**Introduction/Business Problem**

The project aims to find similarity between the neighborhood of Canada with respect to top most common places nearby to each neighborhood.

Project considers Canada Neighborhood , lists out all the nearby places within 500 meter or 1 KM for each neighborhood of Canada. Categorize the venues as Park, Restaurant, store etc and find out top 10 most common places for each neighborhood.

Performs Kmeans clustering and group the places in right category based on the nearby most common venue. It provides the visualization using Folium Map and shows the neighborhood in respective cluster.

**Data**

Data used from the below link which contains postal code, borough, neighborhood

<https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>

Filtered out the postal codes where borough is not assigned.

Used Geospatial\_Coordinates.csv to retrieve longitude and latitude for each postal code which can be used to fetch Foursquare data

For each postal code, API call is made to Foursquare to retrieve the nearby places within 500 meter or 1 km and the max limit has been kept as 100 to return nearby places within 100.

url='https://api.foursquare.com/v2/venues/explore?&client\_id={}&client\_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(

CLIENT\_ID,

CLIENT\_SECRET,

VERSION,

dfnm\_latitude,

dfnm\_longitude,

radius,

LIMIT

)

Below data has been fetched through the API call.

name

address

latitude

longitude

distance

postalCode

cc

city

state

country

formattedAddress

categories

photos.count

photos.groups

venuePage.id

crossStreet

Below attributes retrieved from Foursquare to rank the top most popular venue for each longtitute and latitude and used for clustering.

Neighborhood

Neighborhood Latitude

Neighborhood Longitude

Venue

Venue Latitude

Venue Longitude

Venue Category

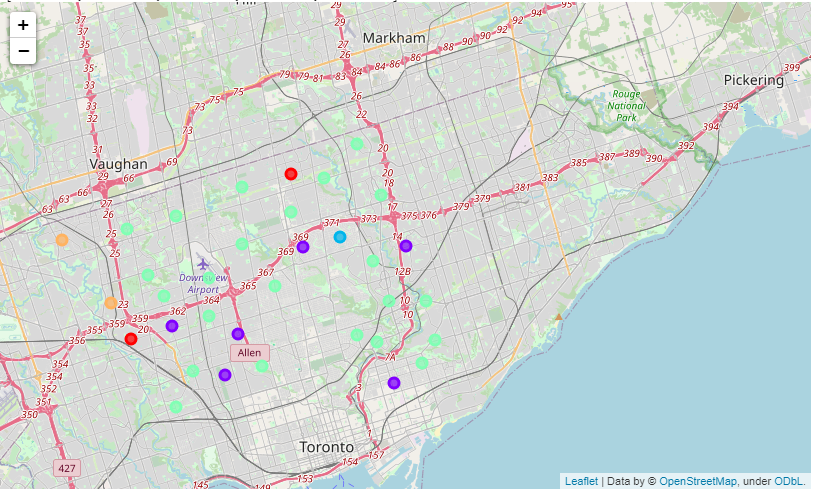
**Methodology**

K Means clustering technique has been used to perform similarity analysis in the neighborhood based on top most common places in the neighborhood.

No of Clusters defined as 5, it will converge after 3 iterations.

**Results**

Folium Map has been used for visualization.

****