IBM DATA SCIENCE

Applied Data Science Capstone

The Battle of Neighbourhoods (Part-1)

Finding Best Venues, Finding Best Neighbourhoods, Segmenting and Clustering Neighbourhoods in Mumbai City

Data Section

Mumbai is the financial, commercial and entertainment capital of India. It is also one of the world's top ten centres of commerce in terms of global financial flow, generating 6.16% of India's GDP. Therefore, there are venues of various categories and given below are the sources which will be used to collect data.

Data needed for this project is as follows:

- Mumbai City data that contains list Boroughs, Neighbourhoods along with their Latitude and Longitude.
  - Data source : https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai
  - Description : This website contains the required information. After cleaning this data, it will be used to explore various neighbourhoods of Mumbai City.

- Venues of certain category given by user, details of venues of certain category and top venues in each neighbourhood of Mumbai City.
  - Data source : Foursquare API
  - Description : By using this API required information can be fetched and after converting it to data frames it can be cleaned and used to get required data.

- Locating the coordinates of addresses and cities.
  - Data source : geopy-geocoders (Third party geocoders and other data sources)
  - Description : This will be used to fetch co-ordinates of locations.