Capstone Project — The Battle of Neighbourhoods - Week 1

Introduction - Business Problem

India is one of the most tourist visited countries in the world for its monuments, culture, traditions and the list could go on and on. There is lot of tourist market in india and New delhi is one of the most visited places for lot of tourists to catch the glimpse of Taj mahal, Red fort etc.,

</p> New Delhi is the capital city of India. It is a part of the city of Delhi's 11 districts. The city itself has a population of 257,803. However, the much larger metro area has a population that exceeds 26 million.</p>

New Delhi are used interchangeably to refer to the National Capital Territory of Delhi (NCT), these are two distinct entities, with New Delhi forming a small part of Delhi. The National Capital Region is a much larger entity comprising the entire NCT along with adjoining districts in neighboring states.

The official language of New Delhi and the one that is most widely spoken is Hindi. However, English is also spoken as a formal language within businesses and government agencies. Over last decades it is continuously grow because of the city's important role in government and commercial business.
With it's diverse culture , comes diverse food items. There are many restaurants in New Delhi City, each belonging to different categories like Chinese , Italian , French etc.

The idea of my project is to present the client and stake holders with answers to the follwing questions

- Which areas have the best restaurants in New Delhi?
- What is best location in New Delhi City for the famous Biryani ?
- What is best location in New Delhi City for Chinese Cuisine ?
- Which areas of all areas have large number of Chinese Resturant Market ?
- Which area of all areas have less number of resturants ?
- Which area of all areas have most number of resturants ?
- Which is the best place to stay if I prefer Chinese Cuisine ?

Data

For this project we need the following data :

- New Delhi Resturants data that contains list Locality, Resturant name,Rating along with their latitude and longitude.
  - Data source : Zomato kaggel dataset (https://www.kaggle.com/shrutimehta/zomato-restaurants-data)
  - Description : This data set contains the required information. And we will use this data set to explore various locality of new delhi city.

- Nearby places in each locality of new delhi city.
  - Data source : Fousquare API (https://developer.foursquare.com/)
  - Description : By using this api we will get all the venues in each neighborhood.

Approach

- Collect the new delhi city data from Zomato kaggel dataset (https://www.kaggle.com/shrutimehta/zomato-restaurants-data)
- Using FourSquare API we will find all venues for each neighborhood.
- Filter out all venues that are nearby by locality.
- Using aggregative rating for each resturant to find the best places.
- Visualize the Ranking of neighborhoods using folium library(python)

In [1]:

```
import pandas as pd
import numpy as np
import requests # library to handle requests
from pandas.io.json import json_normalize # tranform JSON file into a pandas dataframe
# Matplotlib and associated plotting modules
import matplotlib.cm as cm
import matplotlib.colors as colors
# import k-means from clustering stage
from sklearn.cluster import KMeans

!conda install -c conda-forge folium=0.5.0 --yes # uncomment this line if you haven't c
ompleted the Foursquare API lab
import folium # map rendering library
! pip install geocoder
import geocoder
```

```
usage: conda-script.py [-h] [-V] command ...
conda-script.py: error: unrecognized arguments: # uncomment this line if y
ou haven't completed the Foursquare API lab

Requirement already satisfied: geocoder in c:\users\srira\anaconda3\lib\si
te-packages (1.38.1)
Requirement already satisfied: click in c:\users\srira\anaconda3\lib\site-
packages (from geocoder) (7.1.2)
Requirement already satisfied: future in c:\users\srira\anaconda3\lib\site
-packages (from geocoder) (0.18.2)
Requirement already satisfied: requests in c:\users\srira\anaconda3\lib\si
te-packages (from geocoder) (2.24.0)
Requirement already satisfied: six in c:\users\srira\anaconda3\lib\site-pa
ckages (from geocoder) (1.15.0)
Requirement already satisfied: ratelim in c:\users\srira\anaconda3\lib\sit
e-packages (from geocoder) (0.1.6)
Requirement already satisfied: idna<3,>=2.5 in c:\users\srira\anaconda3\li
b\site-packages (from requests->geocoder) (2.10)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\srira\anacon
da3\lib\site-packages (from requests->geocoder) (2020.6.20)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in
c:\users\srira\anaconda3\lib\site-packages (from requests->geocoder) (1.2
5.10)
Requirement already satisfied: chardet<4,>=3.0.2 in c:\users\srira\anacond
a3\lib\site-packages (from requests->geocoder) (3.0.4)
Requirement already satisfied: decorator in c:\users\srira\anaconda3\lib\s
ite-packages (from ratelim->geocoder) (4.4.2)
```

Read the zomato resturant data from csv file

In [2]:

```
df = pd.read_csv('C:/Users/srira/Downloads/zomato.csv',encoding='ISO-8859-1')
df_india = df[df['Country Code'] == 1]
## New Delhi
df_NDLS = df_india[df_india['City'] == 'New Delhi']

df_NDLS.reset_index(drop=True, inplace=True)
df_NDLS.head()
```

Out[2]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitud
0	18287358	Food Cloud	1	New Delhi	Aaya Nagar, New Delhi	Aaya Nagar	Aaya Nagar, New Delhi	0.000000	0.00000
1	18216944	Burger.in	1	New Delhi	84, Near Honda Showroom, Adchini, New Delhi	Adchini	Adchini, New Delhi	77.196923	28.53538
2	313333	Days of the Raj	1	New Delhi	81/3, 1st Floor, Qutub Residency, Adchini, New...	Adchini	Adchini, New Delhi	77.197475	28.53549
3	18384127	Dilli Ka Dhaba	1	New Delhi	66 A, Ground Floor, Sri Aurobindo Marg, Adchin...	Adchini	Adchini, New Delhi	77.198033	28.53754
4	582	Govardhan	1	New Delhi	84, Adjacent Hero Motor Bike Showroom, Main Me...	Adchini	Adchini, New Delhi	77.196924	28.53552

5 rows × 21 columns

Data Cleaning

remove the unwanted columns and rows from dataset

In [3]:

```
df_Res= df_NDLS[df_NDLS.Longitude !=0.000000][['Restaurant Name','Locality','Longitude'
,'Latitude','Cuisines','Aggregate rating','Rating text','Votes']]
```