

CAPSTONE PROJECT: BATTLE OF THE NEIGHBORHOODS

Purpose

This document provides the details of my final peer reviewed assignment for the IBM Data Science Professional Certificate program – Coursera Capstone.

Introduction

As an Indian student, I wanted to visualize the extent of the spread of the Indian cuisines across the world. As we worked on the data sets which were related to the prime locations, like New York and Toronto, so I chose New York as the city of interest to find out the number of restaurants serving the Indian cuisines and based on the ratings of the visitors of these restaurants taken from the Foursquare API, in order decide the best restaurant amongst all and also to find out the neighborhood in New York city with the highest count of the Indian restaurants.

Demographics of New York City (Source: Wikipedia)

New York City's demographics show that it is a large and ethnically diverse metropolis. It is the largest city in the United States with a long history of international immigration. New York City was home to nearly 8.5 million people in 2014, accounting for over 40% of the population of New York State and a slightly lower percentage of the New York metropolitan area, home to approximately 23.6 million. 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.

Throughout its history, New York City has been a major point of entry for immigrants; the term "melting pot" was coined to describe densely populated immigrant neighborhoods on the Lower East Side. As many as 800 languages are spoken in New York, making it the most linguistically diverse city in the world. English remains the most widely spoken language, although there are areas in the outer boroughs in which up to 25% of people speak English as an alternate language, and/or have limited or no English language fluency. English is least spoken in neighborhoods such as Flushing, Sunset Park, and Corona.

Having read about the diversity of New York, it will be great to find out a what this city has in stock for us to find as an outcome to our problem statement.

We will try to answer these questions:

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

Which place has the highest rated Indian Cuisine?

Where to find a variety of Indian Cuisines (highest count of Indian Restaurants)?

Required Datasets

For this project the following data is used:

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 will be used 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.

Working with the Data:

1. Gather the New York city data from https://cocl.us/new_york_dataset (https://cocl.us/new_york_dataset)
2. Using the Foursquare API, we will find all venues for each neighborhood.
3. Filtering out venues that are 'Indian Restaurants'.
4. Collect rating, tips and like count for each Indian Restaurants using Foursquare API.
5. Using rating for each restaurant, we will sort that data.
6. Visualize the Ranking of neighborhoods.

Data Description:

```
: new_york_data=get_new_york_data()

: new_york_data.head()

:
```

	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

Figure: Loading the New York dataset

```
indian_rest_ny.head()
```

	Borough	Neighborhood	ID	Name
0	Bronx	Woodlawn	4c0448d9310fc9b6bf1dc761	Curry Spot
1	Bronx	Parkchester	4c194631838020a13e78e561	Melanies Roti Bar And Grill
2	Bronx	Spuyten Duyvil	4c04544df423a593ac83d116	Cumin Indian Cuisine
3	Bronx	Concourse	551b7f75498e86c00a0ed2e1	Hungry Bird
4	Bronx	Unionport	4c194631838020a13e78e561	Melanies Roti Bar And Grill

Figure: Indian Restaurants with Neighborhood and Borough

```
# Resturant with maximum Tips
indian_rest_stats_ny.iloc[indian_rest_stats_ny['Tips'].idxmax()]
```

```
Borough          Manhattan
Neighborhood      Gramercy
ID              4a70a75bf964a52016d81fe3
Name             Bhatti Indian Grill
Likes            422
Rating           8.7
Tips            160
Name: 49, dtype: object
```

Figure: Restaurants with Max Tips

```
# Resturant with maximum Rating
indian_rest_stats_ny.iloc[indian_rest_stats_ny['Rating'].idxmax()]
```

```
Borough          Manhattan
Neighborhood      Tribeca
ID              4bbb9dbded7776b0e1ad3e51
Name             Tamarind TriBeCa
Likes            590
Rating           9.1
Tips            148
Name: 40, dtype: object
```

Figure: Restaurants with Max Rating

```
# Restaurant with maximum Likes
indian_rest_stats_ny.iloc[indian_rest_stats_ny['Likes'].idxmax()]
```

```
Borough          Manhattan
Neighborhood      Tribeca
ID               4bbb9dbded7776b0e1ad3e51
Name              Tamarind TriBeCa
Likes             590
Rating            9.1
Tips              148
Name: 40, dtype: object
```

Figure: Restaurants with Max Likes

```
ny_neighborhood_stats.sort_values(['Average Rating'],ascending=False).head(10)
```

	Neighborhood	Average Rating
29	Tribeca	9.1
13	Greenwich Village	8.9
33	West Village	8.8
20	Murray Hill	8.8
3	Chelsea	8.7
5	Clinton Hill	8.7
10	Fort Greene	8.7
12	Gramercy	8.7
24	Prospect Lefferts Gardens	8.6
35	Yorkville	8.5

Figure: Indian Restaurants' average ratings with Neighborhood

```
indian_rest_ny[indian_rest_ny['Neighborhood']=='Floral Park']
```

	Borough	Neighborhood	ID	Name
101	Queens	Floral Park	527ffc0811d2d329d5e49abd	Jackson Diner
102	Queens	Floral Park	4e4e3e22bd4101d0d7a5c2d1	Kerala Kitchen
103	Queens	Floral Park	4b647b56f964a520c4b62ae3	Usha Foods & Usha Sweets
104	Queens	Floral Park	4b787c49f964a5209cd12ee3	Santoor Indian Restaurant
105	Queens	Floral Park	4c0c01e0bbc676b00d6b4cd5	Mumbai Xpress
106	Queens	Floral Park	4c76ff35a5676dcb72671721	Flavor Of India
107	Queens	Floral Park	4df0f39dd4c04d0392c853ea	Sagar Chinese
108	Queens	Floral Park	4cc642ed306e224b5bf2a76c	Shahi Darbar
109	Queens	Floral Park	4ec1b54f4690d2338a359d66	Tabaq Kebab Grill
110	Queens	Floral Park	4e6bfe1c7d8b2c711b17bbe5	Surya sweets and snacks

So Floral Park in Queens has the highest number of Indian Restaurants i.e. 10.

Figure: Indian Restaurants in Flora Park

Observations

Plotted various graphs:

1. No of Neighborhood in NY vs Boroughs
2. No of Indian Restaurants in each Borough vs Boroughs
3. No of Indian Restaurants in each Neighborhood vs Boroughs
4. Average Rating of Indian Restaurants for each Borough vs Boroughs

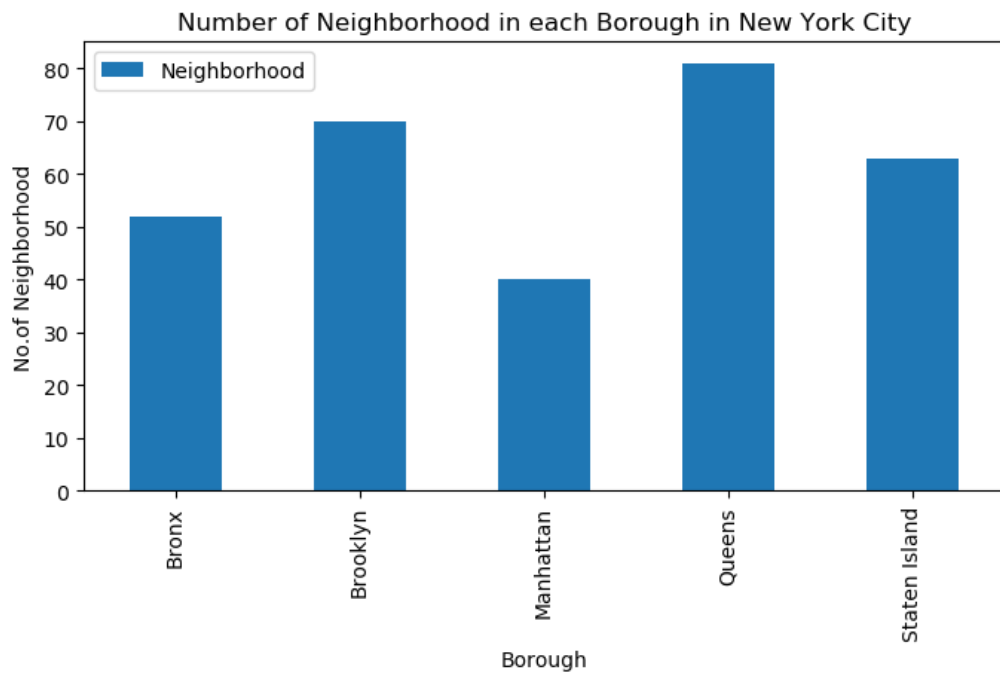


Figure: No of Neighborhood in NY vs Boroughs



Figure: No of Indian Restaurants in each Borough vs Boroughs

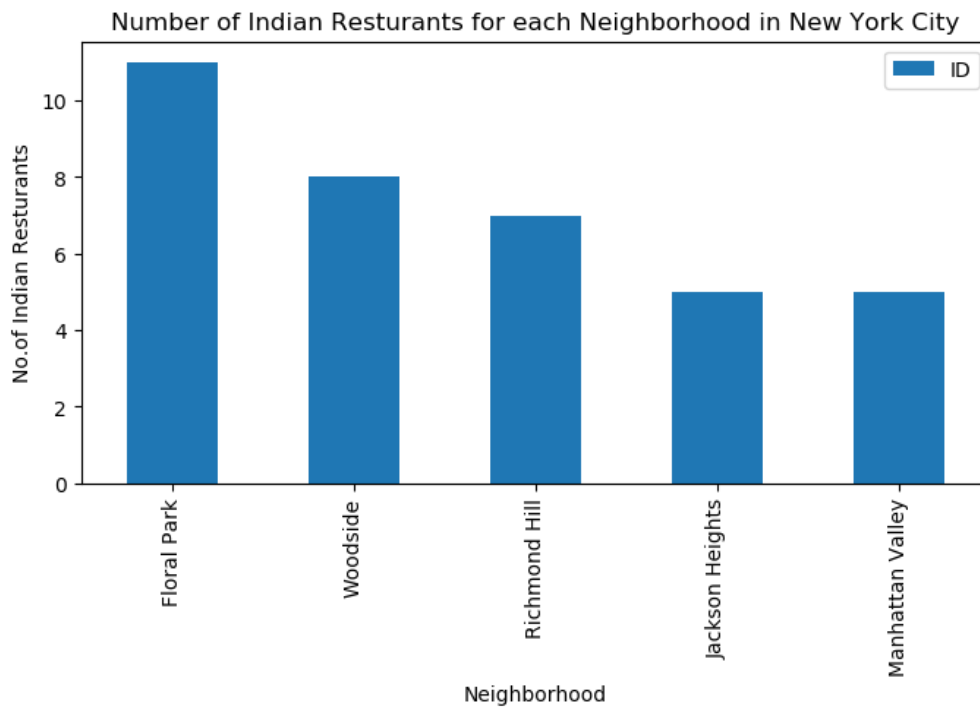


Figure: No of Indian Restaurants in each Neighborhood vs Boroughs

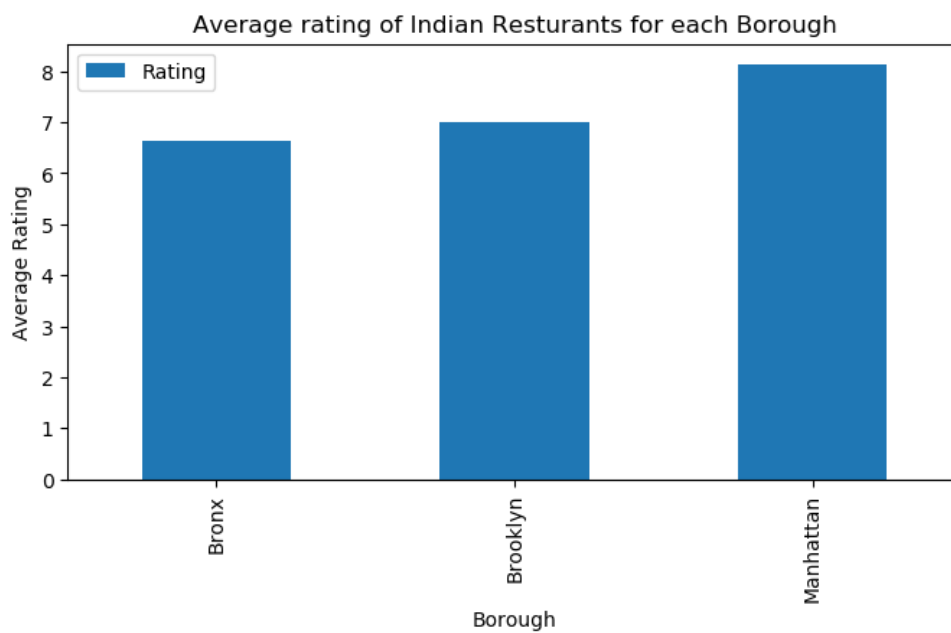


Figure: Average Rating of Indian Restaurants for each Borough vs Boroughs

Discussion and Conclusion

In this notebook, analysis was done to find the best Indian restaurant in New York using the venue recommendations based on Food venue category.

Using Foursquare API, we have collected a good amount of Indian Restaurant recommendations in

New York. Sourcing from the venue recommendations from Foursquare has its limitation; The list of venues is not exhaustive list of all the available venues in the area. Furthermore, not all the venues found in the area have stored ratings. For this reason, the number of analyzed venues is only about 50% of all the available venues initially collected. The results therefore may significantly change, when more information will be collected on those with missing data.

Results

1. Gramercy in Manhattan has received the highest tips.
2. Tribeca in Manhattan has received the maximum likes and ratings.
3. As far as the average ratings are considered than Manhattan has the highest average ratings.

Although Queens has the highest count of Indian restaurants but Manhattan has neighbors with highest ratings.

Limitations

- The ranking is purely on basis of rating of restaurants
- The accuracy of data depends purely on the data provided by Foursquare.

Future Scope

As we used the data set to search for Indian Restaurants, in a similar fashion we can find various Food Centers based on the user's choice, i.e. creating a web application for taking the input from the user and returning the desired search results based on the user preferences.

Thank you.

Apoorv Saxena

email: apoorvsaxena7@gmail.com

LinkedIn: <https://www.linkedin.com/in/apoorvsaxena/>

Created For: **COURSERA IBM Applied Data Science Capstone Project**