Coursera Capstone

IBM Applied Data Science Capstone

Opening a new Authentic Indian Restaurant in New York

By: Omprakash Bharti

April 2021

1. Discussion and Background of the Business Problem:

Indian Restaurants

Introduction Section

This final project explores the best locations for Indian restaurants throughout the New York. New York is a major metropolitan area with more than 8.4 million (Quick Facts, 2018) people living within city limits. New York City is the largest city in the United States with a long history of international immigration. People came from many parts of the world. According to the 2007 American Community Survey estimates, New York City is home to approximately 315,000 people from the Indian subcontinent. With its diverse culture, comes diverse food items. There are many restaurants in New York City, each belonging to different categories like Chinese, Indian, French, etc.

Target Audience

- Business personnel who wants to invest or open a restaurant.

- The freelancer who loves to have their own restaurant as a side business.

- Finding the best location for opening a restaurant.

- Budding Data Scientists, who want to implement some of the most used.

- Exploratory Data Analysis techniques to obtain necessary data, analyze it and, finally be able to tell a story out of it.

Problem Statement

- What is the best location for an Indian restaurant in New York City?

- In what Neighborhood should I open an Indian restaurant to have the best chance of being successful?

Data Section

For this project we need the following data:
1. New York City data that contains Borough, Neighborhoods along with their latitudes and longitudes
   •   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.
2. Indian restaurants in neighborhood of New York city.
   •   Data Source: Foursquare API
   •   Description: By using this API we will get all the venues in the Queens neighborhood. We can filter these venues to get only Indian restaurants.
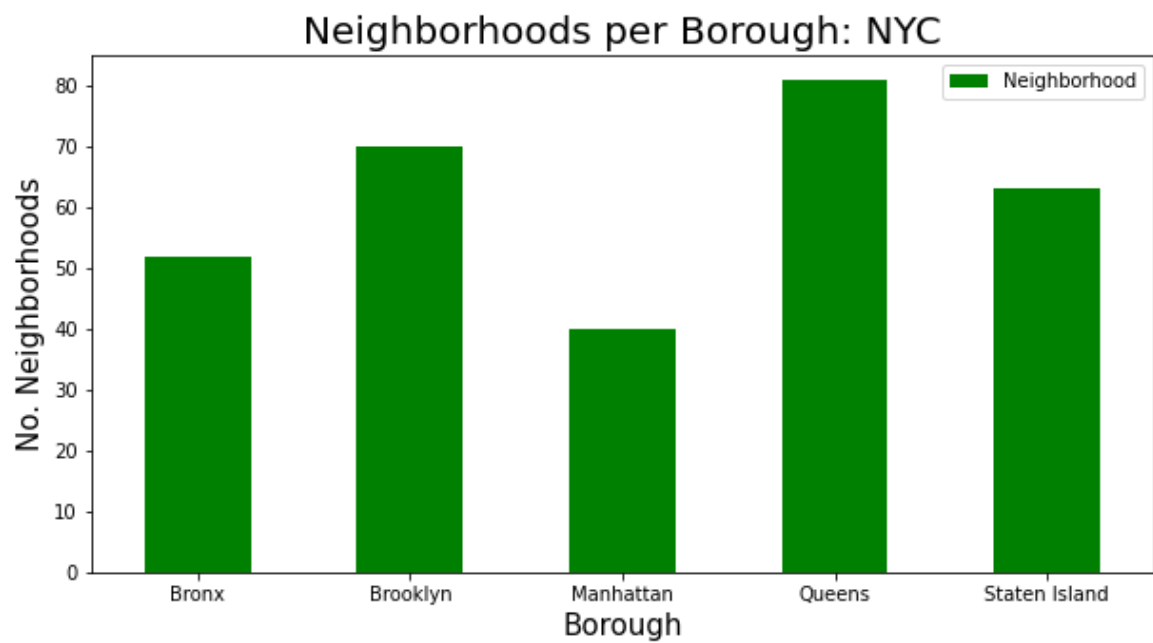
Approach

•   Collect the New York city data from https://cocl.us/new_york_dataset
•   Using Foursquare API we will get all venues for each neighborhood.
•   Filter out all venues which are Indian Restaurants.
•   Data Visualization and some statistical analysis.
•   Analyzing using Clustering (Specially K-Means):
   1. Find the best value of K
   2. Visualize the neighborhood with several Indian Restaurants.
•   Compare the Neighborhoods to Find the Best Place for Starting up a Restaurant.
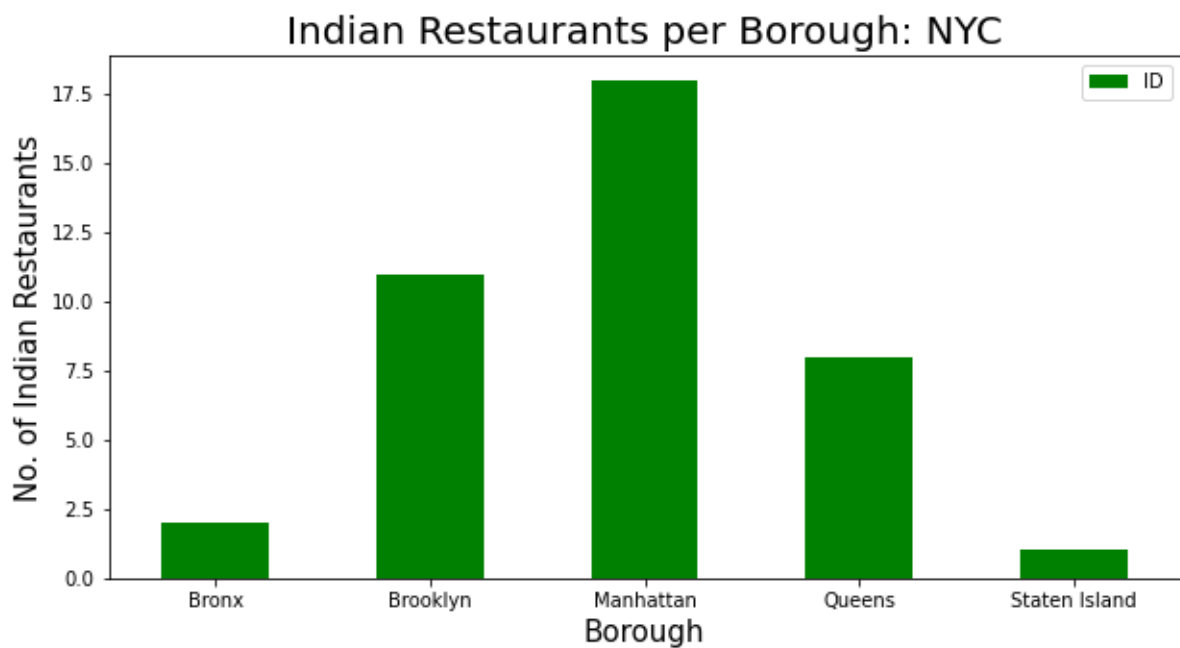•   Inference From these Results and related Conclusions

Results

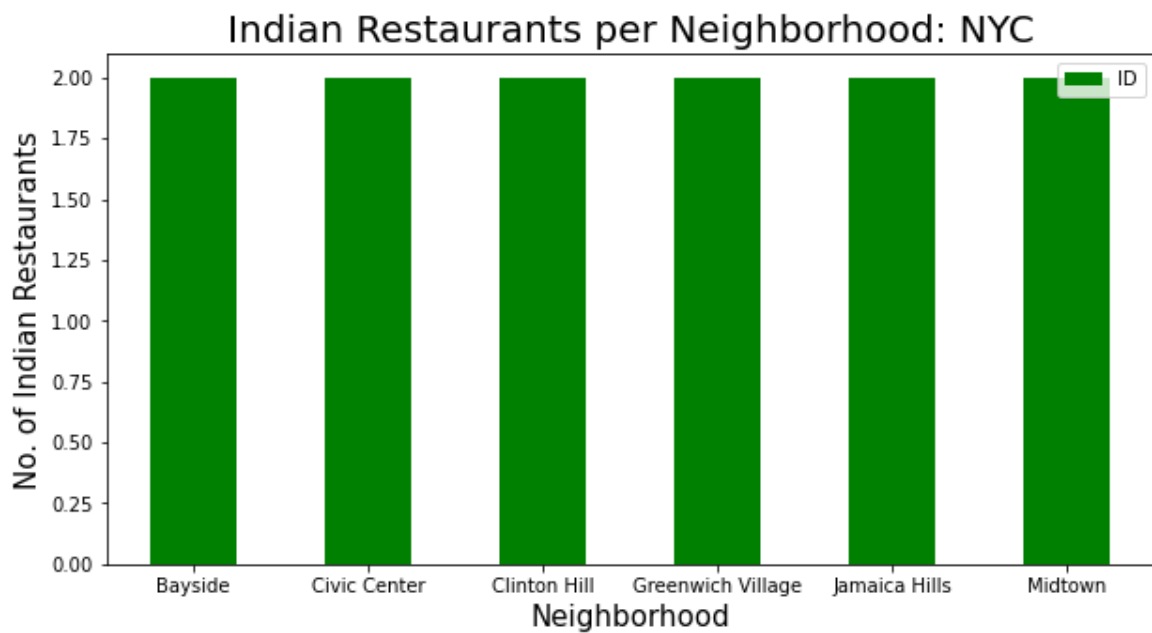The results of our analysis showed below:

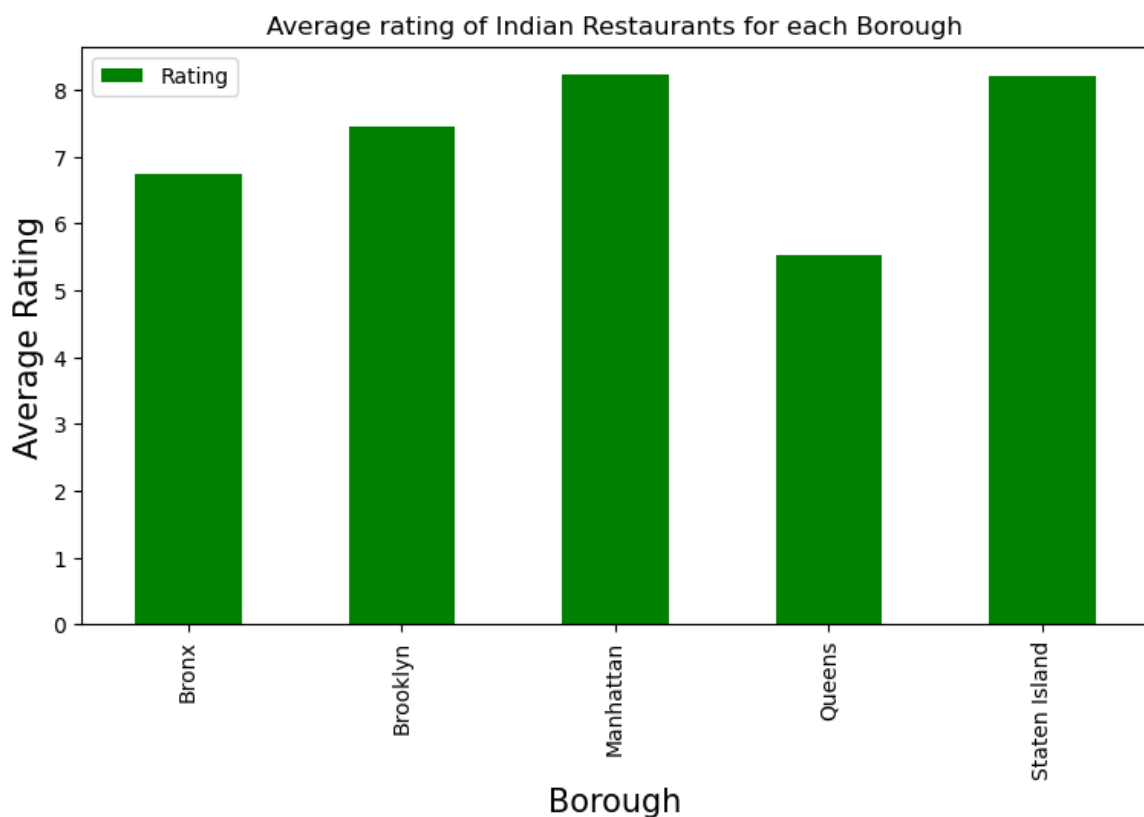1. We see that Queens has the highest number of Neighborhoods.

Neighborhoods per Borough: NYC

2. Although Manhattan had the least number of neighborhoods, it does have the highest number of Indian restaurants.

Indian Restaurants per Borough: NYC

3. The all neighborhood in Manhattan has the highest number of Indian Restaurants.

4. Manhattan and Staten Island has the highest average rating for Indian Restaurants. Brooklyn goes right behind the Manhattan with the average rating of its Indian restaurants only a little lower than Manhattan and Staten Island. Two neighborhoods with the highest average rating for Indian Restaurants are Chelsea and Tribeca.

	Neighborhood	Average Rating
0	Astoria	8.2
3	Boerum Hill	8.2
6	Carnegie Hill	8.1
7	Chelsea	8.9
10	Financial District	8.0
11	Flatiron	8.3
12	Fort Greene	8.6
15	Greenwich Village	8.6
21	Midtown	8.5
22	New Dorp	8.2
23	Noho	8.8
24	North Side	8.4
25	Prospect Lefferts Gardens	8.8
27	Sutton Place	8.3
28	Tribeca	9.1
31	West Village	8.5

Top Sixteen neighborhoods with Indian Restaurants with the best average rating

Discussion Section

Based on the results of our analysis, I would state that Manhattan and Staten Island are the best locations for Indian cuisine in NYC. To have the best shot of success, I would open an Indian restaurant in Staten Island. Staten Island has multiple neighborhoods with average ratings exceeding 8.0 on a scale of 1.0 to 10.0 and has a smaller number of Indian restaurants than Manhattan, making competition easier. In addition, we should keep in mind, that real estate prices in Staten Island are much cheaper than in Manhattan. Particularly, I would recommend considering opening an Indian Restaurant either in Chelsea or Tribeca, because both of these neighborhoods have the highest rating for Indian restaurants.

Limitations and suggestions for future researches

All of the above analysis is depended on the accuracy of Four Square data.

Besides, during this project, we used a free Sandbox Tier Account of Foursquare API that goes with limitations as to the number of API calls and results returned. To get better results, future research work and more comprehensive analysis could consider using a paid account to bypass these limitations as well as incorporating data from other external databases.

Conclusions

In the project we have gone through the process of identifying the business problem, specifying the data required, extracting, and preparing the data, performing data analysis, and lastly providing recommendations to the investors/developers. During the project, we applied different data science methods and instruments to get the answer to our main question: "Where in the City of New York, should the investor open an Indian Restaurant?" The findingsof this project will help the relevant investor better understand the advantages and disadvantages of different New York neighborhoods/boroughs in terms of opening an Indian restaurant.