Capstone Project - The Battle of Neighbourhoods Report

Lakshmi sri sai , August,2020

**1. Introduction**

**1.1 Background**

Getting through lot many options while searching for houses has became very important these days, as people looking for lot many comforts/facilities like nearby Grocery stores, hospitals,School,Super market,medical shops etc. Going through each and every place and searching houses with all these facilities is not possible and take too much of time.

This project will help in such cases like prediction of houses with nearby facilities around their neighborhood in in Central Toranto.This project will people to get awareness of the area and neighborhood before moving to a new city, state, country or place for their work or to start a new fresh life.

This Project include analysis of features for a people migrating to North york to search a best neighborhood. The features include median housing price and better school according to ratings, crime rates of that particular area, road connectivity, weather conditions, medical facailities for emergency, water resources both fresh and waste water etc.

**1.2 Problem**

1. sorting the list of houses with nearby better facilities

2. finding the best house in central toronto neighbourhoods

**1.3 Interest**

People who are considering to relocate to Central Toronto will be interested to identify the safest neighbourhood in Central toronto and explore its neighborhoods and common venues around each neighbourhood.

**2. Data Acquisition and Cleaning**

**2.1 Data Acquisition**

The data acquired for this project is a combination of data from three sources. The first data source of the project uses a List of postal codes of Canada("https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of\_Canada:\_M") that shows the borough and neighbourhoods in central toronto. The dataset contains the following columns:

Postal codes: postal codes of all Canadas boroughs

Borough: names of different boroughs in totonto

Neighbourhoods: areas in that boroughs

The second source of data is scraped from a wikipedia page that contains the list of latitudes and longitudes for boroughs(<https://cocl.us/Geospatial_data>). This page contains additional information about the boroughs, the following are the columns:

Postal codes: postal codes of boroughs

Latitude : latitude of boroughs

Longitude: longitudes of boroughs

The third data source is the list of venues in Canada around their boroughs found from Foursquare API . This dataset contains the following coloumns:

Venue name

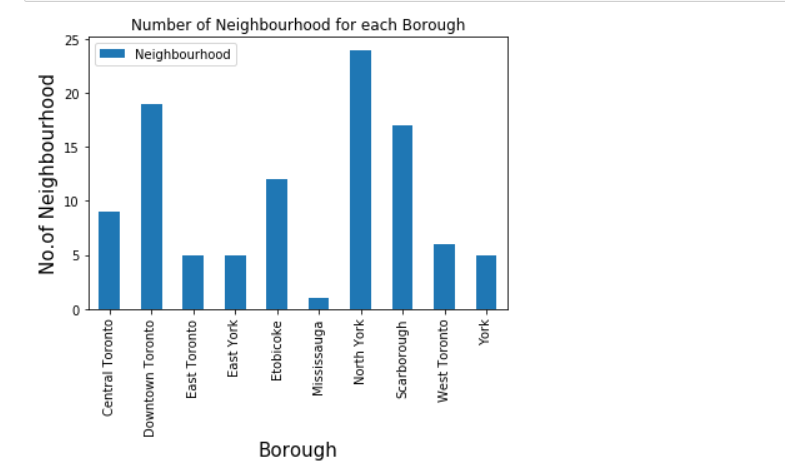
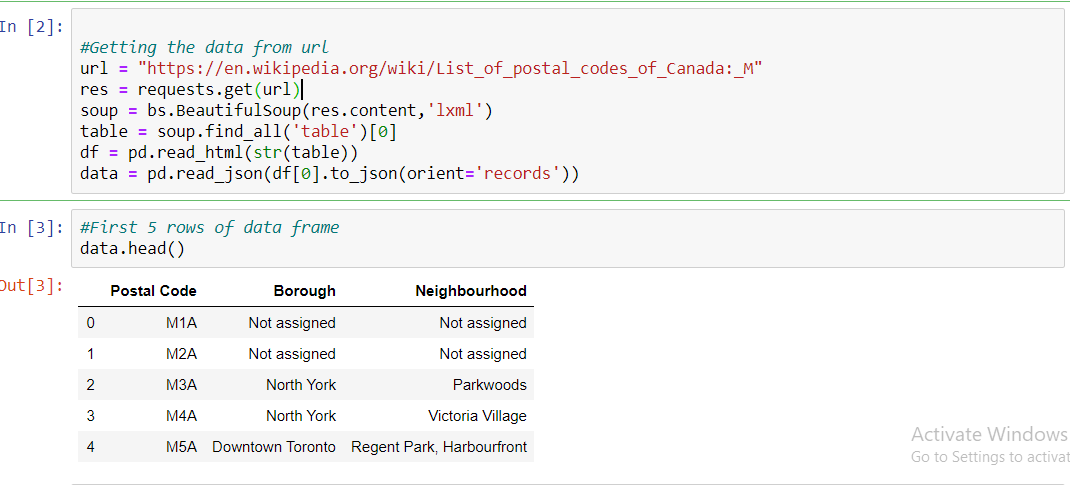
Venue category

Venue location latitude

Venue location longitude

**2.2 Data Cleaning**

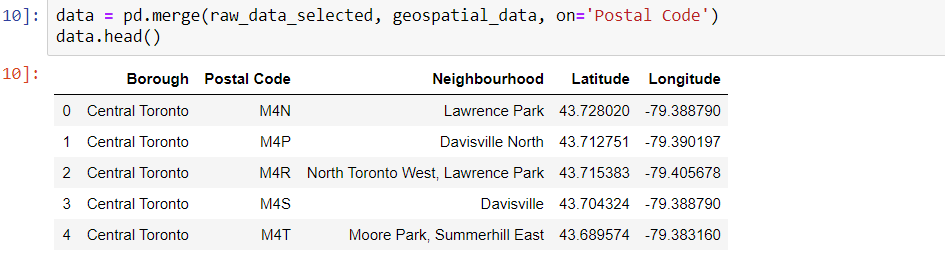
The data preparation for each of the three sources of data is done separately. From list of postal codes of Canada we will make use of all coloumns. This data is scraped from a wikipedia page using the Beautiful Soup library in python. Using this library we can extract the data in the tabular format as shown in the website. This is important because we will be merging the two datasets together using the Borough names and visuialize them the following pictures referes the data.



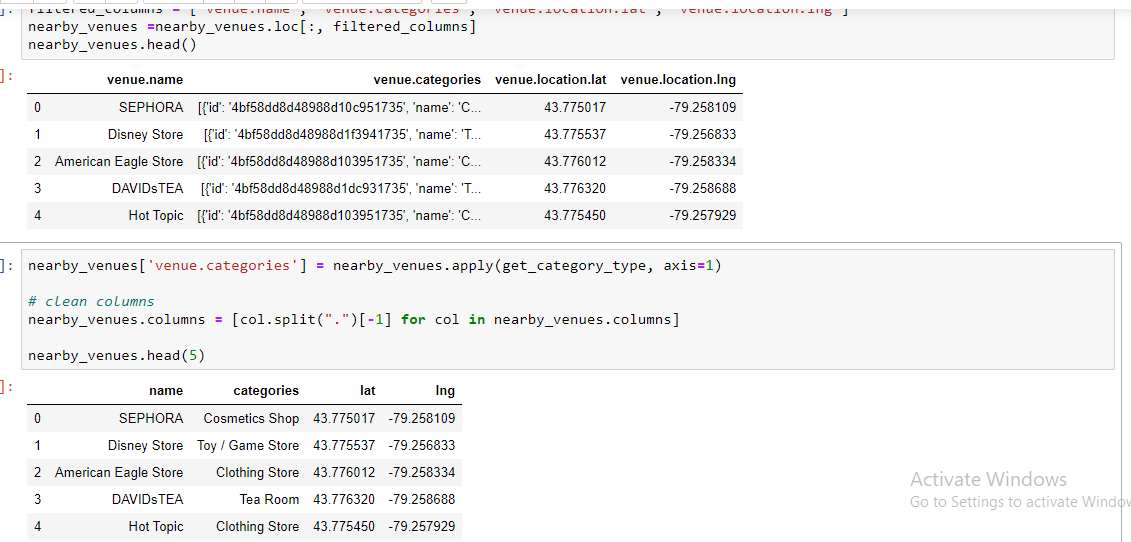
We secod phase of data from web by this link <https://cocl.us/Geospatial_data> and we will reaad that with the help of pandas.



We will merege the abouve two data frames with the help of merger() with the help of postal codes coloumn.



After that we will get thrird data phase with the help of Foursquare account of mine and get the following dataset



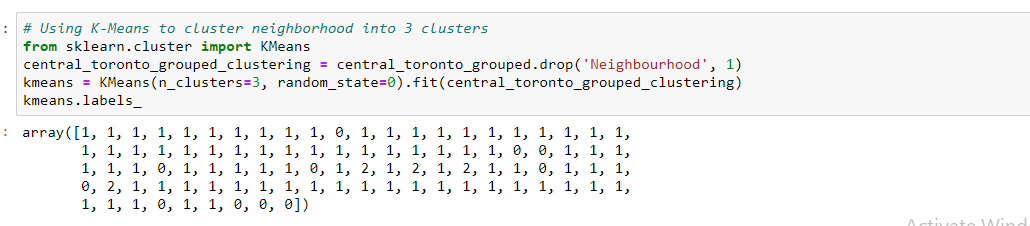
After that we will sort out the venues around the neighbourhoods.

**3.methodology**

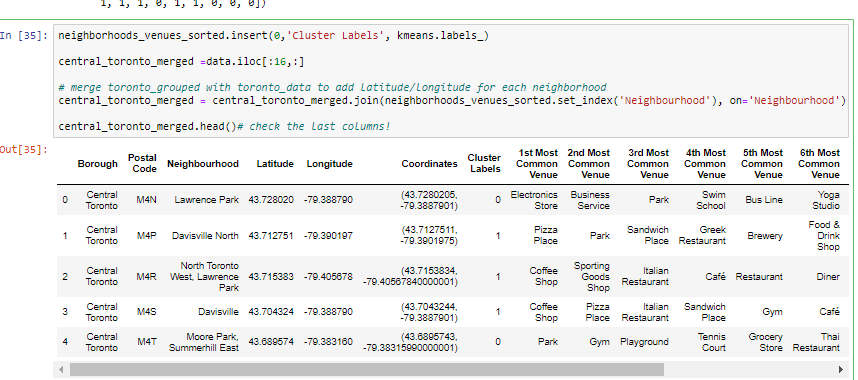
**3.2 Modelling**



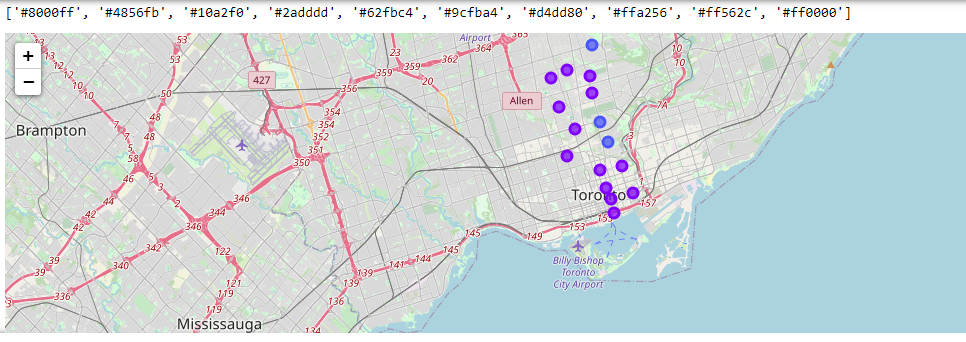
Using K-Means to cluster neighborhood into 3 clusters



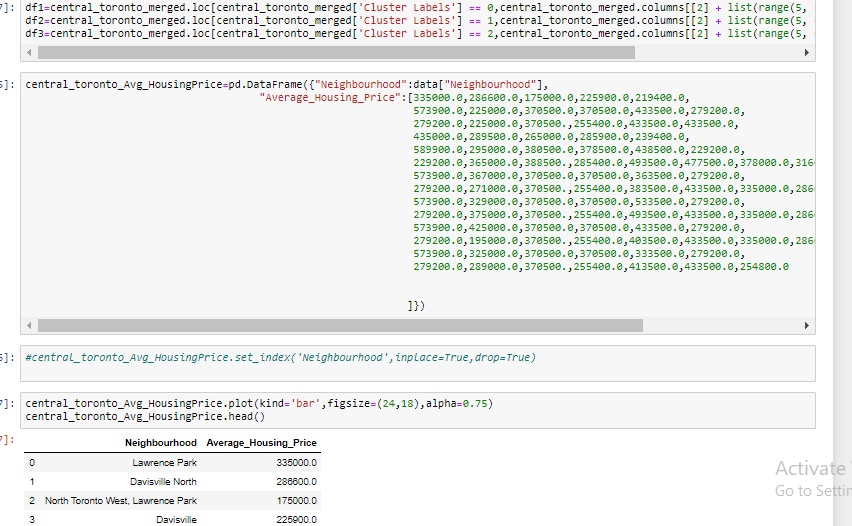
merge toronto\_grouped with toronto\_data to add latitude/longitude for each neighborhood



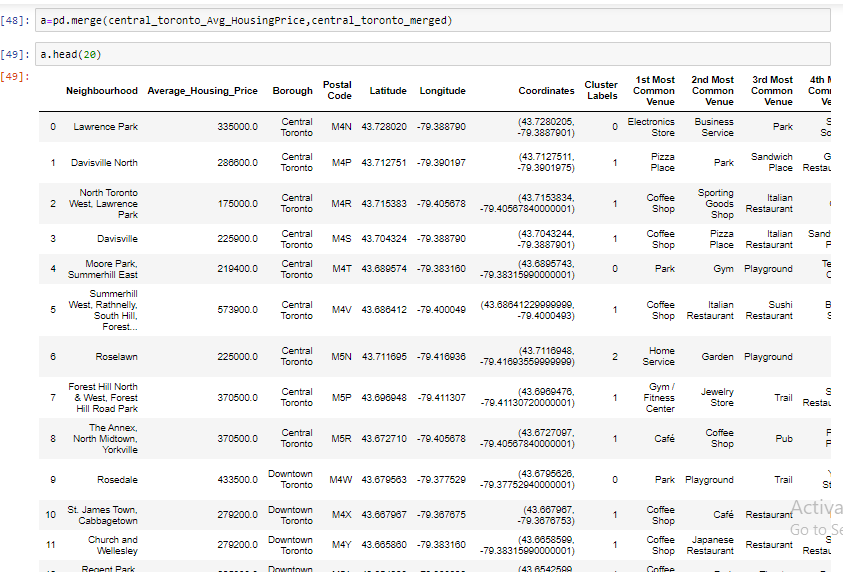
**map of clusters**



adding prices of houses of neighbourhoods in respective boroughs and getting 1st 5 rows of the dataset.



Mergeing the data sets with house price and the data sets with locations venue etc

****

**4. Results**

From 9 neighbourhoods in central totonto we analyse the prices at eac neighbourhood with suitable facilities among them top 3 best places are

1. North Toronto West, Lawrence Park-175000.0-Central Toronto

2. Roselawn-225900.0-Central Toronto

3. Davisville-225900.0-Central Toronto