The Battle of Neighborhoods

Coursera Capstone Final Presentation

By, Sumegha Lazarus

Background

Introduction/Business Problem

In this project we will try to find best locations in New York City for Indian Cuisine. Specifically, this report will be targeted to people interested in eating out in an **Indian restaurant** or want to live in vicinity of an **Indian restaurant** in **New York city**.

Since there are lots of Indian restaurants in New York we will try to detect **locations that have average rating >9**.

We will use our data science powers to generate a few most promising restaurants based on these criteria, so that people can find the following:

- Best location in New York City for Indian Cuisine
- Potential areas to open an Indian Restaurant
- Areas that lack Indian Restaurants
- Areas to stay in vicinity of Indian Restaurants

Data Acquisition

New York City data

- Source: https://cocl.us/new_york_dataset
- Description: This data set contains list of boroughs, neighborhoods with their latitude and longitude. Used to explore various neighborhoods of New York city.

Indian restaurants data

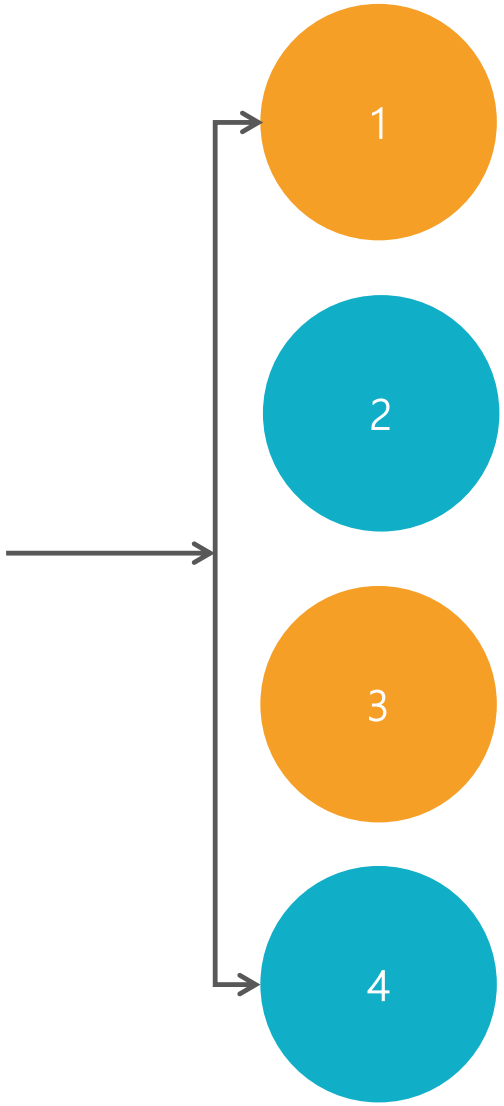
- Source: Foursquare API
- Description: Used to obtain number of restaurants, their type, location in every neighborhood, rating, tips and like count. Filtering these will fetch only Indian restaurants.

GeoSpace data

- Source: https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm
- Description: Used to get the New York borough boundaries and to visualize choropleth map.

Factors

Based on definition of our problem, factors that will influence our decision



1. number of boroughs and neighborhoods in each borough in New York
2. number of existing Indian restaurants in each neighborhood
3. number of top neighborhoods based on average ratings of restaurants in that neighborhood
4. number of top boroughs based on average ratings of restaurants in that borough

Methodology

1. Define Functions

Following functions are defined to extract and process the data for further analysis.

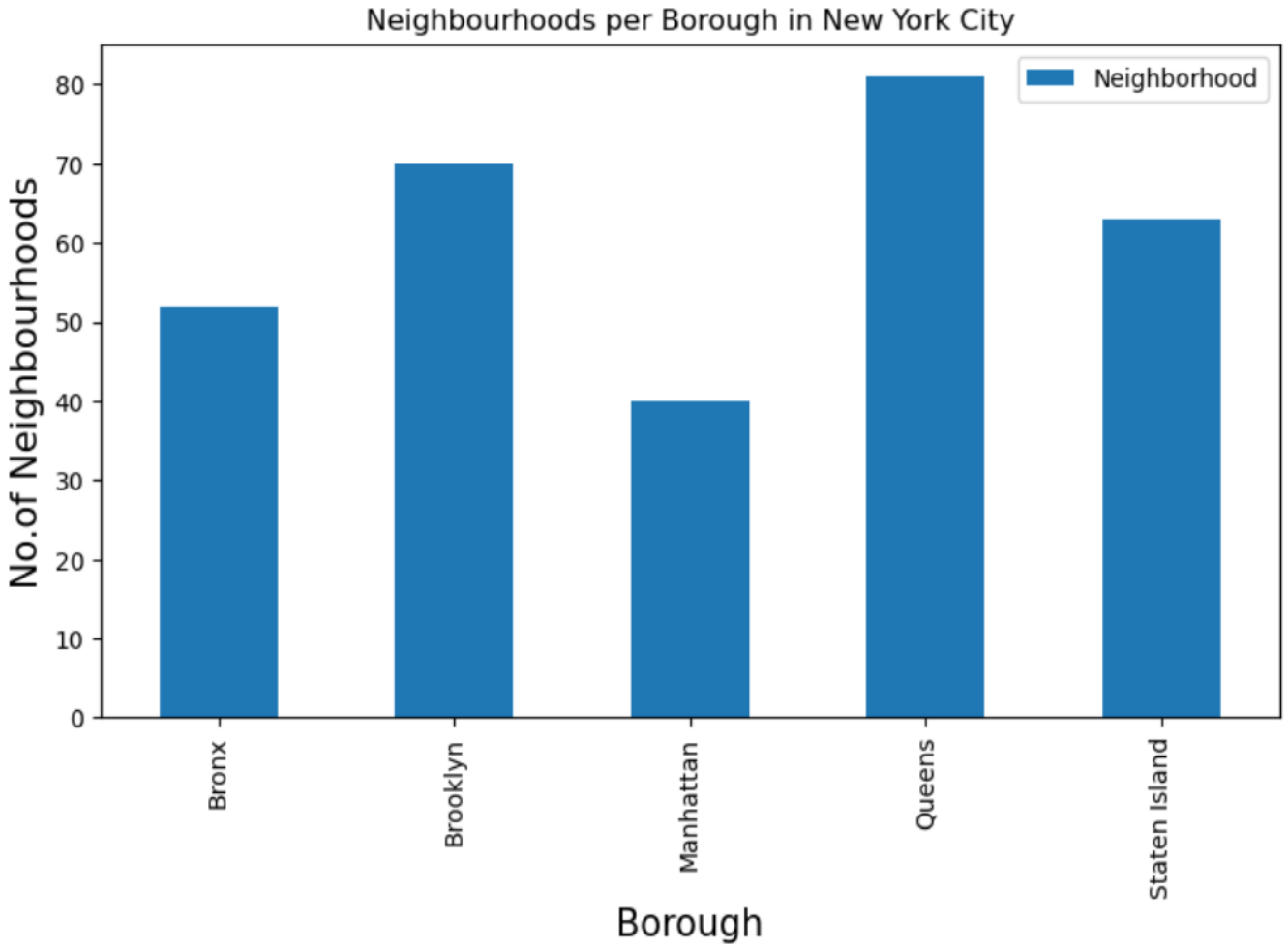
Function	Description
1. geo_location(address)	Using geopy library, this function returns the latitude and longitude for a given location.
2. get_venues(lat,lng)	Using FourSquare API, this function returns defined number of venues & it's requested properties within defined perimeter for a given latitude and longitude.
3. get_venue_details(venue_id)	Using FourSquare API, this function returns ratings, count of likes, count of tips for provided venue. This function will be used for ranking.
4. get_new_york_data()	This function returns New York city data such as Boroughs, Neighborhoods along with their latitude and longitude.

Methodology

2. Number of neighborhoods per borough

To get the geographical perimeter and scope, the first step is to find number of neighborhoods per borough in the New York city. This is fetched using get_new-york_data() function, transferred it to dataframe and visualized using matpotlib.pylot instance.

Inference:
From the adjacent chart, it is quite evident that out of 5 boroughs, Queens borough has the highest number of neighborhoods.



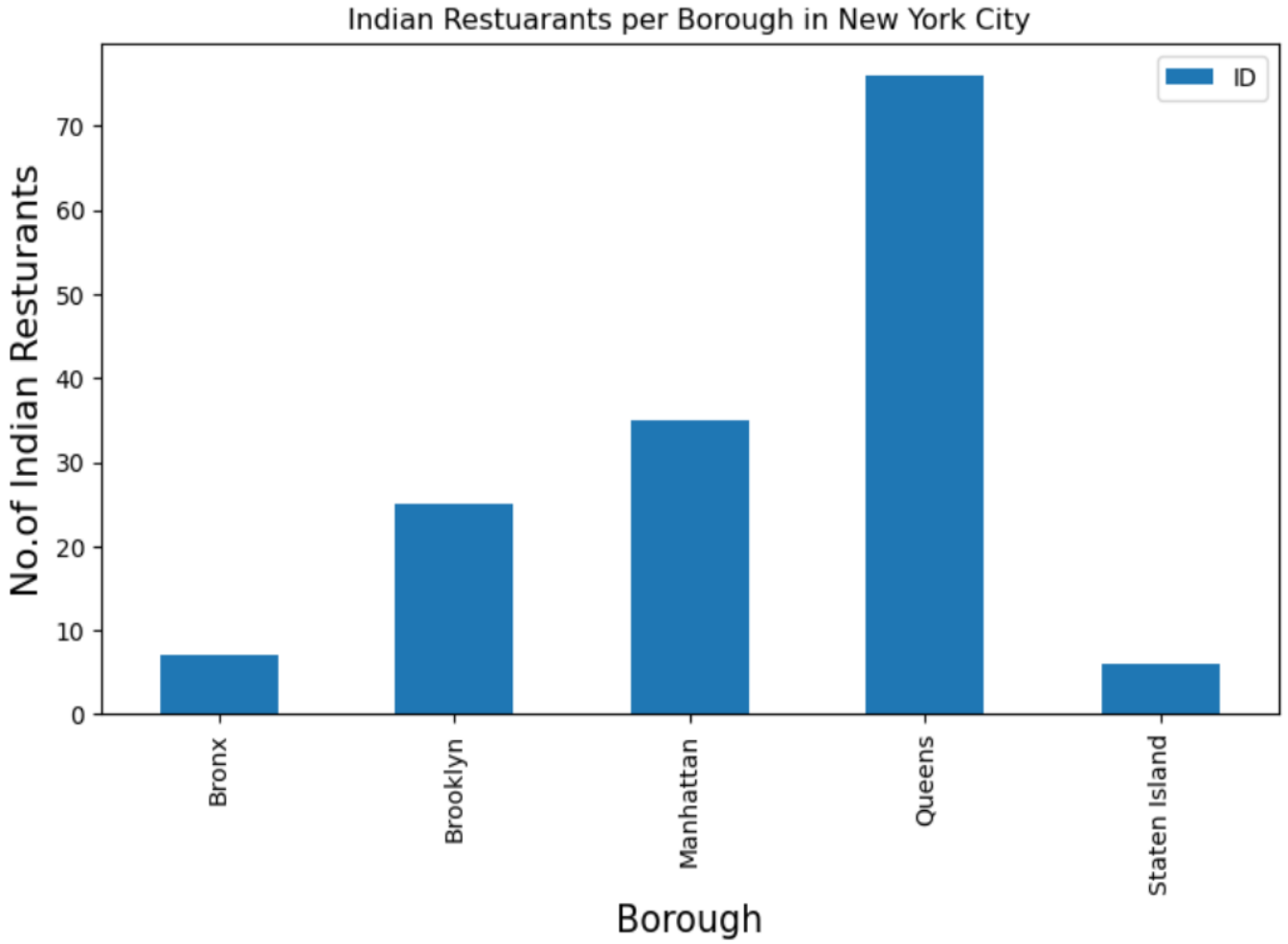
Methodology

3. Number of restaurants per neighborhoods

Using get_venues() function, total number of existing Indian restaurants per neighborhood are fetched.

Finding: There are 149 restaurants across NY city.

Inference:
From the adjacent chart, out of 149 restaurants across 5 boroughs, Queens borough has the highest number of Indian Restaurants.



Methodology

4. Ranking

Based on the data of 149 Indian restaurants that are present in NY city, here are the top restaurants graded based on maximum likes, ratings and tips. Highlighted cells are the maximum numbers in their respective category i.e. rating/like/tips.

Restaurant	Borough	Neighborhood	Rating	Likes	Tips
Tamarind TriBeCa	Manhattan	Tribeca	9	599 ↑	149
Seva Indian Cuisine	Queens	Astoria	9.1 ↑	240	133
Dhabha	Manhattan	Gramercy	8.3	495	164↑

Indian restaurants with maximum average rating > 9

Borough-wise

Borough	Average Rating
Manhattan	8.320000
Brooklyn	7.732000
Queens	6.221053
Bronx	5.685714
Staten Island	5.300000

Inference: Manhattan tops

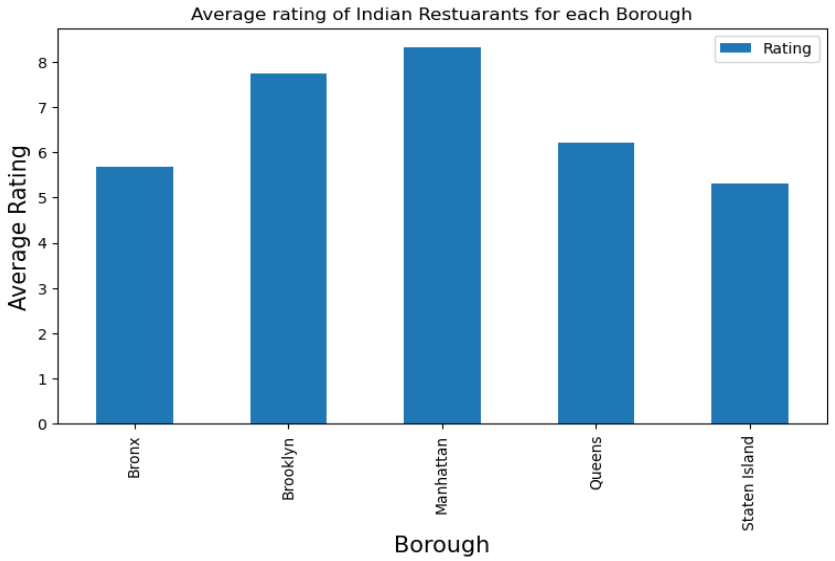
Neighborhood-wise

Neighborhood	Average Rating
Blissville	9.1
Sunnyside	9.1
Greenwich Village	9.0
Tribeca	9.0
Civic Center	9.0
West Village	8.9
Fort Greene	8.8
Noho	8.8
Astoria	8.7
Chelsea	8.6

Inference: Blissville & Sunnyside tops

Methodology

4. Ranking - Visualization



Borough	Neighborhood	Latitude	Longitude	Average Rating	Label
Queens	Blissville	40.737251	-73.932442	9.1	Blissville, Queens(9.1)
Queens	Sunnyside	40.740176	-73.926916	9.1	Sunnyside, Queens(9.1)
Staten Island	Sunnyside	40.612760	-74.097126	9.1	Sunnyside, Staten Island(9.1)
Manhattan	Civic Center	40.715229	-74.005415	9.0	Civic Center, Manhattan(9.0)
Manhattan	Greenwich Village	40.726933	-73.999914	9.0	Greenwich Village, Manhattan(9.0)
Manhattan	Tribeca	40.721522	-74.010683	9.0	Tribeca, Manhattan(9.0)

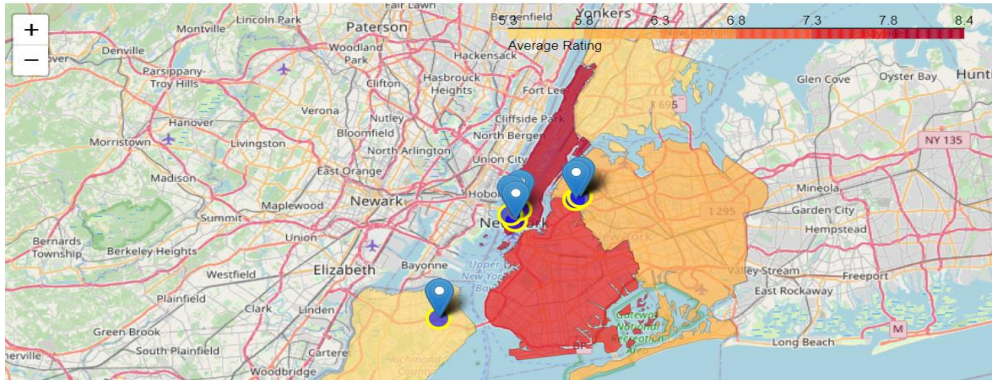
Overall Inference:

Though Manhattan tops the maximum ratings borough-wise however, Blissville and Sunnyside in Queens borough tops the ratings in neighborhood-wise data.

Hence,

1. Best location for Indian cuisine is Blissville and Sunnyside in Queens
2. Manhatten has potential to open new Indian restaurants

Choropleth map



Conclusion

- Best location in New York City for Indian Cuisine - **Blissville and Sunnyside** in Queens Borough with average rating of 9.1
- Potential areas to open an Indian Restuarant - Manhattan Borough as it tops the average rating with 8.3 from all other Boroughs
- Areas that lack Indian Restuarants - **Staten Island**
- Areas to stay in vicinity of Indian Restaurants - **Manhattan** it is! Though Queens has highest number of Indian restaurants, Manhattan scores high in the quality with the score of 9.1

Thank You