FINAL REPORT

<u>BATTLE OF NEIGHBORHOOD</u>

Introduction/Business Problem:

In this project, main focus is on New York City and as we all know that it is the most populous city in the United States. With an estimated 2019 population of 8,336,817 distributed over about 302.6 square miles (784 km2), New York is also the most densely populated major city in the United States. New York City is a global hub of business and commerce, as a center for banking and finance, retailing, world trade, transportation, tourism, real estate, new media, traditional media, advertising, legal services, accountancy, insurance, theater, fashion, thus it attracts many people from all over the world and thus it forms the major entry point to the immigrants and therefore New York City's food culture includes an array of international cuisines influenced by the city's immigrant history.

Project will concentrate on following problems:

1. **What is best place in New York City for Indian Cuisine?**
2. **Which is the best place to stay if you prefer Indian Cuisine?**
3. **Visualizing every parts of New York City that serves Indian foods**
4. **Which are the areas that lack Indian Restaurants?**

Target Audience :

This project report would be beneficial for every person who wishes to try best Indian Cuisine in New York or wanted to know about Indian Cuisine in New York or willing to research about them.

Data Section:

For this particular project the data was collected from various resources available on Google therefore keeping the BUSINESS PROBLEM in mind which is mentioned in INTRODUCTION part, below are the required data and information along with the sources are -:

1. New York City data that contains list of Boroughs, Neighborhoods along with their location i.e. latitude and longitude.

   Data set is available here : " https://cocl.us/new_york_dataset " which is in JSON format.

2. Geospatial data are defined in the ISO/TC 211 series of standards as data and information having an implicit or explicit association with a location relative to Earth.

   Informaton regarding geospatial data is available at :

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

3. Indian restaurants in each part/area of New York City.

   For this purpose Foursquare API was used which provides location based experiences with diverse information about venues, users, photos, and check-ins. The **API** supports real time access to places, Snap-to-Place that assigns users to specific locations, and Geo-tag.

   Foursquare Web site - "https://developer.foursquare.com/"

Methodology :

Step 1 : First step involves importing all the necessary libraries which are mentioned below:

BeautifulSoup : Beautiful Soup is a Python library for pulling data out of HTML and XML files. It works with your favorite parser to provide idiomatic ways of navigating, searching, and modifying the parse tree. It commonly saves programmers hours or days of work

Requests : Requests is a Python HTTP library, released under the Apache License 2.0. The goal of the project is to make HTTP requests simpler and more human-friendly. The current version is 2.23.0.Requests is one of the most popular python libraries that is not included with python, it has been proposed that requests be distributed with python by default.

Pandas : pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series. It is free software released under the three-clause BSD license.

Numpy : Numpy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays

Geocoder : Geocoder is a python library which is used for computational process of transforming a physical address description to a location on the Earth's surface (spatial representation in numerical coordinates)

Folium : Folium is a powerful Python library that helps you create several types of Leaflet maps.It enables both the binding of data to a map for choropleth visualizations as well as passing rich vector/raster/HTML visualizations as markers on the map**.**

Matplotlib Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy.

Step 2: Second step involves utilizing the Foursquare API by applying following credentials:

1. CLIENT ID
2. CLIENT SECRET
3. VERSION
4. LIMIT

Step 3: Third step involves collection of data from the link mentioned in Data Section : " https://cocl.us/new_york_dataset "

'Longitude': nei

```
        return new_york_data

In [8]: new_york_data=get_new_york_data()
        new_york_data.head()

Out[8]:
```

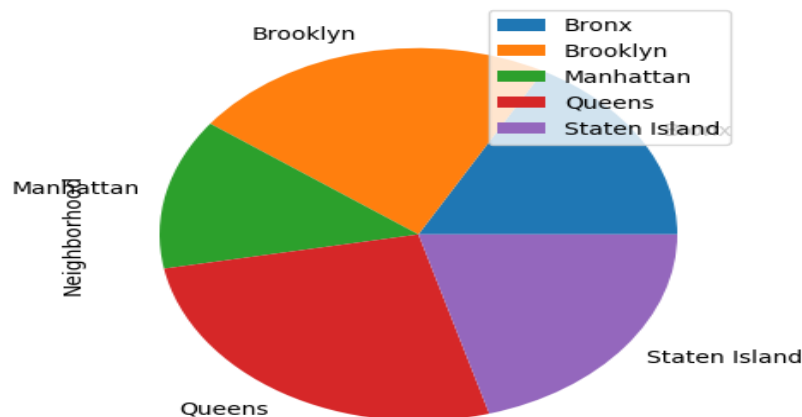|	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

Step 4: Visualizing the above dataset .

Visualising above 306 Neighborhood by creating pie chart

```
In [8]: plt.figure(figsize=(9,5), dpi = 100)
        #giving a bar plot
        new_york_data.groupby('Borough',axis=0)['Neighborhood'].count().plot(kind = 'pie')
        plt.legend(loc = 'upper right')
        #displays the plot
        plt.show()
```

Step 5 : Prepare neighborhood list that contains Indian restaurants

	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
5	Bronx	Unionport	55dfa36a498e164ef19bef7b	Premium Sweets & Restaurant
6	Brooklyn	Bay Ridge	4b5a4dc8f964a520a2bb28e3	Taj Mahal
7	Brooklyn	Bay Ridge	545835a1498e820edc6f636f	Bombay Grill
8	Brooklyn	Brighton Beach	4af0d31bf964a5207ddf21e3	Pak Nasheman
9	Brooklyn	Flatbush	52213c4211d295d4c57a607c	Ashoka Grill
10	Brooklyn	Flatbush	564d283d498e6e851df79d87	Great Indian Curry
11	Brooklyn	East Flatbush	512a9ea9e4b004fb8eeb84e5	Silver Krust
12	Brooklyn	Kensington	4db0f4371a729faa56497f20	Mashallah

Out[29]:

Thus there are total 146 Indian Resturants across New York City

Step 6: Involves plotting a bar chart which shows that **Queens** have highest no. of Indian restaurants.

Number of Indian Resturants for each Borough in New York City

From above bar plot it can be seen that Queens has highest number of Indian restaurents

Step 7 : Now with help of **Foursquare API** related factors such as "Tips", "Rating", "Likes" were found and recorded.

In [39]: indian_rest_stats_ny.head()

Out[39]:

	Borough	Neighborhood	ID	Name	Likes	Rating	Tips
0	Bronx	Woodlawn	4c0448d9310fc9b6bf1dc761	Curry Spot	5	7.8	10
1	Bronx	Parkchester	4c194631838020a13e78e561	Melanies Roti Bar And Grill	3	6.0	2
2	Bronx	Spuyten Duyvil	4c04544df423a593ac83d116	Cumin Indian Cuisine	13	6.0	9
3	Bronx	Concourse	551b7f75498e86c00a0ed2e1	Hungry Bird	8	6.8	3
4	Bronx	Unionport	4c194631838020a13e78e561	Melanies Roti Bar And Grill	3	6.0	2

In [40]: indian rest stats ny.shape

Among those the top the top neighborhoods with top average rating of Indian restaurants are :-

In [52]: ny_neighborhood_stats.sort_values(['Average Rating'],ascending=False).head(10)

Out[52]:

	Neighborhood	Average Rating
0	Astoria	9.0
5	Blissville	9.0
26	Fort Greene	8.8
13	Clinton Hill	8.8
53	Prospect Heights	8.7
46	Murray Hill	8.6
71	Tudor City	8.6
70	Tribeca	8.6
12	Civic Center	8.6
33	Holliswood	8.5

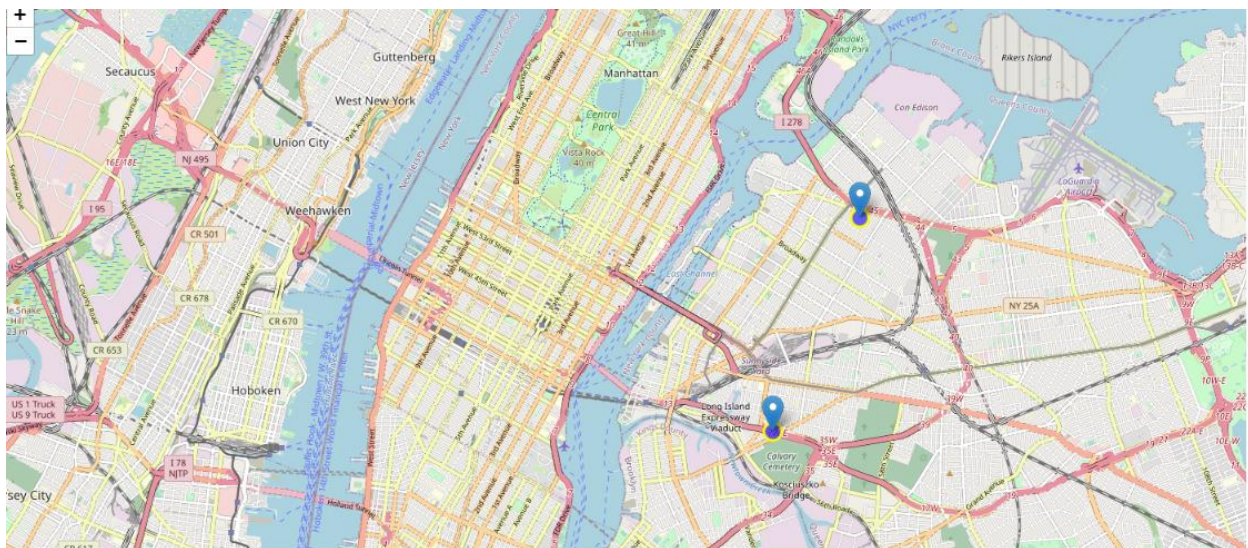**Step 8:** consider all the neighborhoods with average rating greater or equal 9.0 to visualize on map.

NOTE:- The number here can vary depending on the data set, and also we can select average rating >=8.0 also, but I have choosed the rating >=9.0

In [56]: ny_neighborhood_stats

Out[56]:

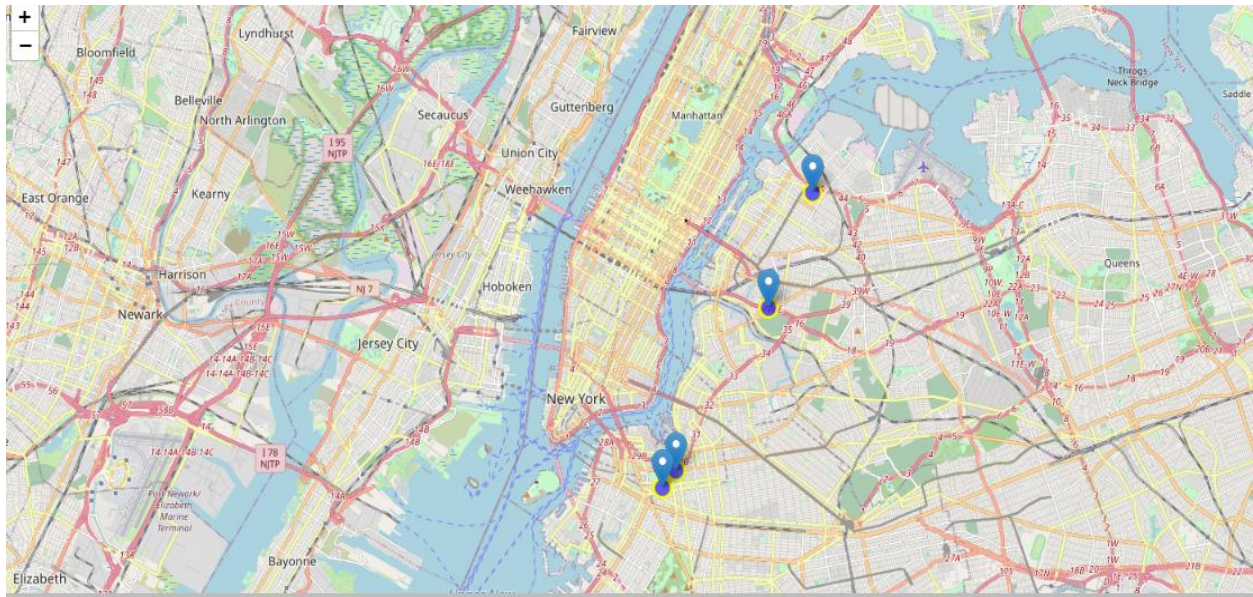|	Neighborhood	Average Rating
0	Astoria	9.0
5	Blissville	9.0

Step 9 : Visualizing Neighborhood and Borough depending on Average Rating using Folium Library

For average rating >=8.8, we have 4 restaurants in total, which are as follows:-

	Borough	Neighborhood	Latitude	Longitude	Average Rating
0	Queens	Astoria	40.768509	-73.915654	9.0
1	Queens	Blissville	40.737251	-73.932442	9.0
2	Brooklyn	Clinton Hill	40.693229	-73.967843	8.8
3	Brooklyn	Fort Greene	40.688527	-73.972906	8.8

Its corresponding map will be:-

Conclusion/Result:

From the anaysis here is the conclusion:

- **Manhattan is the best place to stay if you prefer Indian Cuisine**

- **2) Manhattan have potential Indian Resturant Market**

- **3) Staten Island ranks last in average rating of Indian Resturants.**

- **4)Floral Park in Queens has the highest number of Indian Resturants.**

- **Astoria(Queens), Blissville(Queens) are the only 2 best neighborhood for Indian Cuisine (for avg rating >=9.0)**