CAPSTONE PROJECT - THE BATTLE OF NEIGHBORHOODS
APPLIED DATA SCIENCE CAPSTONE PROJECT

Indian Restaurant in New York

By: Kanika Sehgal

Introduction

- New York City (NYC), often simply called New York, is the most populous city in the United States. With an estimated 2019 population of 8,336,817 distributed over about 302.6 square miles (784 km2), New York City is also the most densely populated major city in the United States.

- The restaurant industry is essential to New York City's social and economic fabric. From small, family-owned restaurants and food carts to four-star world-famous establishments, the City abounds with restaurants that offer cuisines from every corner of the globe. Since March 2020, the restaurant industry has been hit very hard by the COVID-19 pandemic. Mandatory closures, stay-at-home and social distancing orders, the onset of a severe economic recession, and travel restrictions have resulted in unprecedented upheaval for the industry. Due to these situations the Indian population living in New York is neither able to get food from their culture nor they could travel back to their country. As part of this project, we need to explore the best locations for Indian restaurant throughout the city of New York for the owner such that it could reach to the target customer and could have great success and consistent profit despite the pandemic situation.

- However, as with any business, opening a new restaurant will require serious considerations specifically due to pandemic situation.

Business Problem

Using Data Science methodology and instruments such as Data Analysis and Visualization, this project aims to provide solutions to answer the business question:

- Where in the city of New York, should the investor open an Indian Restaurant?

- Which areas have large number of Indian Restaurant Market?

- Which all areas have a smaller number of Restaurants?

- Which areas have highest population of Indian customers?

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

- What places are having best restaurant in New York?

Data

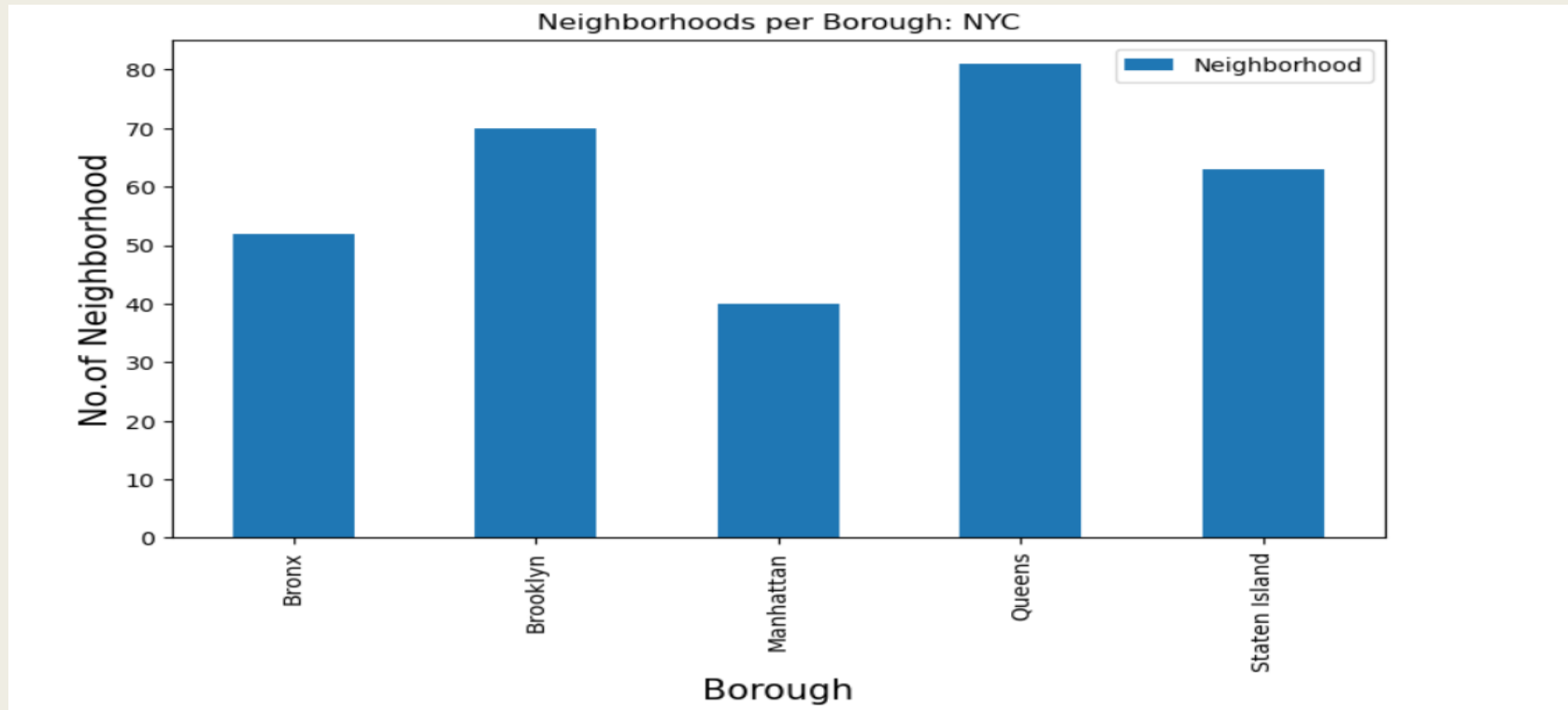To solve the problem, we will need the following data:

- New York City data containing the neighborhoods and boroughs.

- Latitude and longitude coordinates of those neighborhoods. This is required to plot the map and get the venue data.

- Venue data, particularly data related to restaurants. We are going to use this data to perform further analysis of the neighborhoods.

- Python packages and Dependencies:
  - *Pandas - Library for Data Analysis*
  - *NumPy – Library to handle data in a vectorized manner*
  - *JSON – Library to handle JSON files*
  - *Geopy – To retrieve Location Data*
  - *Requests – Library to handle http requests*
  - *Matplotlib – Python Plotting Module*
  - *Folium – Map rendering Library*

Methodology

- New York City data containing the neighborhoods and boroughs will be obtained from the open data source: https://cocl.us/new_york_dataset.

- Geographical coordinates of the neighborhoods (latitude and longitude) using Python Geocoder package.

- Finally, we will use Foursquare API to get the venue data for the neighborhoods defined at the previous step.

- This project will require using of many data science skills, from web scrapping (open source dataset), working with API (Foursquare), data cleaning, data wrangling, to map visualization (Folium).
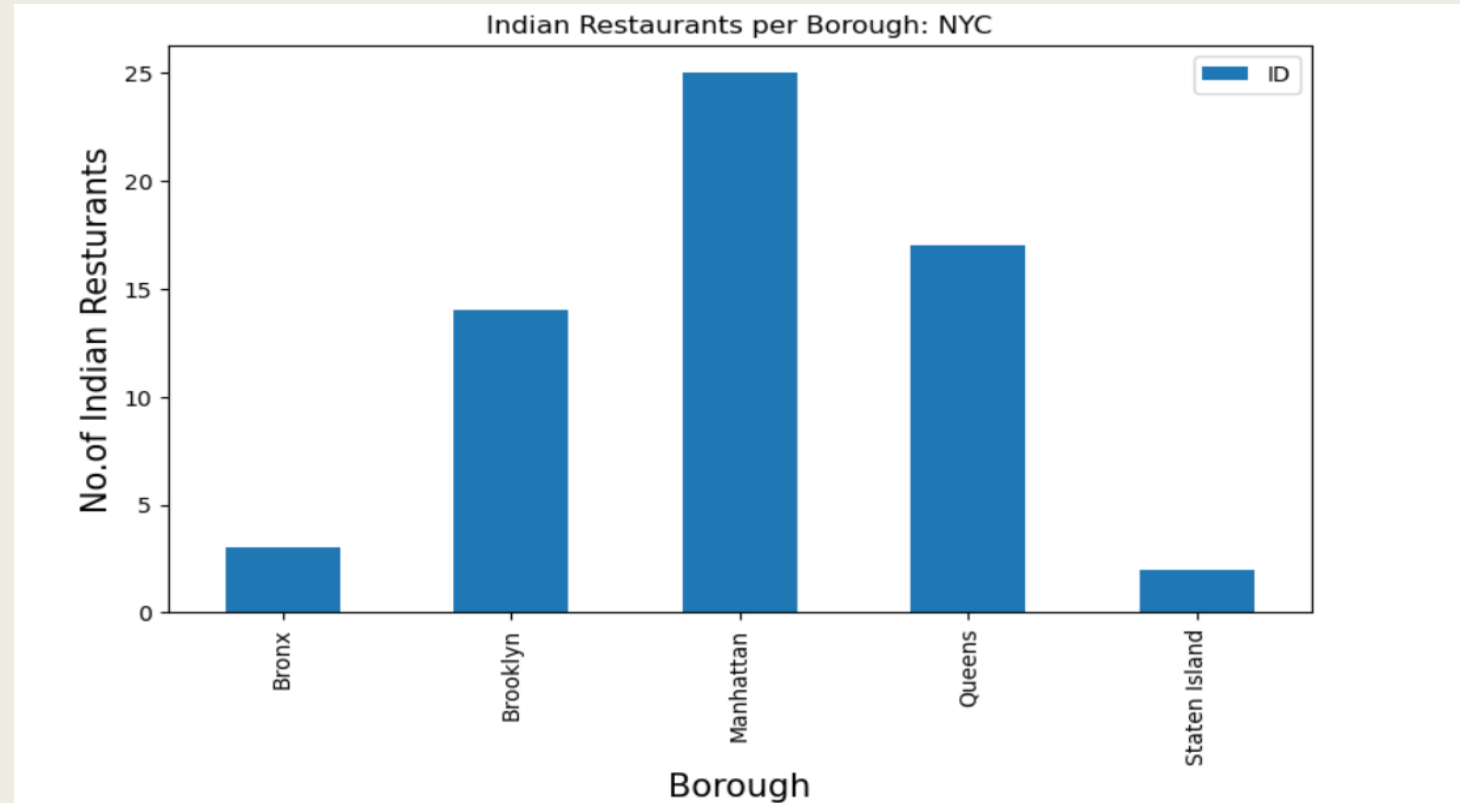
Results(1)

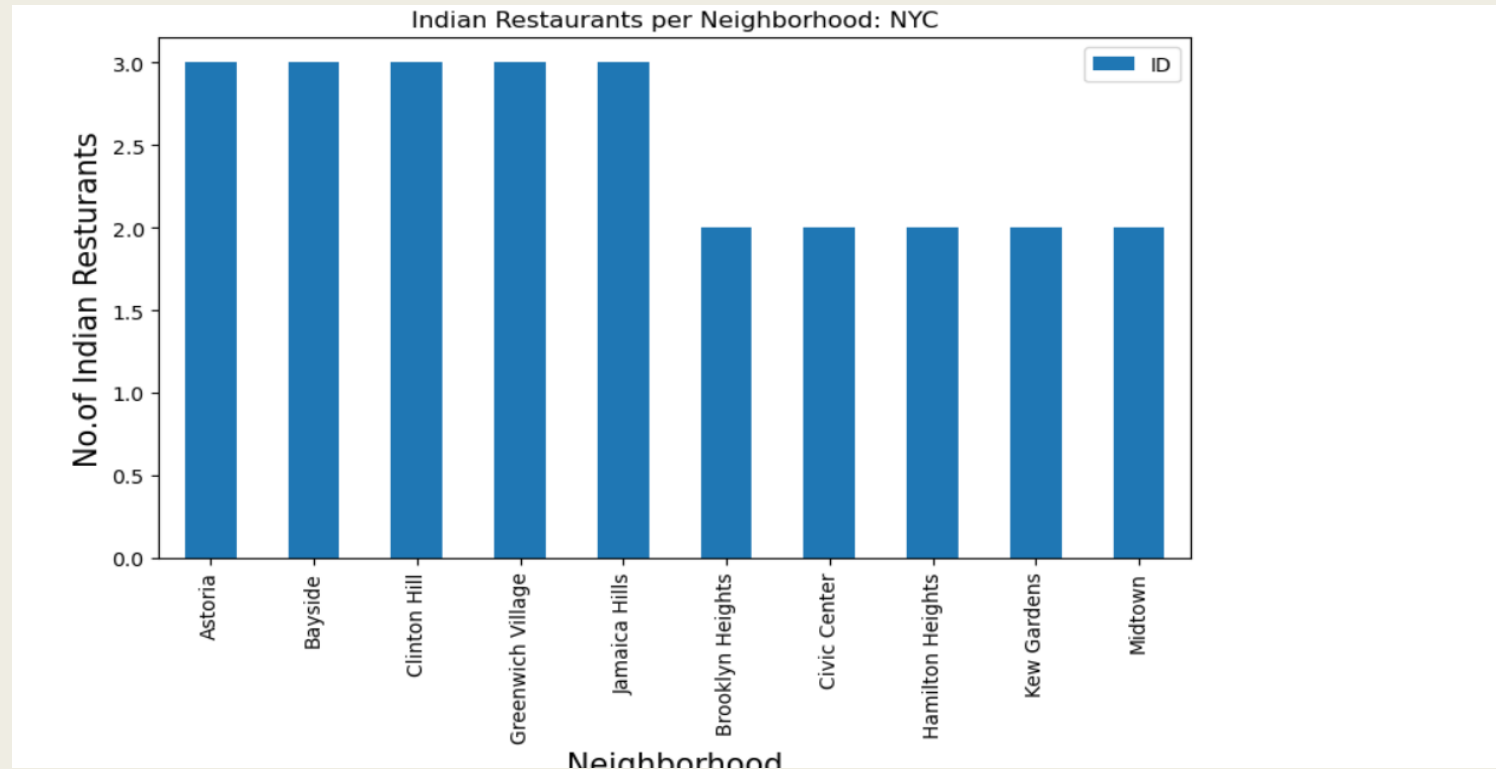■ Queens has the highest number of Neighborhoods in New York City.

Neighborhoods per Borough: NYC

A bar chart titled "Neighborhoods per Borough: NYC" with y-axis labeled "No.of Neighborhood" (0 to 80) and x-axis labeled "Borough". Bars: Bronx ~52, Brooklyn ~70, Manhattan ~40, Queens ~81, Staten Island ~63. Legend: Neighborhood.

Results(2)

■ Manhattan has the highest number of Indian restaurants.

Results(3)

- Astoria, Bayside, Clinton Hill, Greenwich Village and Jamaica Hills have the highest number of Indian Restaurants.

Results(4)

- Restaurant with maximum likes:

```
Borough                        Manhattan
Neighborhood                     Midtown
ID               49d91c12f964a520015e1fe3
Name              The Kati Roll Company
Likes                             834.0
Rating                              8.7
Tips                              256.0
Name: 21, dtype: object
```

- Restaurant with maximum tips:

```
Borough                        Manhattan
Neighborhood                     Midtown
ID               49d91c12f964a520015e1fe3
Name              The Kati Roll Company
Likes                             834.0
Rating                              8.7
Tips                              256.0
Name: 21, dtype: object
```
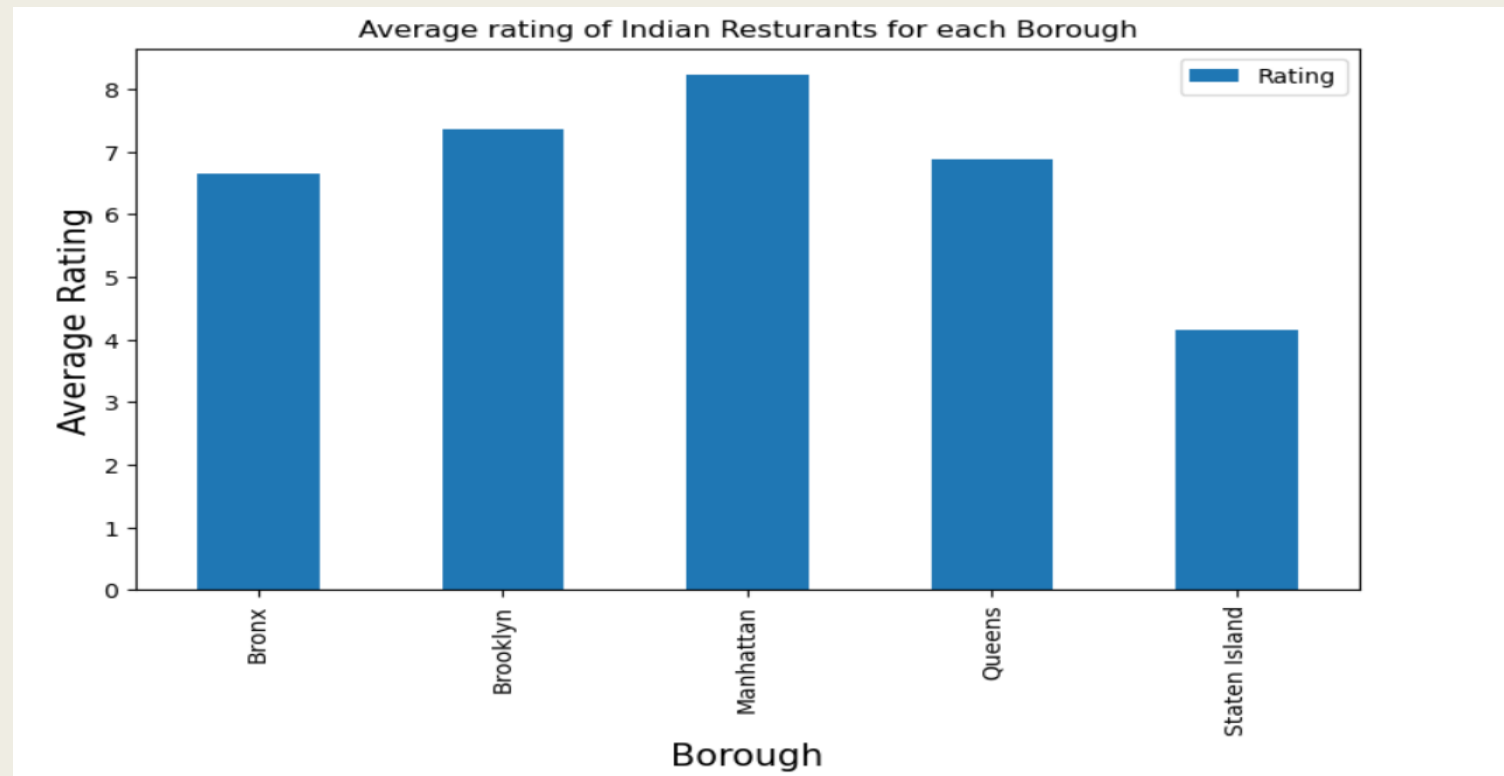
- Restaurant with maximum ratings:

```
Borough                        Manhattan
Neighborhood                      Tribeca
ID               4bbb9dbded7776b0e1ad3e51
Name               Tamarind TriBeCa
Likes                             602.0
Rating                              9.1
Tips                              152.0
Name: 27, dtype: object
```
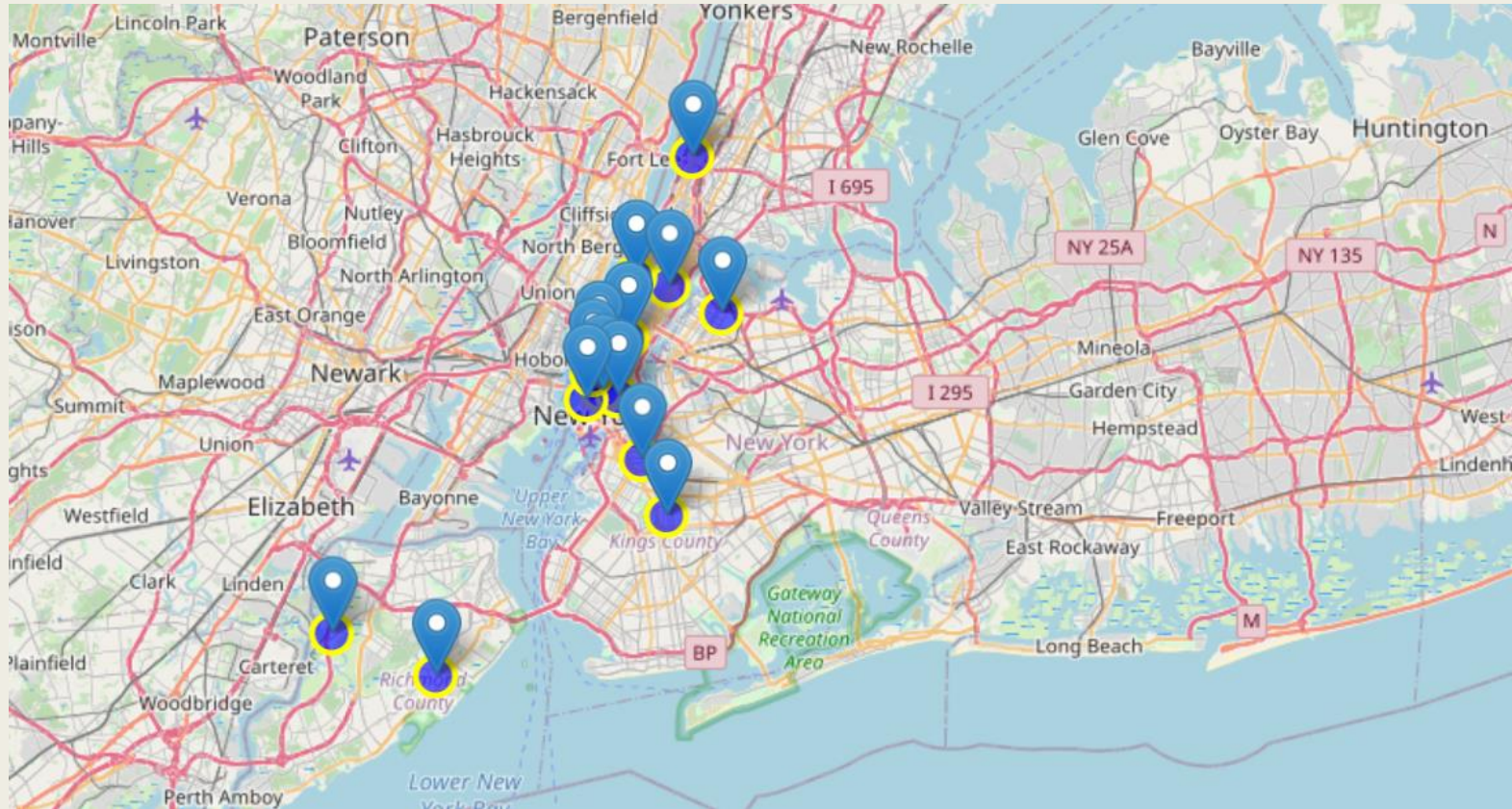
Results(5)

■ Average rating of Indian restaurants per Borough.

Results(6)

■ Neighborhoods with top 10 average rating.

	Neighborhood	Average Rating
29	Tribeca	9.100000
25	Noho	8.800000
9	Chelsea	8.700000
17	Greenwich Village	8.666667
31	Upper West Side	8.600000
13	Fort Greene	8.600000
22	Midtown	8.550000
33	West Village	8.500000
0	Astoria	8.500000
24	New Dorp	8.300000

Results(7)

- Neighborhoods with average rating greater than 8.

Conclusion

Based on the results of our analysis:

- Manhattan is the best location for Indian cuisine in NYC.

- To have the best shot of success, I would open an Indian restaurant in Manhattan.

- Manhattan has multiple neighborhoods with average ratings exceeding 8.0 on a scale of 1.0 to 10.0.

- Tribeca and NoHo have the highest rating for Indian restaurants.

- Staten Island is least considered for Indian Cuisine.

- The Kati Roll Company in Manhattan is the best for Indian food based on its likes and tips.