**Suitable Place For Pharmacy In Toronto**

1. **Introduction**
   1. **Background**

ABC Pharmacy is one of the leading pharmaceutical company in the world and wanted to setup some stores in Toronto. They are finding the best possible location in Toronto to setup their store.

* 1. **Problem**

We need to find the best possible place in Toronto for ABC Pharmacy which can help them to generate maximum revenue, by analysing the entire city and its neighbour.

1. **Data Acquisition & Cleaning**
   1. **Data Source**

We would need the data of all locality info of Toronto, local pharmacy, locality information, etc. This data we can get from <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M> , foursquare.com & [http://cocl.us](https://cocl.us/Geospatial_data). We will get all the required information from here.

* 1. **Feature Selection & Data Cleaning**

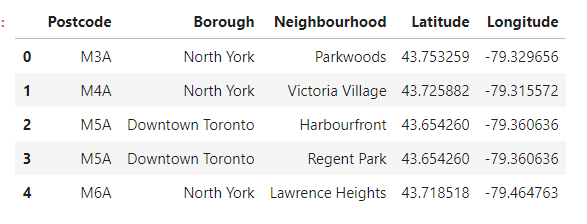
We will combine the data from all the sources into one table and then take 5 feature out of that are Postal code, Borough, Neighbourhood, Longitude & Latitude. After this we will clean and process all NaN. We will also capture the data for all the public places, stores, parks, etc for every locality.

* 1. **Data Utilisation**

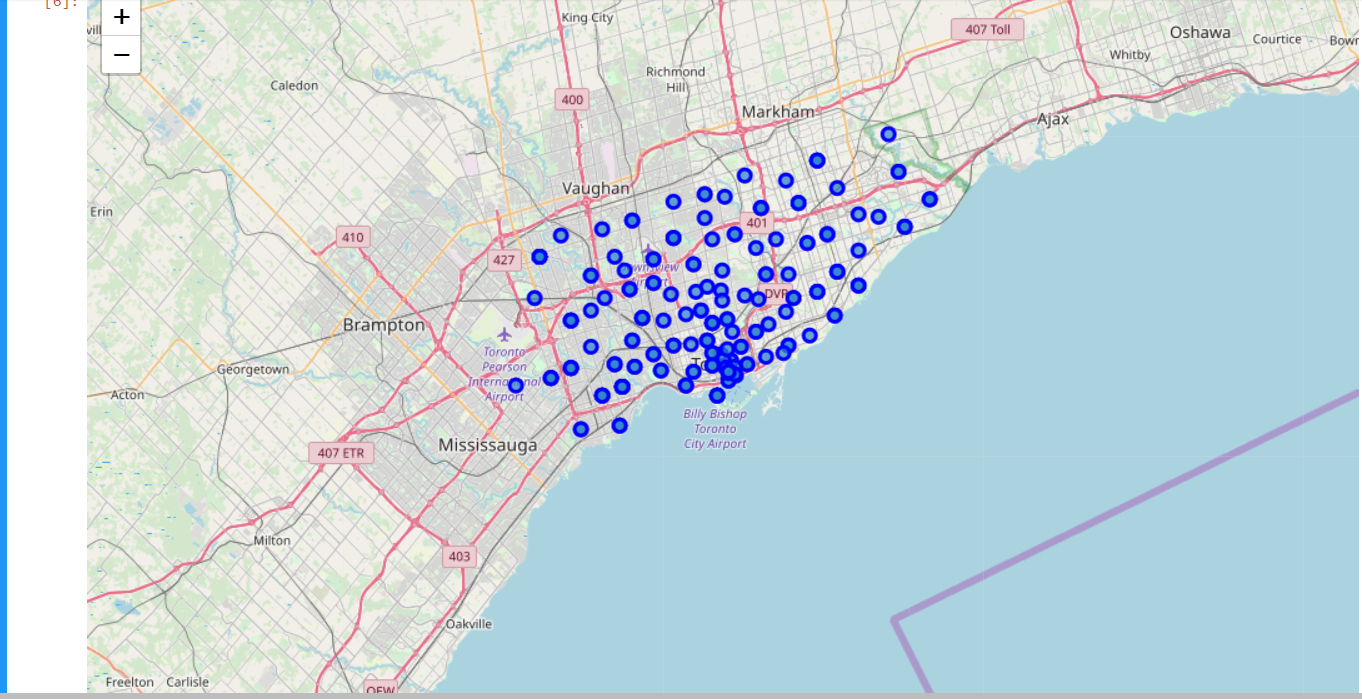
After cleaning the data we will try to figure the place where number the number of pharmacies are very less and the other localities like sports club, residential societies, grocery market, malls, etc are near. So that the place is crowded enough to attract the customers

1. **Modelling**

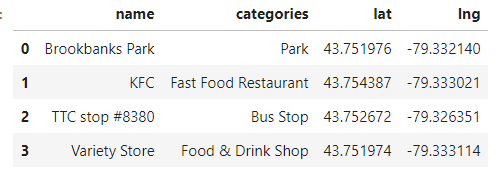
As a database, I used GitHub repository in my study. My master data which has the main components *Postal code, Borough, Neighbourhood, Latitude & Longitude* information of the city.



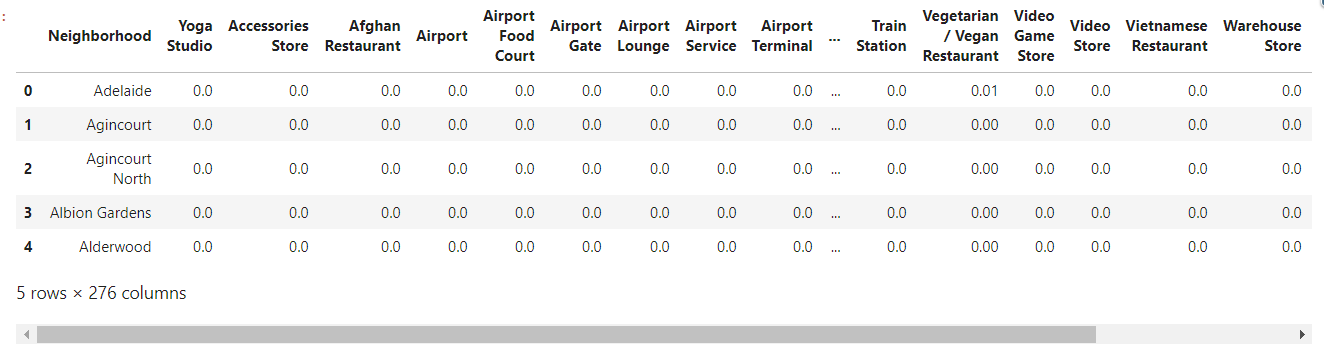
I used python **folium** library to visualize geographic details of Toronto and its boroughs and I created a map of Toronto with boroughs superimposed on top. I used latitude and longitude values of postal code to get the visual as below:



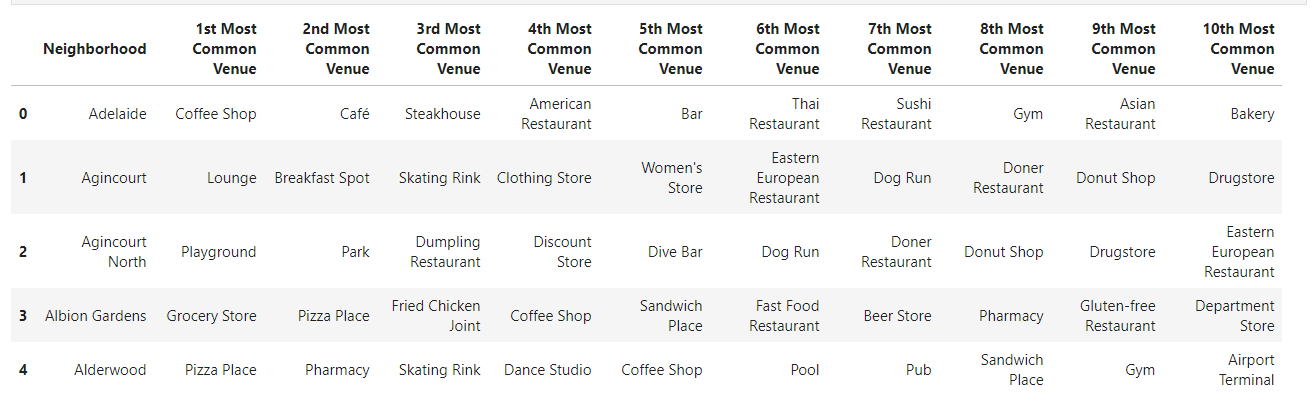
I utilized the Foursquare API to explore the boroughs and segment them. I designed the limit as **100 venue** and the radius **500 meter** for each borough from their given latitude and longitude information. Here is a head of the list Venues name, category, latitude and longitude information from Forsquare API.



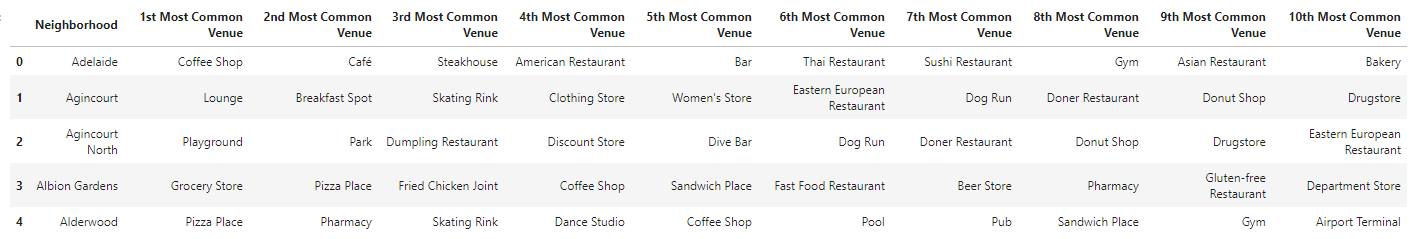
Now as I have the Venue details of all postal code of Toronto I will determine the frequency of occurrence of each venue in each Neighbourhood



Now I have created a dataframe to display the top 10 venue of each neighbourhood.

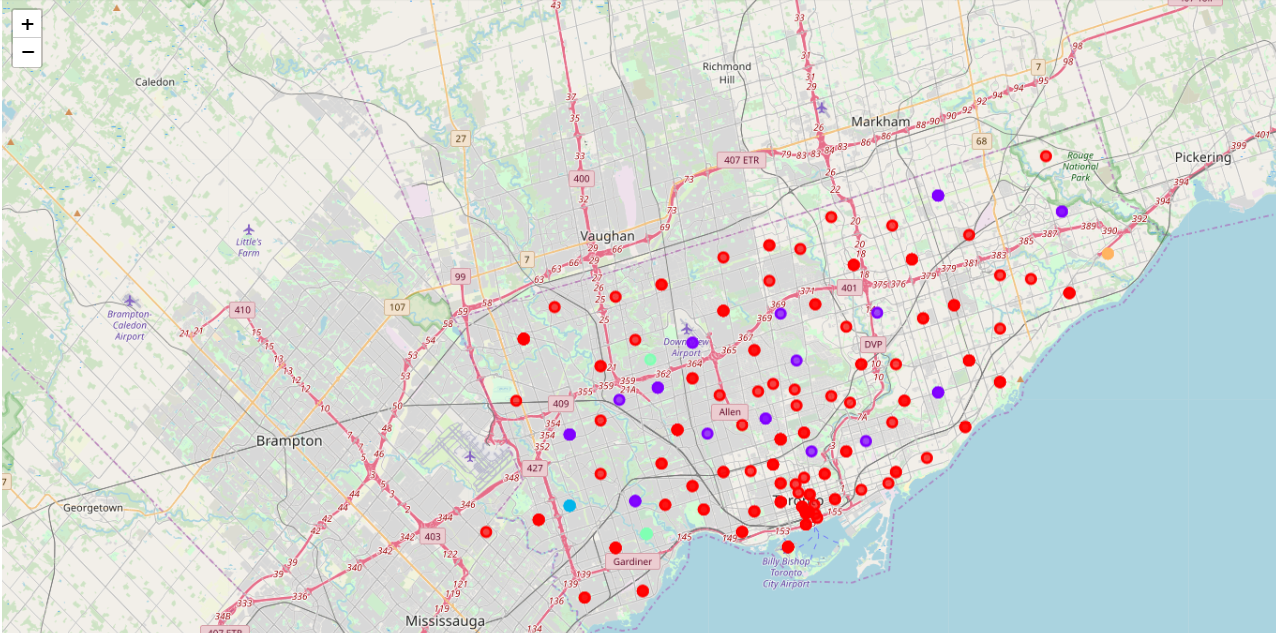


We have some common venue categories in boroughs. In this reason I used unsupervised learning **K-means algorithm** to cluster the boroughs. K-Means algorithm is one of the most common cluster method of unsupervised learning. Here is my merged table with cluster labels for each borough

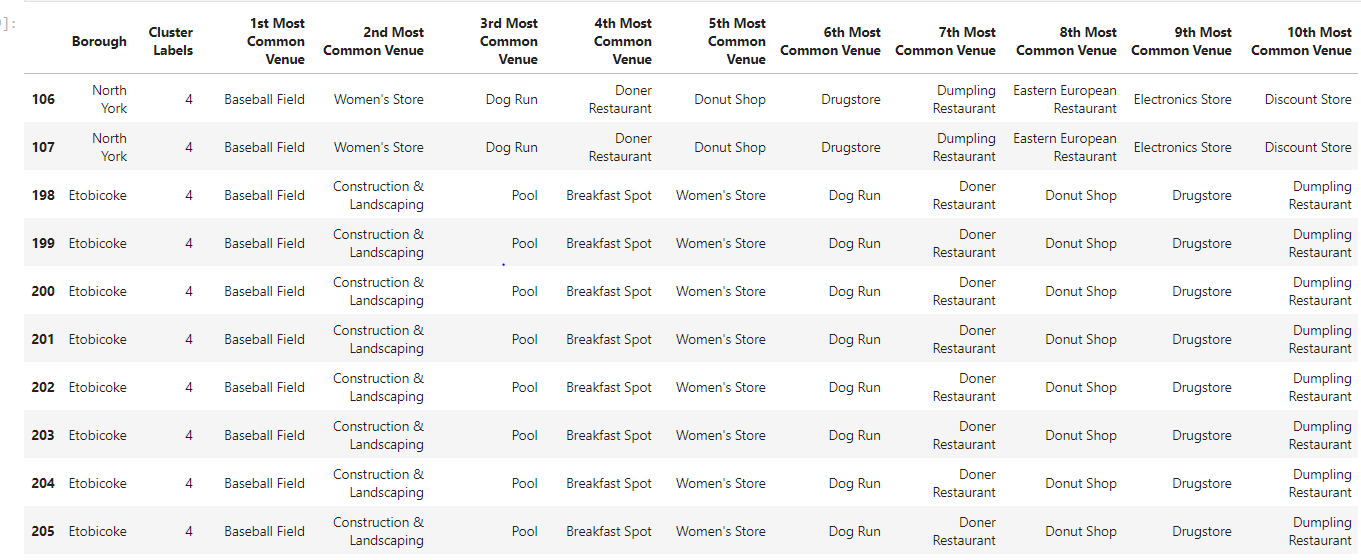


1. **Results**

Below is the clustered map of Toronto.



Below is the list of all the Brought where we do not have pharmacy in top 10 venues



1. **Discussion**

In the above result have found 2 localities in Toronto where pharmacy is not in top 10 venues they are North Youk & Etobicoke. We can also observe from the data that Baseball Field, Construction & Landscaping & Pool are amongst the top 3 venues. In such places there are a lot of chances of accident.

1. **Conclusion**

As the pharmacy is not so easily available in North Youk & Etobicoke and these places have venues where accident can happen very easily & also as the construction sites are here so we expect more danced area in near future so I would recommend ABC Pharmacy to setup there store Somewhere here.