BEST PLACE TO OPEN RESTURANT RECOMMENDER

1. Introduction:

The purpose of this Project is to help people in to find the best place to open new restaurant by exploring facilities and population around their neighborhood. It will help entrepreneur to making smart and efficient decision on selecting neighborhood out of numbers of other neighborhoods in Etobicoke, Canada.

Lots of people are staying to various Borough of Canada. We have taken Etobicoke that contain various neighborhood and done research for exploring the neighborhood to find a best place to open a new restaurant by taking the population and frequency of the other restaurant and stores . So this project is for those people who are looking for better neighborhoods to start a restaurant.

This Project aim to create an analysis of Etobicoke to search a best neighborhood as a comparative analysis between neighborhoods. The features include Population and frequency of other various categories of stores in the locality.

It will help people to get awareness of the area and neighborhood before to open to a restaurant in this Canadian city.

2. Data Section

Data Link: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

We used use Etobicoke dataset from city of Toronto datasets which we scrapped from on Week 3 Assignment and the population dataset based on postal code is obtain from canadas official website (https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/pd-pl/Tables/File.cfm?T=1201&SR=1&RPP=9999&PR=0&CMA=0&CSD=0&S=22&O=A&Lang=Eng&OFT=CSV). Dataset consisting postal code and population .

Foursquare API Data:

We will need data about different venues in different neighborhoods of that specific borough. In order to gain that information we will use "Foursquare" locational information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API.

After finding the list of neighborhoods, we then connect to the Foursquare API to gather information about venues inside each and every neighborhood. For each neighborhood, we have chosen the radius to be 1000meter.

The data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

1. Neighborhood
2. Neighborhood Latitude
3. Neighborhood Longitude
4. Venue
5. Name of the venue e.g. the name of a store or restaurant and type
6. Venue Latitude
7. Venue Longitude
8. Venue Category

Canada Geography Dataset from week 3

	Postal Code	Borough	Neighbourhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park, Harbourfront
3	M6A	North York	Lawrence Manor, Lawrence Heights
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government

Neighborhood Latitude and Longitude Dataset Week 3

Out[9]:

	Postal Code	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

Population dataset

Out[10]:

	Geographic code	Geographic name	Province or territory	Incompletely enumerated Indian reserves and Indian settlements, 2016	Population, 2016	Total private dwellings, 2016	Private dwellings occupied by usual residents, 2016
0	01	Canada	NaN	T	35151728.0	15412443.0	14072079.0
1	A0A	A0A	Newfoundland and Labrador	NaN	46587.0	26155.0	19426.0
2	A0B	A0B	Newfoundland and Labrador	NaN	19792.0	13658.0	8792.0
3	A0C	A0C	Newfoundland and Labrador	NaN	12587.0	8010.0	5606.0
4	A0E	A0E	Newfoundland and Labrador	NaN	22294.0	12293.0	9603.0

Map of Etobicoke

```
The geograpical coordinate of Etobicoke are 43.6435559, -79.5656326.
```

3. Methodology Section

Clustering Approach**:**Clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group than those in other groups. This is a form of unsupervised machine learning called k-means clustering algorithm.

Here we are the population of the neighborhood the various categories of restaurant and other stores as the feature and trying to find the cluster .Here there the Total 41 features passed

Features passed

Out[41]:

	Population	American Restaurant	Bakery	Bar	Baseball Field	Beer Store	Burger Joint	Burrito Place	Bus Line	Café	Chinese Restaurant	Coffee Shop	Construction & Landscaping	Convenience Store	Dessert Shop	Discount Store	Drugstore	Electr
0	0.056592	0.000000	0.000000	0.00	0.0	0.000000	0.0000	0.0000	0.000000	0.000000	0.000000	0.142857	0.0	0.000000	0.000000	0.0000	0.00	0.00
1	0.104816	0.000000	0.000000	0.00	0.0	0.111111	0.0000	0.0000	0.000000	0.111111	0.000000	0.111111	0.0	0.111111	0.000000	0.0000	0.00	0.1
2	0.097433	0.000000	0.000000	0.00	0.0	0.000000	0.0000	0.0000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0000	0.00	0.00
3	0.092366	0.000000	0.000000	0.00	0.0	0.166667	0.0000	0.0000	0.166667	0.000000	0.166667	0.000000	0.0	0.000000	0.000000	0.0000	0.00	0.00
4	0.046639	0.000000	0.062500	0.00	0.0	0.000000	0.0625	0.0625	0.000000	0.000000	0.000000	0.000000	0.0	0.062500	0.000000	0.0625	0.00	0.00
5	0.103951	0.058824	0.058824	0.00	0.0	0.000000	0.0000	0.0000	0.000000	0.117647	0.000000	0.117647	0.0	0.000000	0.058824	0.0000	0.00	0.00
6	0.111366	0.000000	0.000000	0.25	0.0	0.000000	0.0000	0.0000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0000	0.25	0.00
7	0.058303	0.000000	0.000000	0.00	0.5	0.000000	0.0000	0.0000	0.000000	0.000000	0.000000	0.000000	0.5	0.000000	0.000000	0.0000	0.00	0.00
8	0.153179	0.000000	0.000000	0.00	0.0	0.100000	0.0000	0.0000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0000	0.00	0.00
9	0.029528	0.000000	0.500000	0.00	0.0	0.000000	0.0000	0.0000	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0000	0.00	0.00
10	0.088690	0.000000	0.000000	0.00	0.0	0.000000	0.0000	0.0000	0.000000	0.000000	0.200000	0.000000	0.0	0.000000	0.000000	0.0000	0.00	0.00
11	0.057139	0.000000	0.000000	0.00	0.0	0.000000	0.0000	0.0000	0.000000	0.000000	0.125000	0.125000	0.0	0.000000	0.000000	0.1250	0.00	0.00

Clustering the neighbor

Cluster the Neighbour

```
# set number of clusters
kclusters = 3

Etobicoke_grouped_clustering = Etobicoke_grouped.drop('Neighbourhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(Etobicoke_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]
# add clustering labels
```

```
]:  array([0, 0, 2, 0, 0, 0, 0, 1, 0, 0], dtype=int32)
```

The final Clustered output that contain population and most common 10 places in the neighborhood.

:

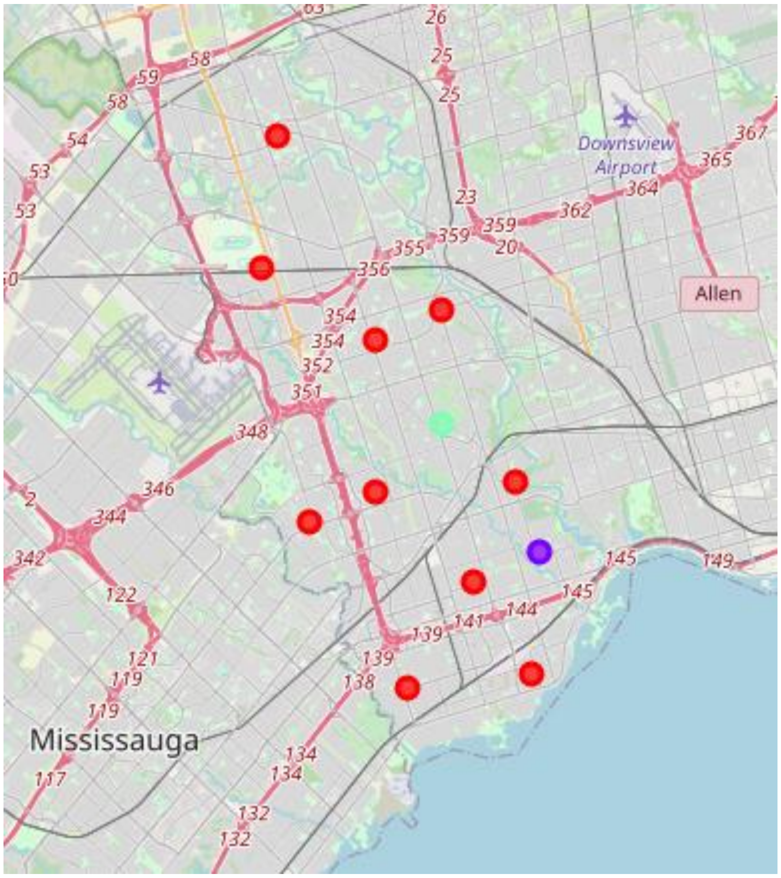
	Postal Code	Borough	Neighbourhood	Latitude	Longitude	Population	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	M8W	Etobicoke	Alderwood, Long Branch	43.602414	-79.543484	20674.0	0	Pizza Place	Gym	Coffee Shop	Sandwich Place	Pharmacy	Pub	Fast Food Restaurant	Electronics Store
1	M9C	Etobicoke	Eringate, Bloordale Gardens, Old Burnhamthorpe...	43.643515	-79.577201	38291.0	0	Coffee Shop	Pizza Place	Electronics Store	Beer Store	Shopping Plaza	Liquor Store	Café	Convenience Store
2	M9A	Etobicoke	Islington Avenue, Humber Valley Village	43.667856	-79.532242	35594.0	2	Pharmacy	Wings Joint	Construction & Landscaping	Fried Chicken Joint	Flea Market	Fast Food Restaurant	Electronics Store	Drugstore
3	M9R	Etobicoke	Kingsview Village, St. Phillips, Martin Grove ...	43.688905	-79.554724	33743.0	0	Pizza Place	Beer Store	Sandwich Place	Bus Line	Chinese Restaurant	Pharmacy	Wings Joint	Flea Market
4	M8Z	Etobicoke	Mimico NW, The Queensway West, South of Bloor,...	43.628841	-79.520999	17038.0	0	Wings Joint	Kids Store	Bakery	Burger Joint	Burrito Place	Convenience Store	Discount Store	Fast Food Restaurant

Work Flow:

Using credentials of Foursquare API features of near-by places of the neighborhoods would be mined. Due to http request limitations the number of places per neighborhood parameter would reasonably be set to 100 and the radius parameter would be set to 1000.

4. Results Section

Clustered Map of Etobicoke on the bases of the Above Features



Cluster 1

```
In [45]: Etobicoke_merged.loc[Etobicoke_merged['Cluster Labels'] == 0, Etobicoke_merged.columns[[1] + list(range(5, Etobicoke_merged.shape[1]))]]
```

Out[45]:

	Borough	Population	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Etobicoke	20674.0	0	Pizza Place	Gym	Coffee Shop	Sandwich Place	Pharmacy	Pub	Fast Food Restaurant	Electronics Store	Drugstore	Discount Store
1	Etobicoke	38291.0	0	Coffee Shop	Pizza Place	Electronics Store	Beer Store	Shopping Plaza	Liquor Store	Café	Convenience Store	Pharmacy	Fast Food Restaurant
3	Etobicoke	33743.0	0	Pizza Place	Beer Store	Sandwich Place	Bus Line	Chinese Restaurant	Pharmacy	Wings Joint	Flea Market	Fast Food Restaurant	Electronics Store
4	Etobicoke	17038.0	0	Wings Joint	Kids Store	Bakery	Burger Joint	Burrito Place	Convenience Store	Discount Store	Fast Food Restaurant	Grocery Store	Thrift / Vintage Store
5	Etobicoke	37975.0	0	Café	Coffee Shop	Mexican Restaurant	Gym	Pizza Place	Bakery	Dessert Shop	Fast Food Restaurant	Fried Chicken Joint	Grocery Store
6	Etobicoke	40684.0	0	Garden Center	Bar	Drugstore	Rental Car Location	Construction & Landscaping	Fried Chicken Joint	Flea Market	Fast Food Restaurant	Electronics Store	Discount Store
8	Etobicoke	55959.0	0	Pizza Place	Grocery Store	Fast Food Restaurant	Beer Store	Sandwich Place	Liquor Store	Fried Chicken Joint	Pharmacy	Wings Joint	Electronics Store
9	Etobicoke	10787.0	0	Bakery	Pool	River	Wings Joint	Flea Market	Fast Food Restaurant	Electronics Store	Drugstore	Discount Store	Dessert Shop

Cluster2

```
Etobicoke_merged.loc[Etobicoke_merged['Cluster Labels'] == 1, Etobicoke_merged.columns[[1] + list(range(5, Etobicoke_merged.shape[1]))]]
```

]:

	Borough	Population	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
7	Etobicoke	21299.0	1	Construction & Landscaping	Baseball Field	Wings Joint	Fried Chicken Joint	Flea Market	Fast Food Restaurant	Electronics Store	Drugstore	Discount Store	Dessert Shop

Cluster3

```
Etobicoke_merged.loc[Etobicoke_merged['Cluster Labels'] == 2, Etobicoke_merged.columns[[1] + list(range(5, Etobicoke_merged.shape[1]))]]
```

']:

	Borough	Population	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
2	Etobicoke	35594.0	2	Pharmacy	Wings Joint	Construction & Landscaping	Fried Chicken Joint	Flea Market	Fast Food Restaurant	Electronics Store	Drugstore	Discount Store	Dessert Shop

Foursquare API:

This Capstone project have used Four-square API as its prime data gathering source as it has a database of millions of places, especially their places API which provides the ability to perform location search, location sharing and details about a business.

5. Discussion Section

Problem Which Tried to Solve: The major purpose of this project is to suggest a better neighborhood in a new city for the person who wanted to open a new restaurant. The Problem is solved by using taking

- The population of the neighborhood
- The frequency of different categories of the location

6. Conclusion Section

In this Capstone project, using k-means cluster algorithm I separated the neighborhood into 3 different clusters from the dataset, which have very-similar neighborhoods around them based on Population. From the tables above shows that cluster n=2 is the best neighbor for start a new restaurant.

In Cluster N=0:
The Population is too high and most common venues are resturants
In Cluster N=1
The Population as compare to other neighbor its too low
In Cluster N=2
The Population is too high but the most common venues are not restaurant.

Thus from the above analysis we can conclude that cluster =2 Islington Avenue, Humber Valley Village is best place in Etobicoke is the best place to start a new resturant .

I feel rewarded with the efforts and believe this course with all the topics covered is well worthy of appreciation.
This project has shown me a practical application to resolve a real situation that has impacting personal and financial impact using Data Science tools.
The mapping with Folium is a very powerful technique to consolidate information and make the analysis and decision better with confidence.

Future Works:

This Capstone project can be continued of best type of restaurant(whether place need chines,indian …….resturants) . For that we again need more immigrant and other data……