The Battle of Neighborhood: Indian Restaurants In New York

1. Introduction

New York City often called as City of New York, is the most populous city in United States. With an estimated 2018 population of 8,398,748 distributed over about 302.6 square miles (784 km2).

New York is often considered one of the greatest cultural hubs in the country, having welcomed immigrants long before the days of Ellis Island. Immigrants now account for one-fifth of the state's total population and make up a staggering 25 percent of its labor force. As workers, business owners, taxpayers, and neighbors, immigrants are an integral part of New York's diverse and thriving communities and make extensive contributions that benefit all.

FACT: One in five New Yorkers is an immigrant.

Indians in the New York City metropolitan region constitute one of the largest and fastest growing ethnicities in the New York City metropolitan area of the United States.

FACT: New York City was home to 64 percent of New York State's Indian residents.

New York City is composed of five boroughs, each of which is a county of the State of New York. The five boroughs–Brooklyn, Queens, Manhattan, the Bronx, and Staten Island–were consolidated into a single city.

Rank	Borough	City	Indian Americans	Density of Indian Americans per square mile	Percentage of Indian Americans in municipality's population
1	Queens (2014)[33]	New York City	144,896	1,326.5	6.2
2	Brooklyn (2012)	New York City	25,270	357.9	1.0
3	Manhattan (2012)	New York City	24,359	1,060.9	1.5
4	The Bronx (2012)	New York City	16,748	398.6	1.2
5	Staten Island (2012)	New York City	6,646	113.6	1.4
	Total (2014)[31]	**New York City**	227,994	753.4	2.7

Food Culture in New York City: Due to such huge number of immigrants in New York, we can see diversity in Food Culture as well. In 2018, there were over 27,000 restaurants in New York.

In this Project, I will be analyzing Indian Restaurants in all the parts of New York City.

2. Data Requirement and its Usage

- New York City's data which contains list of Boroughs, Neighborhoods, latitudes and longitudes.

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

  Usage: This dataset will be used to explore boroughs, neighborhoods, latitude and longitude of new york city.

- To identify Indian restaurants in each neighborhood of new york city.

  Data source: Foursquare API

  Usage: This api is used to identify all venues in each neighborhood of New York city which in turn will be filtered to extract Indian Restaurants only.

- **GeoSpace data**

  Data source: https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm

Usage: Geo Space data will be used to get the Borough boundaries of New York which will help us in visualizing Choropleth Map.

3. Methodology

- Extract New York City Data which specify all boroughs, neighborhoods, latitudes and longitudes.
- Using FourSquare API to get all venues for all neighborhoods.
- Filter venues to extract only Indian Restaurants.
- Find Ratings/Like Counts for each Indian Restaurants.
- Visualize the Ranking of Restaurants using Folium Library.

- Finding Best Indian Restaurants.

Questions that can be answered by this analysis is given by:

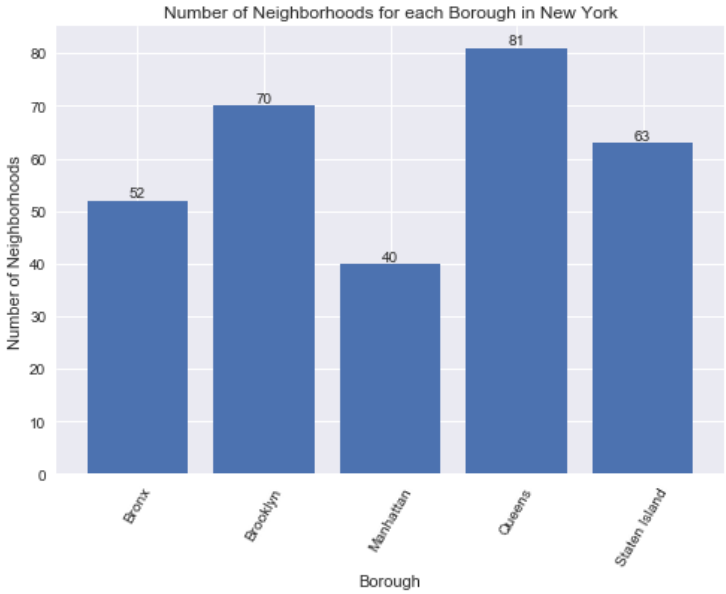
- What is the best location in New York City for Indian Cuisine?
- Which areas have potential Indian Restaurants?
- Which area do not have Indian Restaurants?
- Which is the best place to stay for Indian Cuisine?

4. Evaluation

First we get data about New York City from the link specified in the Data section.

This data contains all Boroughs, neighborhoods, Latitudes and Longitudes. Then we proceed to find out the total number of neighborhoods for each borough.

There are five boroughs in New York and each borough have different number of Neighborhoods.

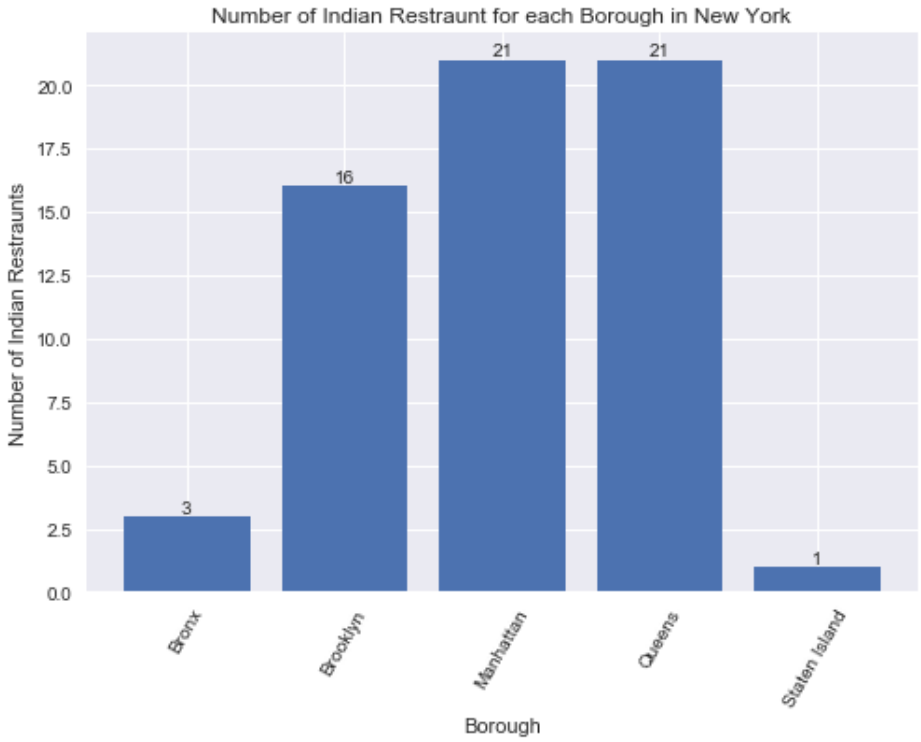


Findings

We have seen from the graph that Queens has highest number of Neighborhoods.

Next, all venues are collected by making call requests to Foursquare api for New York City.

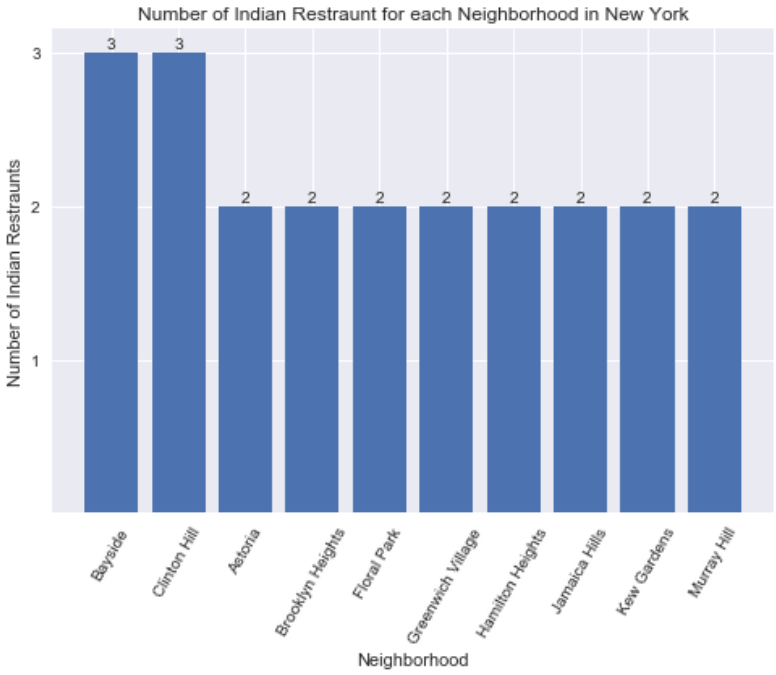
Filter these venues to extract only Indian Restaurants. By plotting these data we came to know about the number of Indian Restaurants in each Borough of New York.



Findings

We noticed that Queens and Manhattan has largest number of Indian Restaurants.

We will get deeper into neighborhoods to see which neighborhood got the most Indian Restaurants.



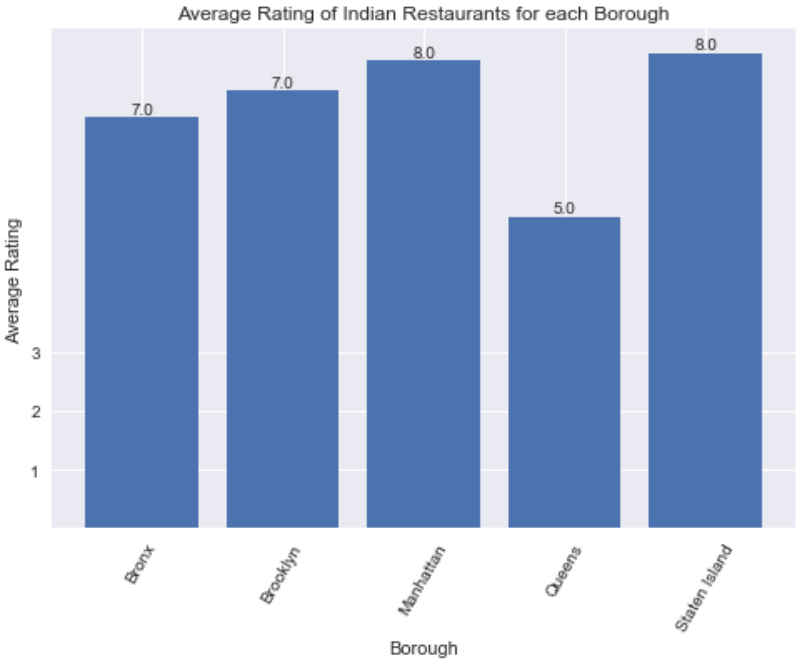
Findings

(Total: 6) We find that Bayside in Queens and Clinton Hill in Broklyn has the highest number of Authentic Indian Restaurants.

Ratings, Likes and Tips of Restaurants

We will consider ratings, likes and tips of all Indian Restaurants. This will help in identifying the best areas and restaurants according to ratings and likes given by their customers.

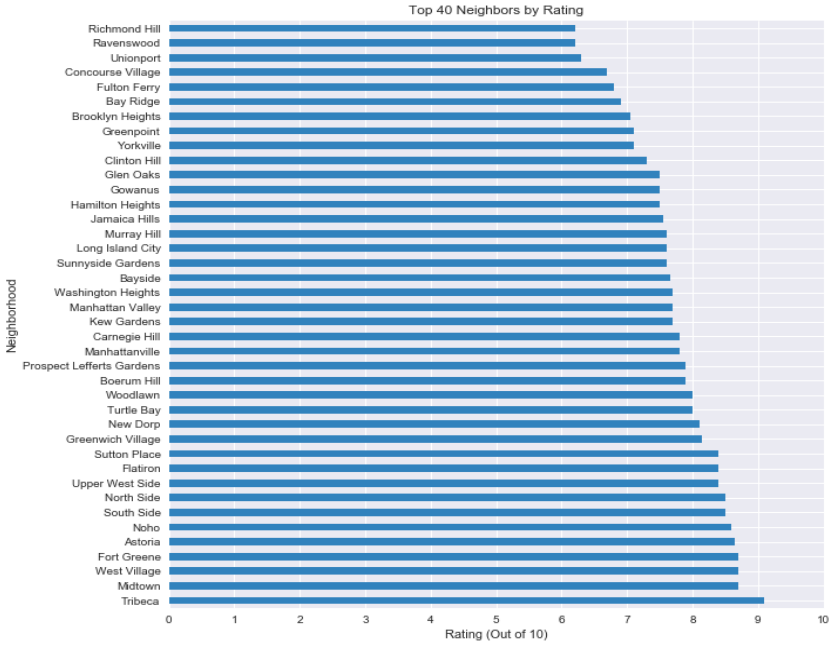
- **Borough with best Average Ratings for their Restaurants**



Findings

We find that Manhattan, Staten Island, Brooklyn has highest average ratings.

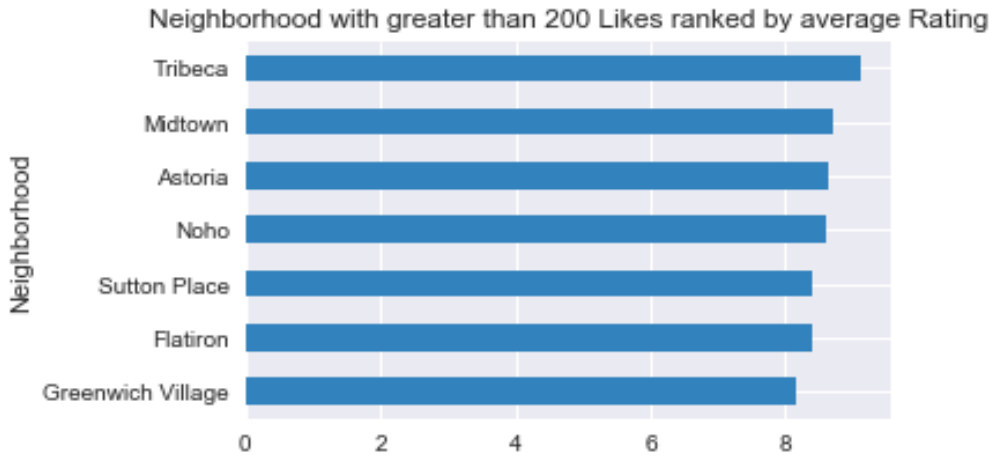
- **Neighborhood with most Likes, Tips and best Ratings for their Restaurants**



Findings

We find that Tribeca, MidTown, Astoria, Noho, South Side, Sutton Place, Flatiron are some of the neighborhood with highest average rating.

- **Neighborhood with more than 200 Likes and highest Average Ratings**

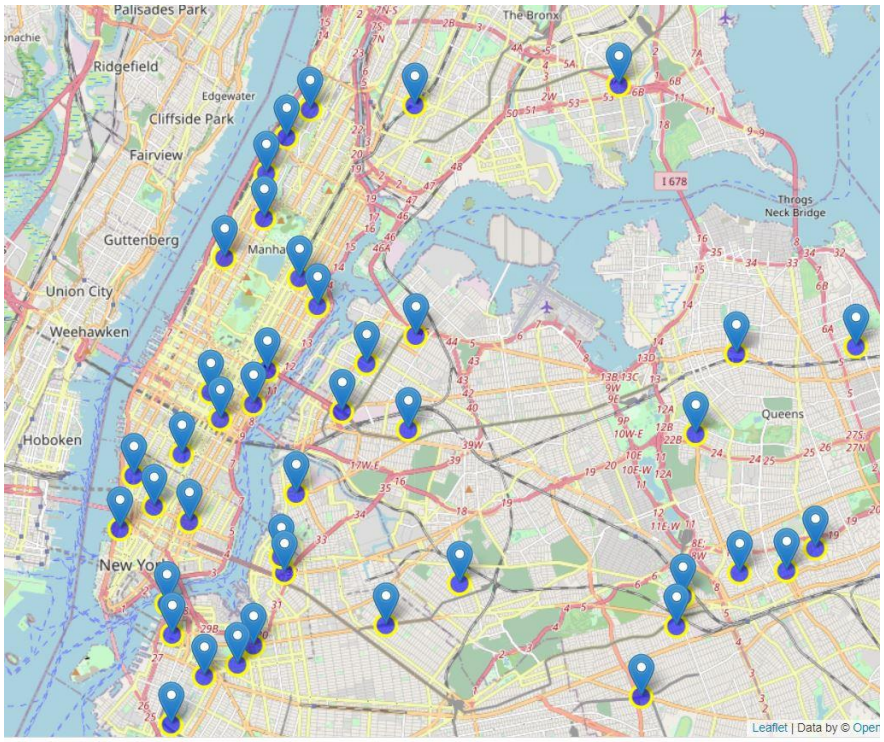


Results

By exploring all of the possible angles of this project, I have used folium to visualize the results on map which can better explain the scenario.

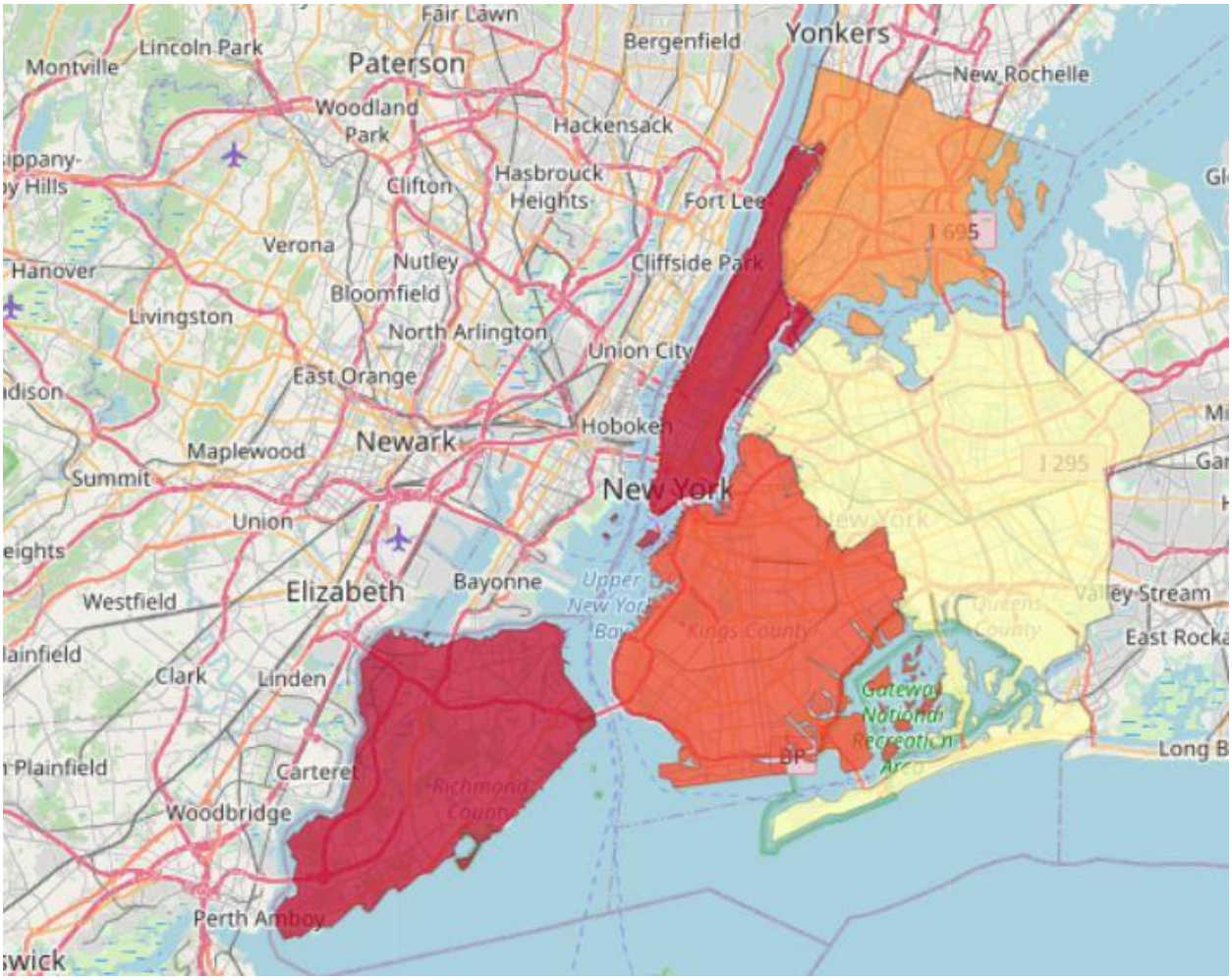
I have used folium to draw results on map.

- **Indian Restaurants in each part of New York is shown as below:**



- **Top Rating Indian Restaurants**

- **Average Rating according to each Borough in New York**



5. Conclusion

- Some of the Best Neighborhood for Indian Cuisine are (Bayside)Queens, Clinton Hill(Broklyn), Tribeca(Manhatton)
- Manhattan and Queens have highest number of Indian Restaurants.
- As Rating and Likes are considered: Tribeca, MidTown, Astoria, Noho, South Side, Sutton Place, Flatiron are some of the best places for Indian Restaurants to explore according to Likes and Ratings.
- According to Analysis Manhattan is the best place to stay as it is surrounded by Indian Restaurants.

6. Limitations

- The Data is dependent on FourSquare API. It may vary from other API's.
- The Ranking is based on rating and likes of Restaurants which does not tell about the updated data as of now, so its outdated and not completely reliable.

7. Future Work

More work can be done to refine venues and get closer look by considering more accurate and granular data.

We did not take into account the case where users reviewed the same restaurant multiple times. This means they might be rating each of their visit experiences rather than the restaurant itself. If this is the case, we can take the average rating of the same user about a restaurant rather than counting all of the user's reviews separately.

8. References

[1] T. Joachims, Optimizing Search Engines Using Clickthrough Data, Proceedings of the ACM Con-A Preference-Based Restaurant Recommendation System for Individuals and Groupsference on Knowledge Discovery and Data Mining (KDD), ACM, 2002

[2] D. Cosley, J. A. Konstan, J. Riedl, PolyLens: A Recommender System for Groups of Users, Proceedings of the 7th European Conference on Computer Supported Cooperative Work, 2001.

[3] H. Sajnani, V. Saini, K. Kumar, E. Gabrielova, P. Choudary, C.Lopes, Classifying Yelp reviews into relevant categories, http://www.ics.uci.edu/ ~vpsaini/.

[4] A. Ihler et al., Recommender Systems Designed for Yelp.com http://www.math.uci.edu/icamp/summer/research/student_research/recommender_systems_slides.pdf