# **Applied Data Science Capstone Project Report**

# ****Introduction****

New York is a major central city for diversity since many people from different cultural atmospheres have brought their families and dreams to New York City. The city has consistently seen people from around the world move to the city and call it home. It has been a center for trade and economic growth. New York is known world wide as a cultural melting pot. While other states have had immigration surges, none have compared to the diversity and sheer number of immigrants that have made their way to the City. So these number of varies cultures combined to create a great diversity for itself. Since People from all over the world tend to come up here, we can see some many of their cultural aspects Transport, Food, Clothing, and so on…

Restaurant is a place where people come to have food and drinks for a cost, People love to do many things and try something new or stick with their own routines, it depends on the individual and there are so many of them with different cultural and various aspects in New York city. There are so many cuisines, which is based on the style of cooking, the ingredients, dishes and techniques. For our problem lets stick with Indian cuisine.

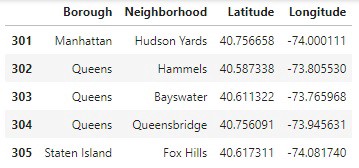
Let’s assume in this one of the world’s most diverse region we want to open an Indian restaurant, so what are all the factors we have to take into account such as follows,

* Market Places
* Competition in particular location
* Aiding places that make people come to restaurants like Gym, Entertaining Public places
* Population
* Menu from competitors

And so on… So our solution needs to be data driven for avoiding or considering low risk criteria and high success rate and thus apply our overall knowledge in the techniques and the tools gained so far in this course.

# Data

New York Neighborhood Data which will be used to get to know about the various neighborhoods which are gonna take into consideration for our Project.



New York Neighborhood data set

*Data source :*[*https://cocl.us/new\_york\_dataset*](https://cocl.us/new_york_dataset)

Geo-spacial data of the New York to get a better understanding of the neighborhoods in it and their corresponding locations in the Folium map would make certain things clear for the Project. This will be achieved using the acquired data and visualize the same using Choropleth maps.



Geo Spacial Data of New York City

*Data source :*[*https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm*](https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm)

Finally the data going to be collected/acquired from the Foursquare API about the various restaurants in each neighborhood of New York city especially Indian cuisine, which will be used for acquiring the information regarding same for all the venues of each neighborhood.

*Data source :*[*Foursquare*](https://enterprise.foursquare.com/products/places)

# ****Step by Step approach****

* Acquire the data from the various sources as mentioned below,
* New York neighborhood data from <https://cocl.us/new_york_dataset>
* Borough Boundaries data from <https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm>
* Indian Cuisine related data from [Foursquare](https://enterprise.foursquare.com/products/places)
* Filter the data for the required data acquired from Foursqaure which is only about Indian Cuisine.
* Acquire the tips and from that get the rating of every restaurant in the neighborhood of New York.
* Visualize the findings using Matplot lib and Folium for better understanding.

# ****Interests****

The main interest of the Business to be started as an Indian Cuisine restaurant would be to know about the many criteria that leads to come to a data driven conclusion for selecting the best neighborhood for opening one. But some of the criteria we are gonna be focusing on in our project is as follows,

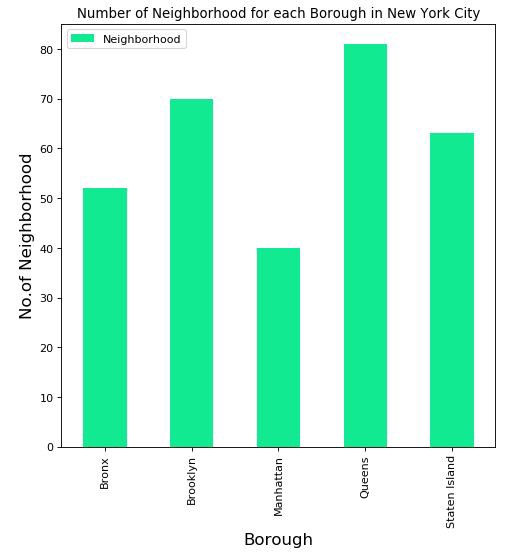
* Best Neighborhood for opening Restaurant (Indian Cuisine) in New York.
* Area which lack Indian Restaurants.
* Places with best Indian Cuisine.
* Areas with Potential Indian Restaurants.

# ****Exploratory Data Analysis****

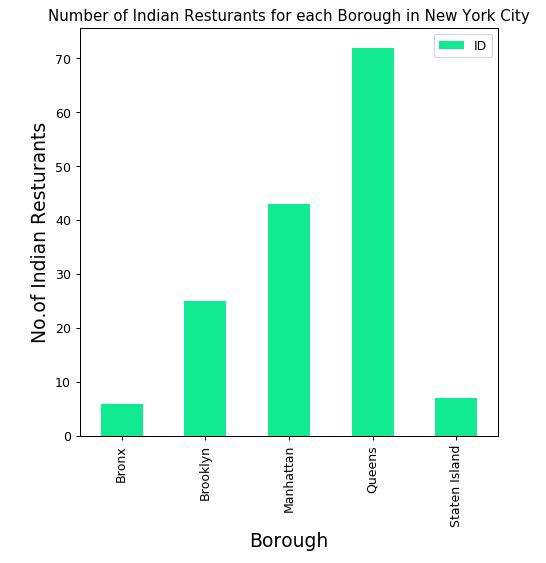
I have done analysis with the data after acquisition and cleaning and filtering the required information, For complete analysis details check out … [this](https://github.com/judeleander/Coursera_Capstone/blob/master/Battle%20of%20Neighborhoods/Battle-of-Neighborhoods.ipynb)

# ****Visualizations****

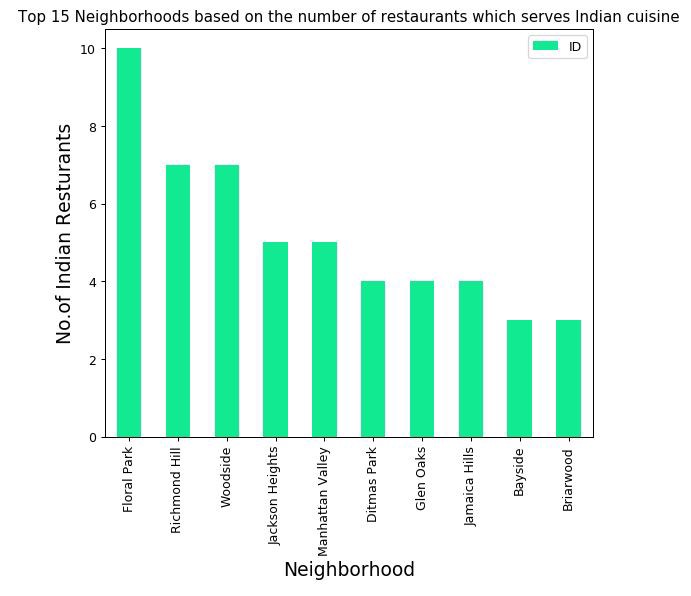
Boroughs in New York charted based on the number of Neighborhoods.



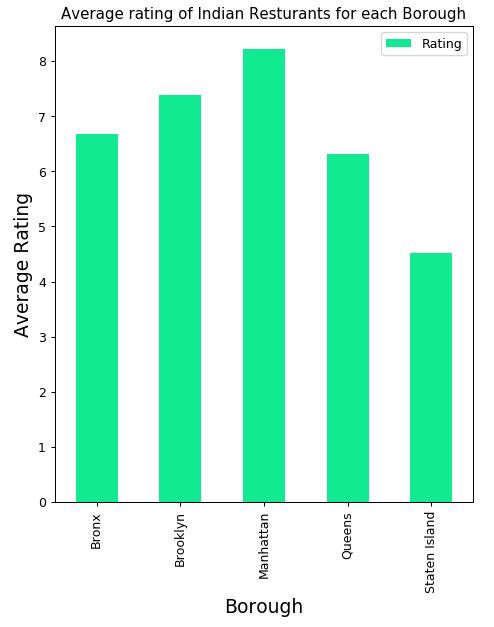
Boroughs in New York and the corresponding number of restaurants with Indian cuisines.



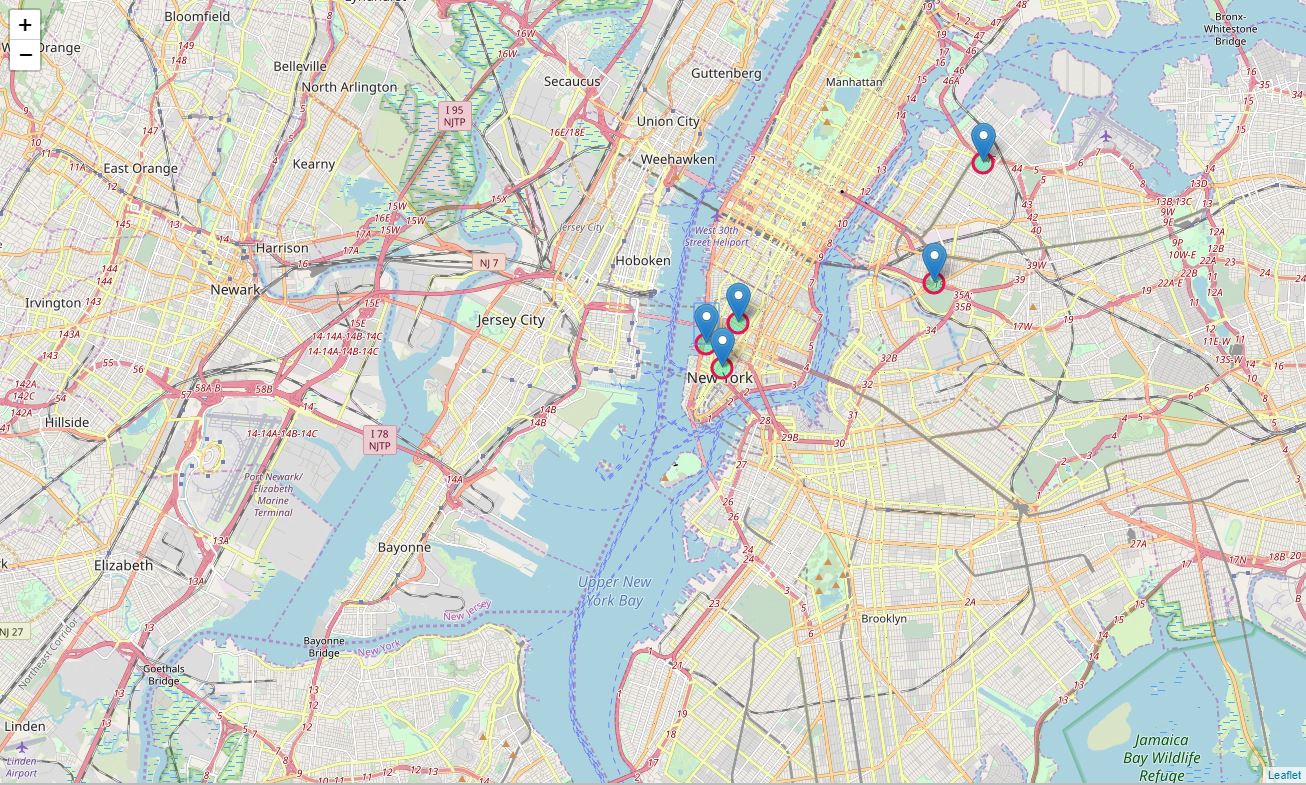
Top 15 Neighborhoods in New York city based on number of restaurants serving Indian cuisine.



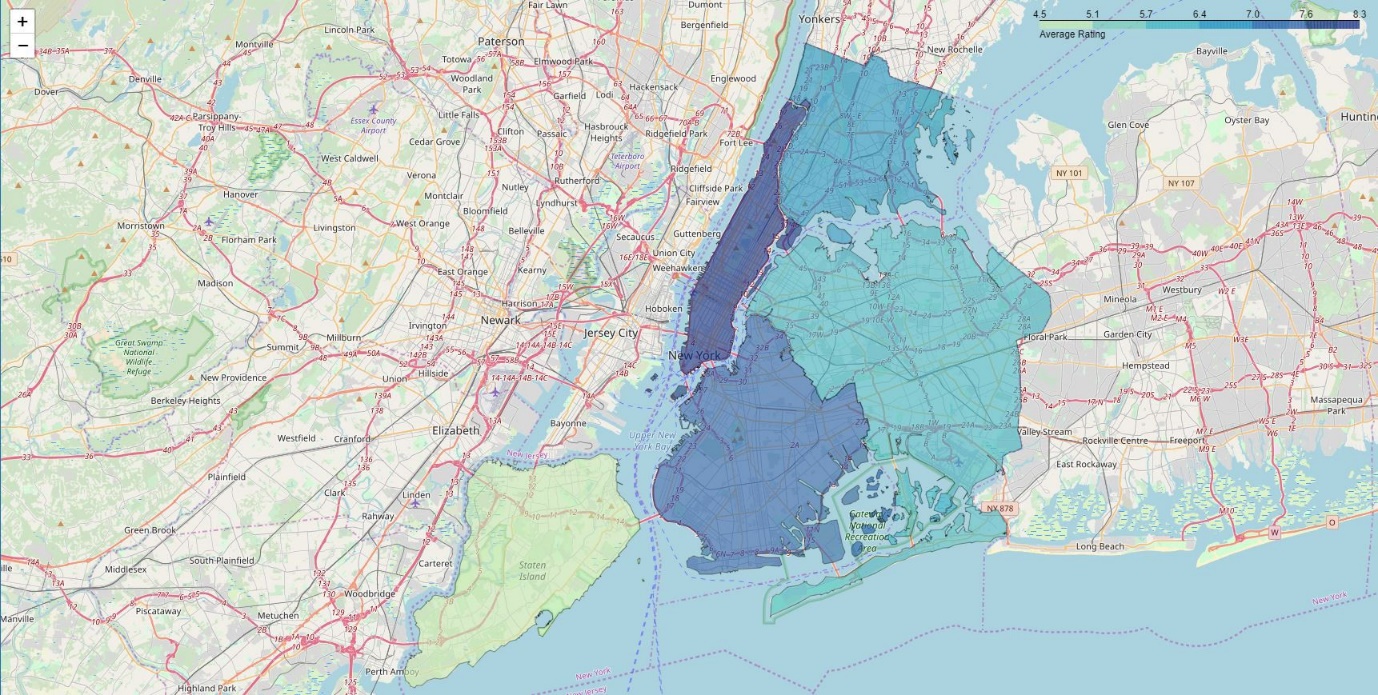
Boroughs vs average rating for the Indian cuisine restaurants in the same in New York.



Neighborhoods which has top rated Indian restaurants in New York city.



Choropleth map of Boroughs in New York city based on the average rating of the restaurants with Indian cuisine.



# ****Conclusion****

Based on the Data Analysis and Visualization using Folium and Matplotlib we have rendered various bar charts and two maps from which we can come to a better understanding of the interests we had earlier to be discovered which are specified as follows.

* Astoria(Queens), Blissville(Queens), Civic Center(Manhattan) are some of the best neighborhoods for indian cuisine.
* Staten Island ranks last in average rating of Indian Resturants.
* Manhattan is the best place to stay if you prefer Indian Cuisine.
* Manhattan have potential Indian Resturant Market