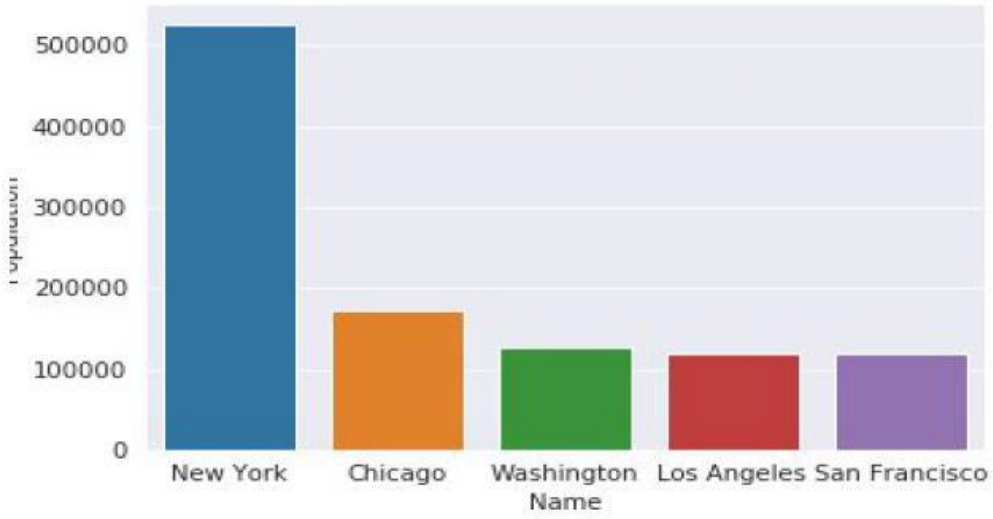
Capstone Project - The Battle of Neighborhoods

A. INTRODUCTION:

a. Description & Discussion of the Background

An entrepreneur wants to start a new Multi Cuisine Indian restaurant in **NEW YORK CITY.**Since he believes that New York is one of the most densely populated Indian immigrant states in US. So He approaches a Data Scientist for a suggestion where to start his business.



(source from Wikipedia)

Since being a Data Scientist we need to solve his problem and recommend the best neighbourhood of New York to start his business.

b. Problem :

1. Finding all the Neighborhoods of New York.
2. Finding all the Neighborhoods with INDIAN Restaurants.
3. Finding all the Neighborhood with-out single INDIAN Restaurants.
4. Finding the best place and most visited INDIAN Restaurant.
5. Suggest best Neighborhood for starting business

B. Data acquisition and cleaning:

New York City is one of the largest metropolises in the USA, it has various Neighborhoods and various kinds of restaurants within its neighborhood, but in our scenario we only need INDIAN restaurants, so we take help of FourSquare API to accomplish our task.

For this project we need following data :

❑ First we need all the neighborhoods of New York along with their latitude and longitude coordinates.

➔**Data Source** : https://cocl.us/new_york_dataset

➔**Description** : cocl.us has a list of Neighborhood of New York, we scrap it using python libraries and get its location coordinates using geopy.geocoder library in python.

❑ Indian restaurants in each Neighborhood of New York

➔**Data Source** : **FourSquare API**

➔**Description** : By using this API we will get all the venues in each neighborhood. We can filter this venues to grab only INDIAN restaurants

C. Methodology :

As every data science project has different phases involved in its life cycle.

This project also has different phases like

➔Business problem

➔Approach followed

➔Requirements Gathering (Data collection)

➔Modeling and evaluation

C.1: Business Problem

This is already stated in the Introduction part (refer it)

C.2 : Approach Followed

We consider some metrics to solve this problem

1. Number of INDIAN restaurants in neighborhood
2. Distance of neighborhood from the center of the city

3. Is there any tourist places to visit

The optimal place to start business is a location where there are **less** INDIAN

restaurants, **Minimum distance** from the center of city and there are some

popular **tourist places** to visit.

C.3 : Requirement gathering (Data collection):

As mentioned in the Data acquisition section, we gather data Foursquare API.

➔ Data from cocl.us that has all Neighborhoods of New York City.

	Borough	Neighborhood
0	Bronx	Wakefield
1	Bronx	Co-op City
2	Bronx	Eastchester
3	Bronx	Fieldston
4	Bronx	Riverdale

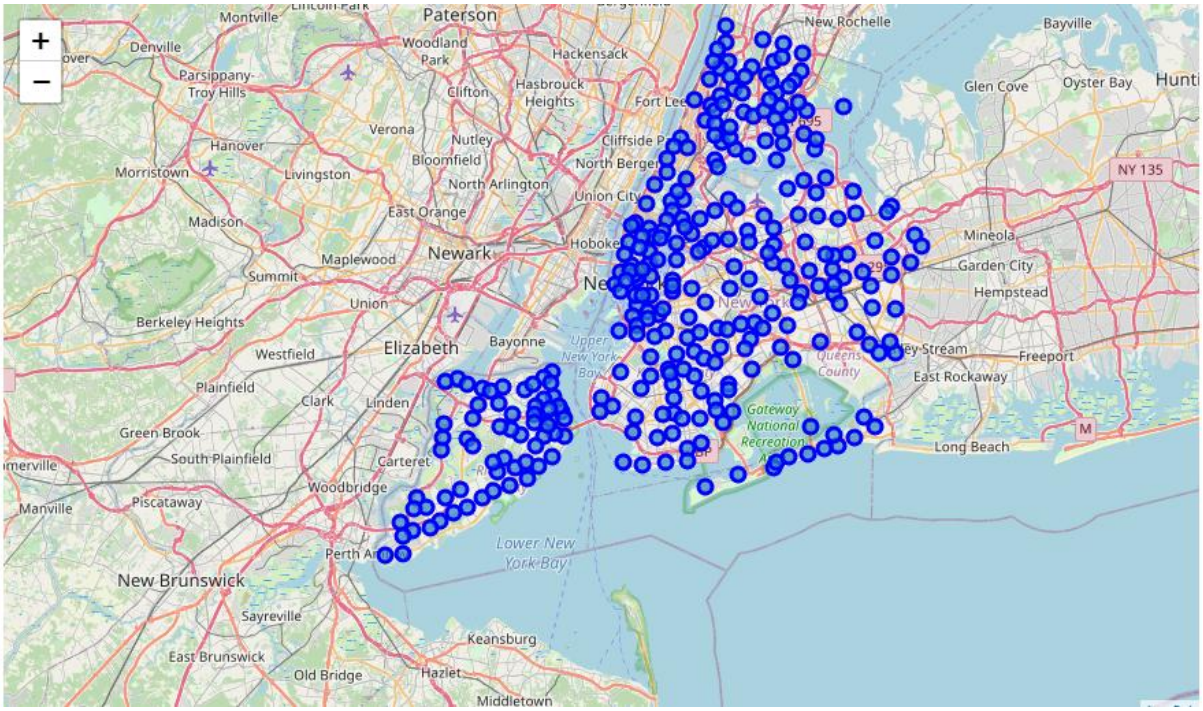
➔ Assigning each Neighborhood its corresponding latitude using Geopy Library in python

	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

We will use geopy library for getting coordinates of Queens,NY for further use:

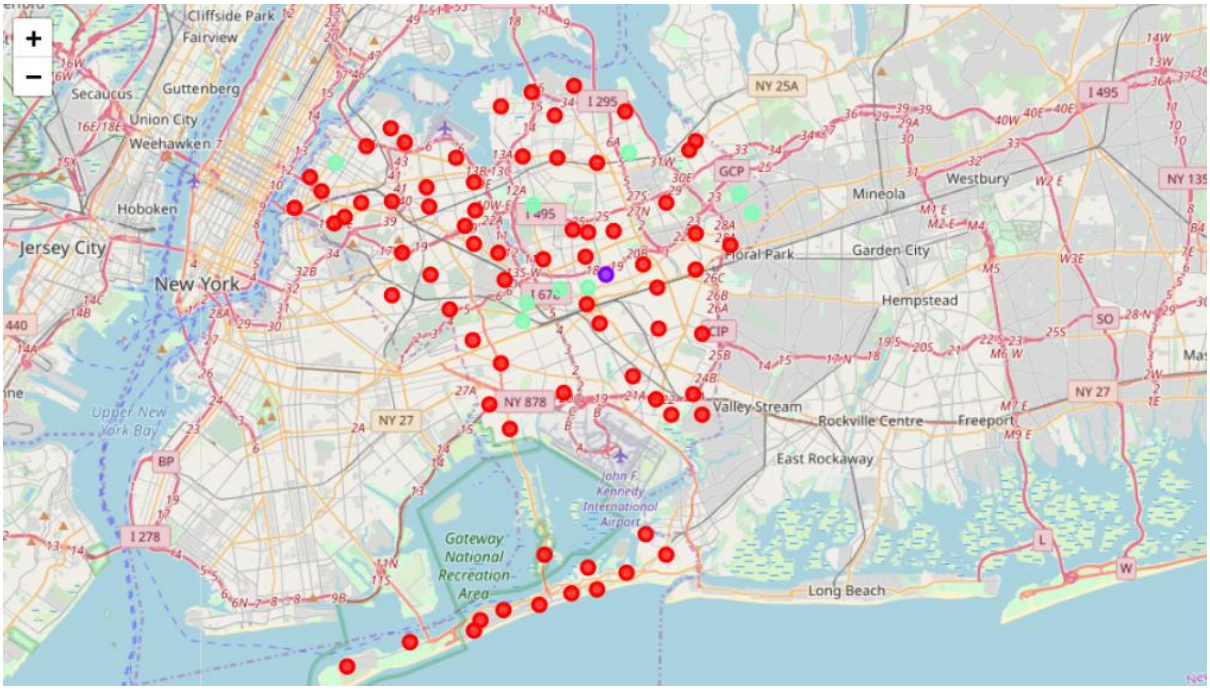
```
address = 'New York City, NY'

geolocator = Nominatim(user_agent="ny_explorer")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of New York City are {}, {}.'.format(latitude, longitude))
```

```
The geograpical coordinate of New York City are 40.7127281, -74.0060152.
```



.

Visualization of the INDIAN Restaurants in NEW YORK



The red marker show the Location where there are larger INDIAN Restaurants

D. Conclusion :

1. Best place in NEW YORK to find INDIAN restaurant : Prospect Leffert Gardens.

2. Best restaurant in Prospect Leffert Gardens to have dinner : Gandhi Fine Indian Cuisine Restaurent (it has got 8.6 ratings and 82 likes)