The Battle of Neighborhoods – Data Sources (Week-1)

Data Sources

The population of New York city is distributed into 5 boroughs and 306 neighborhoods. To explore the data about Thai restaurants, we need to get the access to the data containing the boroughs and their geospatial coordinates.

The data has been freely downloaded from the below website:
https://geo.nyu.edu/catalog/nyu_2451_34572

I will convert the data into Pandas data frame for easy data analysis and visualization. Then I will use Foursquare API to get the required information about venues with filtering applied to find the Thai restaurants. Graphical representation of number of boroughs vs number of Thai restaurants and neighborhoods plotted against number of Thai restaurants would give us a good and clear idea about the distribution of the restaurants in New York city.

Foursquare data will also be used to retrieve the information about rating and number of average tips that will help us to fulfil the requirement as stated in the problem description for the location of both expensive and cheap restaurants. The recommendation for the location of both restaurants will be plotted and viewed on the folium based maps using Jupyter notebook. Various data analysis and plotting libraries and algorithms will be used during the course of the capstone project.

Methodology

Various data analysis and plotting libraries and algorithms will be used during the course of the capstone project. Pandas for data cleaning, analysis and statistical plots. Possible use of seaborn and plotly where deemed necessary for high quality graphics. Numpy will be used to handle the data in vector format.

Folium and geopy will help us plot the geospatial data on the maps with street layer in the background and the location of neighbourhood and restaurants as a foreground layer. The top locations for the both the restaurants will be selected based on the client criteria.

Highest tips and ratings for upbeat restaurant. Neighbourhood with low number of restaurants for affordable and express version of the restaurant.