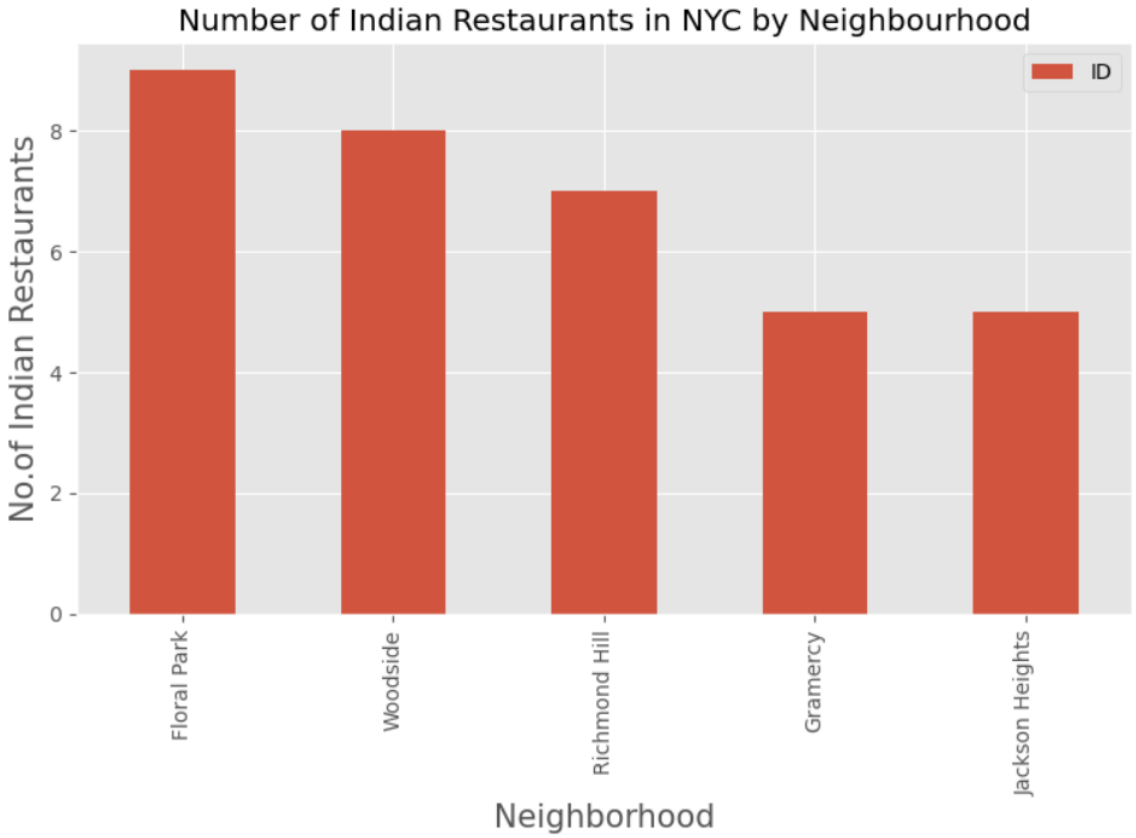
3.1 Number of Neighborhoods in New York city

Below visualization gives us the understanding of the boroughs in New York city and the number of neighborhoods in it.



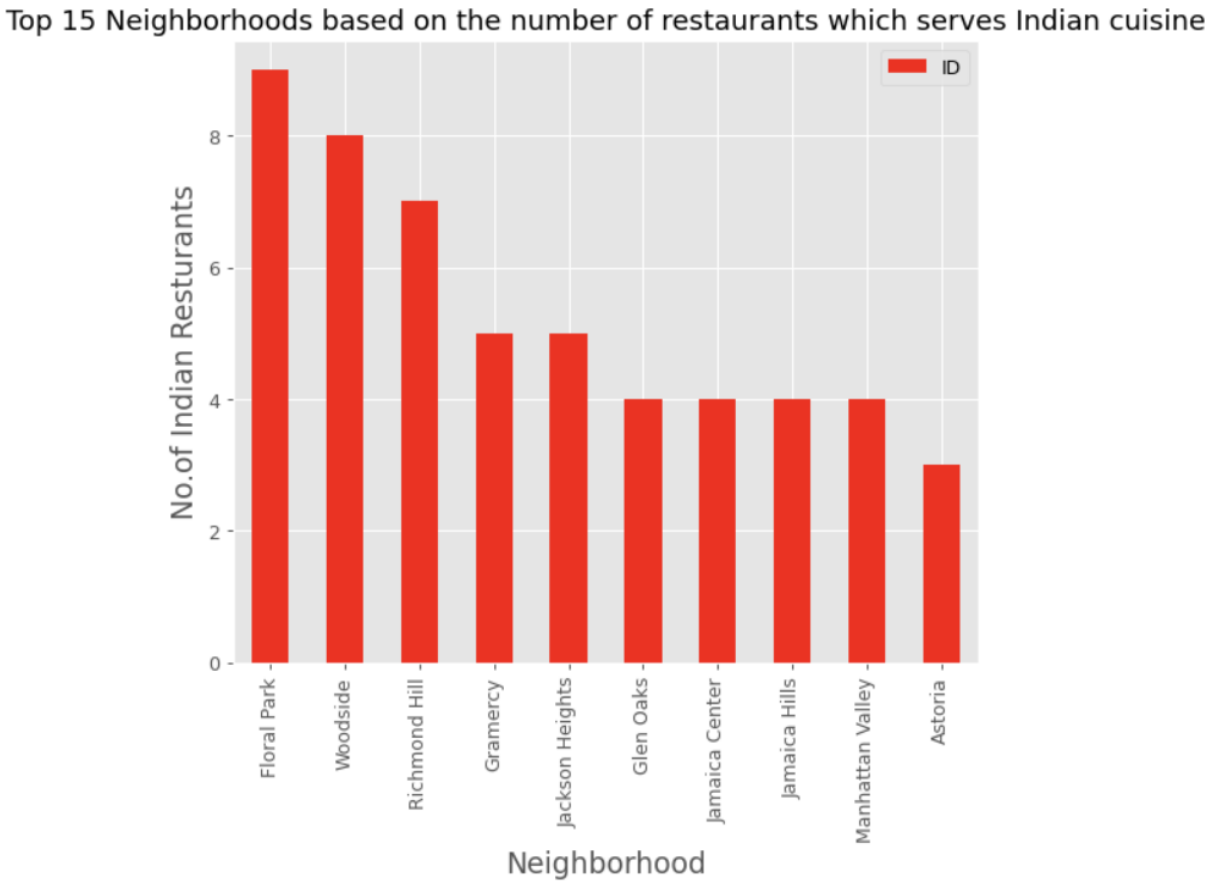
3.2 Boroughs vs. Indian cuisine restaurants – New York

The Queens seems to have the highest number of Indian cuisine restaurants in the New York City and Bronx seems to have the least number of Indian restaurants in New York city, this is just something to be with the number of restaurants and we have so much to explore based on ratings.



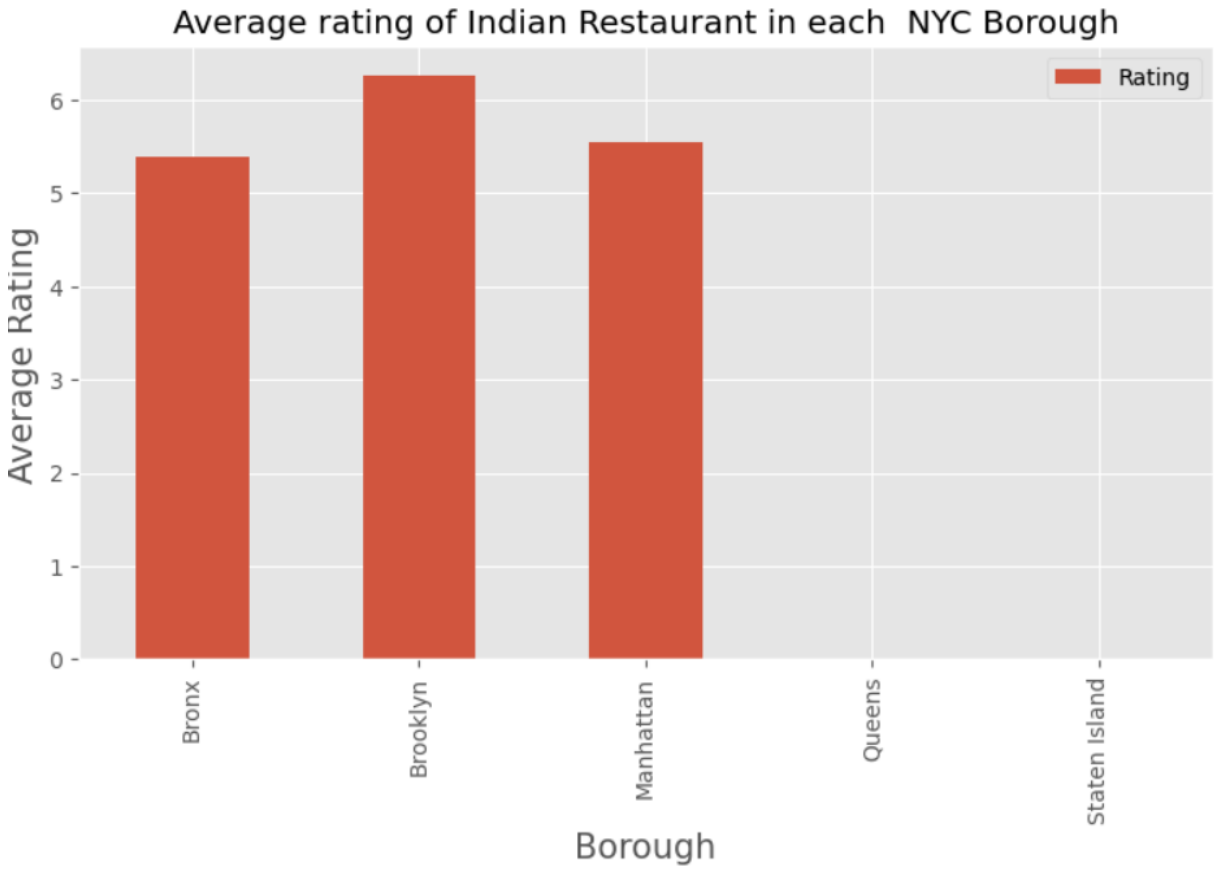
3.3 Top Neighborhoods based on number of Indian restaurants

Floral park neighborhood seems to have highest number of Indian cuisine restaurants in New York. And Richmond hill and Woodside both holds the second position in the numbers. Other 12 top neighborhoods are for your eyes to see



3.4 Boroughs based on Average rating of Indian restaurants

Manhattan seems to hold the first place with the highest average rating of Indian restaurants, and Staten Island seems to go to the last position in this regard, Bar chart below will show details about others:



Recommendations

Midtown or Tribeca in Manhattan would be the best choice to start an Indian restaurant given that:

1. It is the third most populous urban area in New York City (NYC).
2. It has a population density of 27,826 people per square km, highest of any borough in the United States.
3. It has the second highest Asian ethnic minority population in NYC.
4. It has some of the top-rated Indian restaurants located in that area (i.e., ready-made customer base).