**Capstone Project — The Battle of Neighbourhoods**

**INTRODUCTION**

New York City is the most populous city in United States and one of the important cities in the world.

It is the cultural, financial and media capital of the world. It influences many domains like fashion, education, politics, trade and commerce, entertainment, tourism, art and sports.

It is composed of five boroughs Brooklyn, Queens, Manhattan, The Bronx and Staten Island.

[Manhattan](https://en.wikipedia.org/wiki/Manhattan) is the smallest and most densely populated borough. NYC is a home to several ethnicities of which the Indians constitute one of the largest and fastest growing ethnicities. The Asian Indian population has been growing rapidly.

NYC is a home to several restaurants ranging from family owned to Michelin starred restaurants. Of these there are a good number of Indian restaurants too. You can definitely take out but having food in the restaurant has its own experience.

My Project is going to deal with scenarios related to the Indian restaurant industry in NYC. Say a person wants to start a new Indian restaurant. Obviously, there are several factors that a person needs to consider in addition to the budget. The most import is the location. Some of the factors to consider while choosing the location are

1. Best to open the new restaurant away from already highly successful restaurants.
2. Best to open restaurants near famous sites like tourist sites, transit hubs, parks etc since it can attract more people.
3. Identifying work districts to attract office goers for lunch/dinner

Also consider the tourists. Tourists are often lost in a big city. It would be helpful for the tourism companies to obtain some statistics information and provide to their customers.

So, this project will produce a report which will study and analyze some locations in NYC and provide valuable information regarding restaurants. This information can be used by the relevant stakeholders. Tourism companies, business owners planning to open restaurants are some of the few stakeholders.

**DATA DESCRIPTION**

We will be using the following data

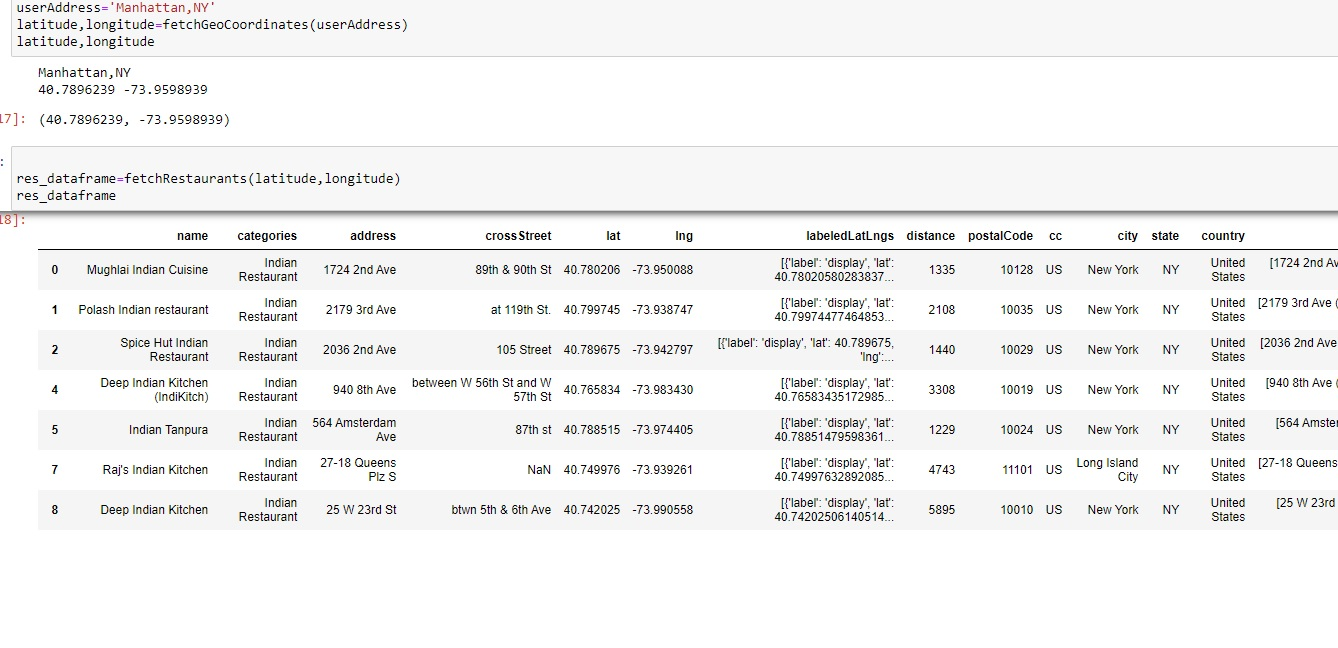
* + [Fousquare API](https://developer.foursquare.com/) : which provides us useful information about venues for analysis
  + <https://cocl.us/new_york_dataset>: provides information on New York city boroughs
  + https://en.wikipedia.org/wiki/List\_of\_Manhattan\_neighborhoods:Provides information about Manhattan Neighborhoods
  + Geopy library: geopy offers geocoding web services. We can locate the latitude and longitude of various locations using third-party geocoders and other data sources. We import the Nominatim library, chose Nominatim geocoder, and create a new instance of it.
  + Folium: Thisis used to visualize the geographic details of New York city and to show the locations of the various restaurants

**METHODOLOGY**

* Throughout the project, data has been extracted from the dataset into the Pandas data frames. Its then cleaned to hold only relevant columns and data.
* Folium has been used for data visualization.
* FourSquare API has been used to obtain distribution of Indian restaurants across boroughs and its details.

Below are some of the use cases that are handled

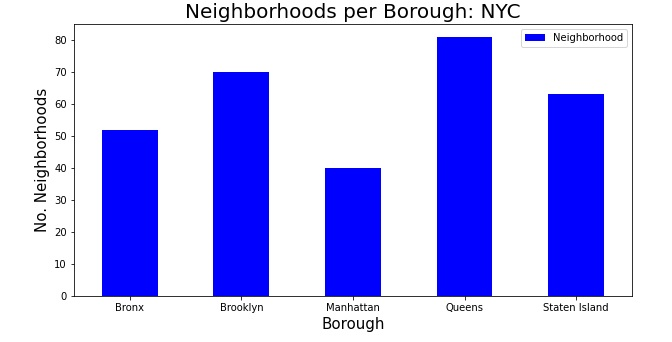
1. **Obtain restaurants around a given location**



Manhattan has been given as the address and all Indian restaurants in a radius of 5000 are listed.

2.**An analysis has been made of the boroughs and its neighborhoods of NYC obtained from the dataset**



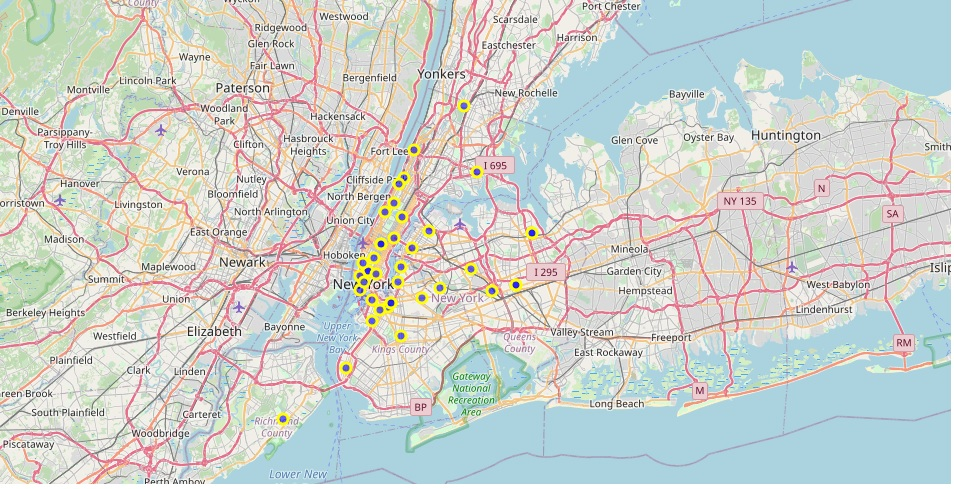


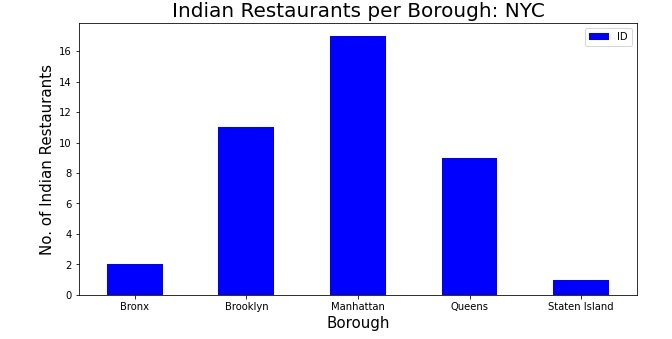
It can be observed that Manhattan has the least number of neighborhoods and Queens the most.

3.**Study all Indian restaurants across the boroughs**







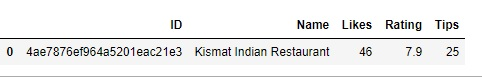


Manhattan has the most number and Staten Island the least.

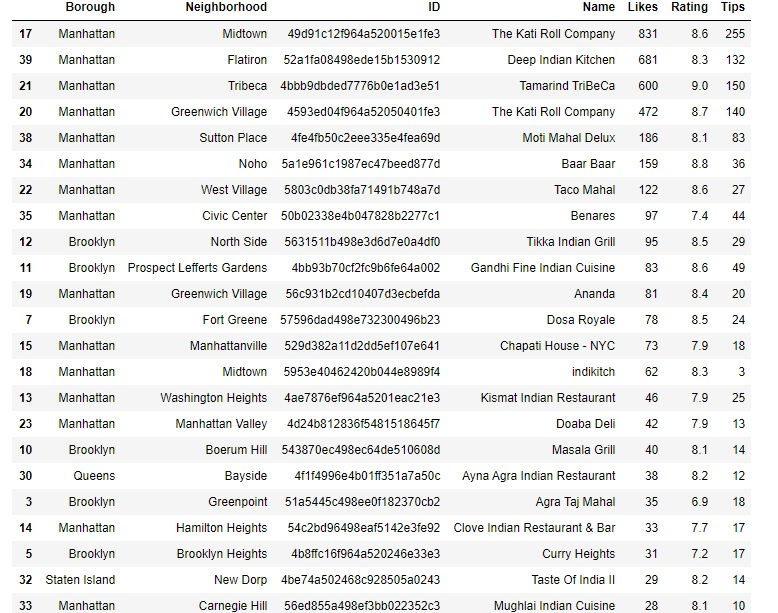
1. **One can also do analysis per borough. Manhattan has been chosen**

****

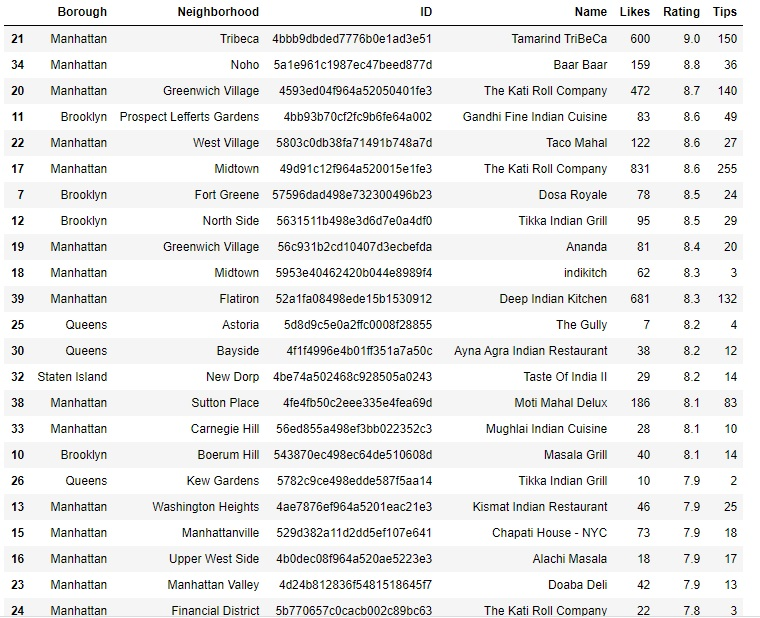
1. **Using Foursquare for exploring restaurant based on its venue ID.**

****

1. **Obtaining likes, ratings to do a comparison of restaurants**

****

Kati Roll company has the maximum likes

****

Tamarind Tribeca has the maximum Ratings

**RESULTS**

We got to study the distribution of Indian restaurants in NYC and were able to make some inferences which will be useful to both travelers and persons having business restaurants in NYC and were able to find out some interesting details which might be useful to travelers as well as people with business interests. Let’s summarize our findings:

* Manhattan is the borough in NYC that has least number of neighborhoods.
* Manhattan has the maximum number of Indian restaurants.
* Staten Island has the least number of Indian restaurants.
* The Katti roll Company in Manhattan has the most number of likes.
* Tamarind Tribeca in Manhattan has the highest rating.

Hence for a stakeholder like tourist, he can use this data to pick a good restaurant based on location, ratings or likes.

A person with business interest can choose a location for new restaurant by considering factors like restaurant numbers at a particular location. Eg Staten Island has the least number and hence a good choice, competitive restaurants etc

**DISCUSSION**

* In further analysis we can explore more data facts about restaurants like menu, prices, convenience to reach from important locations etc. Also, we are highly dependent on the accuracy of Foursquare data.
* Maybe census data also can be employed to obtain facts which can play another factor.

**CONCLUSION**

* In today’s world, data and its analysis can be used to find the solutions to many problems.
* Like seen in this project, we could project the distribution of the various Indian restaurants and its locations across the boroughs of New York city. These results can help tourism companies and persons with business interest. In the world before analytics, businesses would not thrive only because of gap in knowledge. Now the data can benefit businesses to make informed decisions.
* Python has some excellent libraries to explore and analyze data.
* Foursquare API and Folium are excellent for exploration and analysis of data.

GitHub Link to code:

https://github.com/Jaz-1999/Coursera\_Capstone/blob/main/Capstone%20Project\_%20The%20Battle%20of%20Neighborhoods.ipynb