**CAPSTONE PROJECT**

**BATTLE OF THE NEIGHBOURHOOD**

**INTRODUCTION:**

In this final assignment we choose a problem such that it must possess a simple methodologies which we had been studied in these 9 courses and incorporate in this final report , defining a problem as follows

**Business Problem:**

A Asian restaurant food joint thinks of opening few stores in and around Toronto region; as they possess a good brand value around Asian regions, they want to concentrate much of that region people. Such that much of the stores will attract that people from that region. There are several factors we need to consider while solving this problem. In the initial stages they are thinking to open 3 stores and then based on the response they are willing to open subsequent stores.

**PROBLEMS TO ADDRESS:**

1. They need to find the subsequent locations where they have to start the restaurants based on the location of people in that region .
2. They have to choose the locations mostly where the Asian population is higher.in order to do that we could locate the similar Asian food joints too from the foursquare location data apps.
3. They need got to identify the other type of food joints present in that regions and study ther profitability (which is hard mostly and it can be neglected if needed)
4. Using Four square data location apps we need to locate the prime locations in the neighbourhood where similar food joints are kept and prioritize the regions accordingly

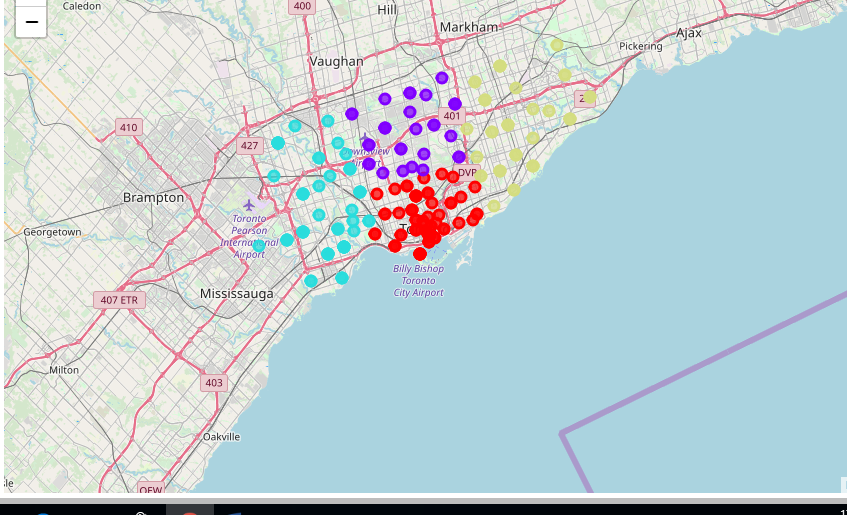
**Data**

The data required is available on the Wikipedia page and foursquare API can be used.

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

**Methodology:**

1. We use the data we acquired from the four square location data app and then render it accordingly to create different clusters which are separated by about a distance of about 1000 m apart.
2. Using the neighbourhood postal codes we do the above process and in this problem we divide into 4-clusters .
3. We then use the forlim df such that we map and colour the cluster point locations as shown below



1. Then, we data set the cluster locations use four square data location and then divide into four clusters here.
2. We then obtains venues for each cluster and then we pick the top three location venues by calling food in the data frame as there is not Asian returants specified in the data set.
3. After this stage, the column "Venue's Category" will be One-hot encoded and different venues will have different feature-columns. After On-hot encoding we will integrate all venues columns to one column and joint columns to "Total Joints" column. After applying k-Means Clustering of Machine Learning Techniques.
4. Thus, from above mentioned methodologies we have to carry out the step to identify the top location venues for the other three clusters.

**RESULTS:**

from performing the following methodology we otain the following results for the clusters as

1. In cluster-1 top 3 picks in this cluster are vicroria hotel, commerce court, esplanade.
2. In cluster-2 top three picks is choosen i.e oriole, fairview, henry farm
3. In cluster-3 top three picks is choosen i.e swasan,high park, junction south.
4. In the final cluster only top 3 has more than 10 common points , so top 3 is choosen i.e steelies west, woodbin gardens, L'Amoureax west.

**DISCUSSION:**

As we want to start a new Asian restaurant joint in that region and there no other one we choose to fint the regions with high fallout and it can we justified using the explainations as well as we could be sure of people visiting that venues such that there is possibility of being an Asian around that region to and to let them know as prime locations gives an added advantage.

Thus, we could see that places with above results gives us an idea about spread of people in that region.so, we could end it saying it is advisable to start a restaurant over that region venues.

**CONCULSION:**

Finally, we can conclude by saying that the following requirements are met as defined and problem is solved accordingly using python labs.