Predicting A New Neighbourhood In Toronto,ON

INTRODUCTION

This project helps people who are shifting from one place to toronto and are looking for an ideal place,which is similar to the one they are currently staying in.This project uses a KMeans clustering model for finding the best neighbourhood, as KMeans is an excellent unsupervised learning algorithm

This project also helps in understanding the futre prospects in the city, as it shows us the types of places with the highest volume

DATA

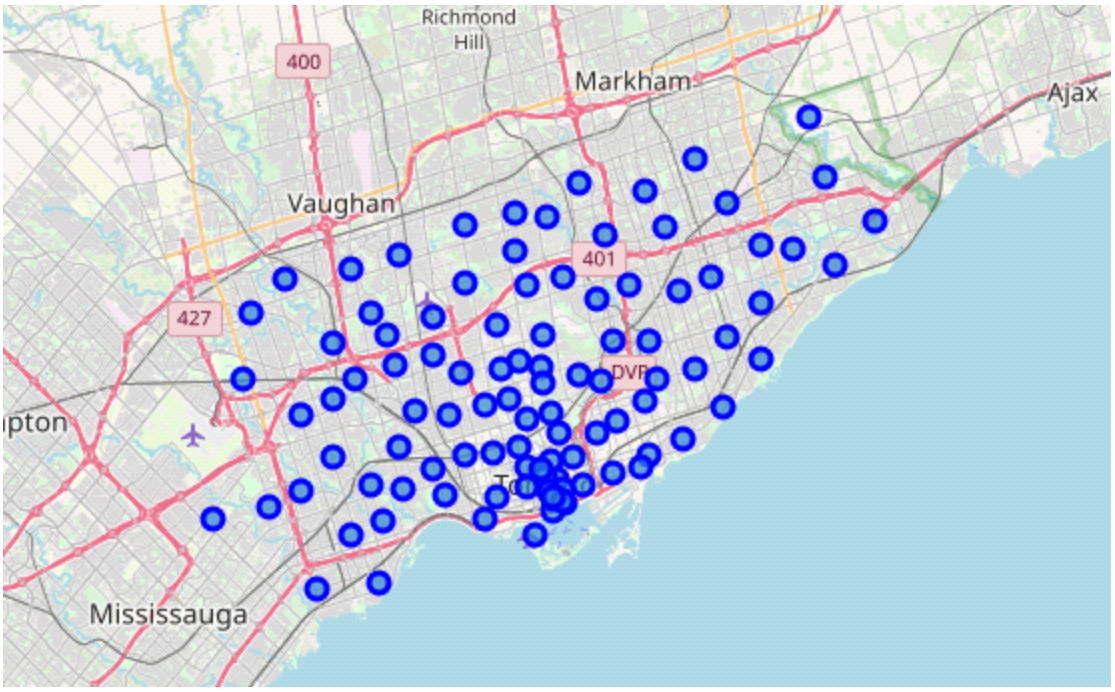
The data required for training the model was obtained from a Wikipedia page, the link to the data used is:-

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

This is a list of postal codes in Canada where the first letter is M. Postal codes beginning with M are located within the city of Toronto in the province of Ontario. Only the first three characters are listed, corresponding to the Forward Sortation Area.

A lot of steps were taken to clean the scrapped data and add other features like latitude and longitude values, Foursquare API was used to find the details about different venues, their type and other information

Using the scrapped data a map was formed to get some basic visual understanding



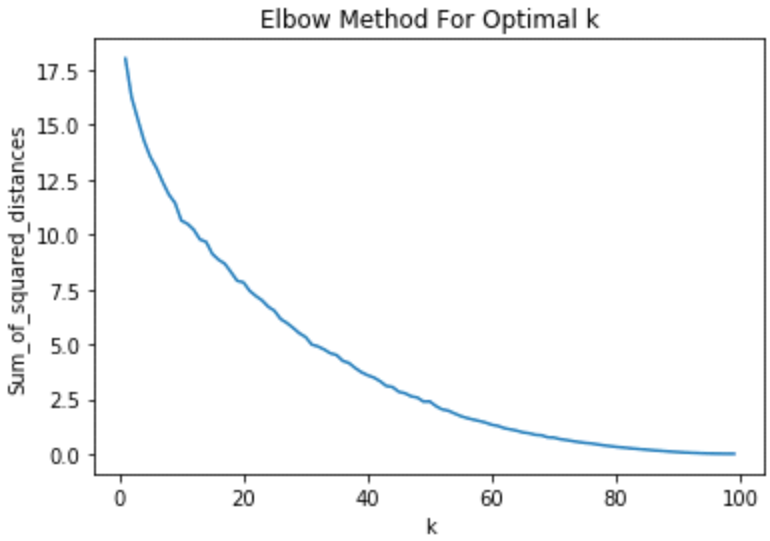
METHODOLOGY

After getting the basic data like the neighbourhoods, boroughs and postal codes, some steps were taken to make this data more usable

Firstly, the entries without a borough were removed, then the entries containing a borough but no neighbourhood were corrected by assigning the corresponding borough value to the neighbourhood part.

Then the latitude and longitude values were added, proceeded by aquiring information about different venues and convrting them to a usable dataframe.

A 'KMeans' machine learning model was trained as it is one of the best unsupervised learning algoritm, an appropriate 'n_clusters' value was obtained by using the elbow method.

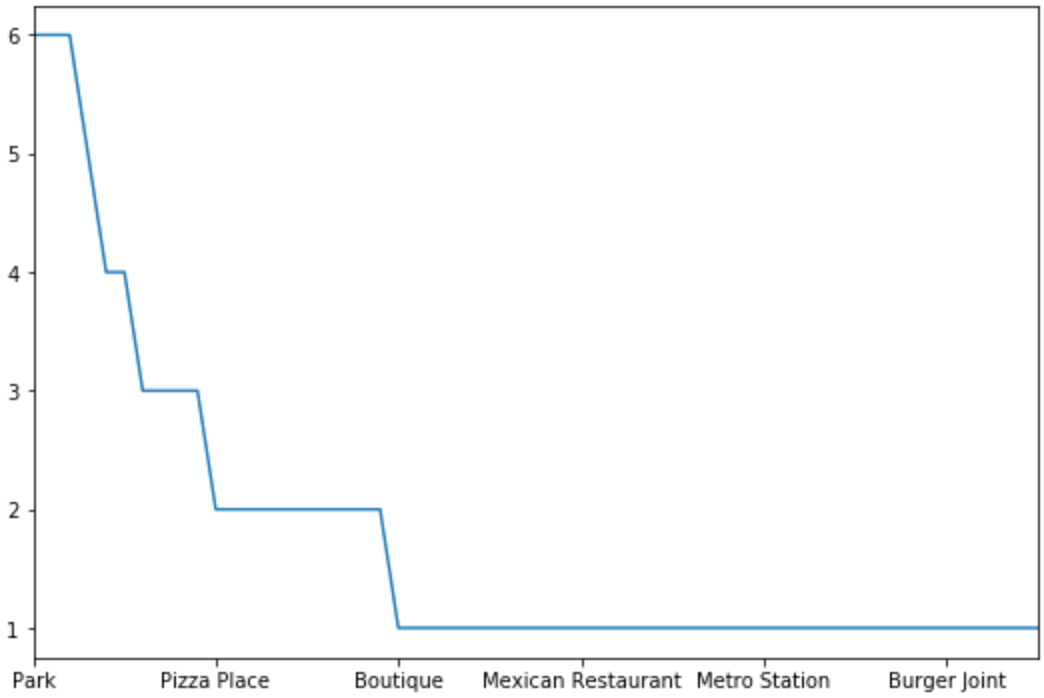


A dataframe was also obtained of the five most common venues, which can give us a better understanding about that particular place.

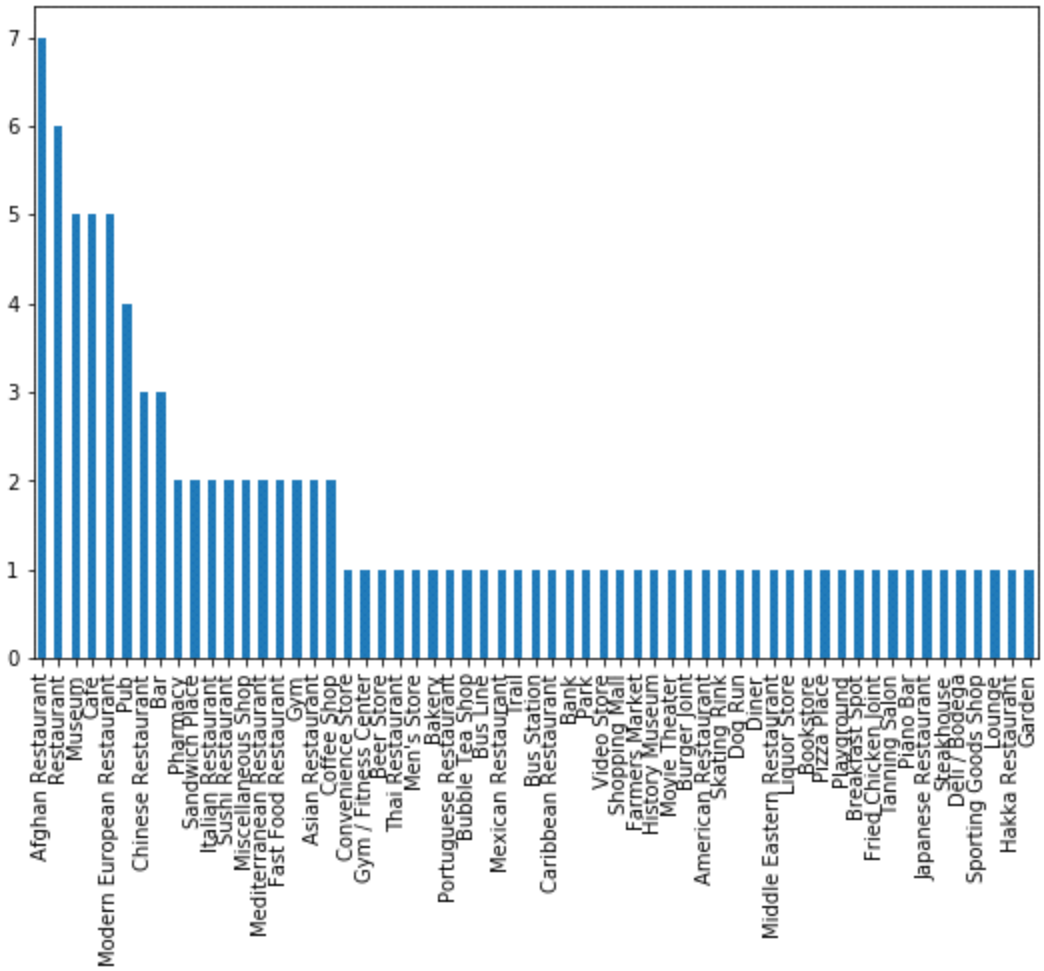
RESULTS

a proper predicting model was obtained which could be used for obtainning correct information about the possible futre scenarios and what kind of atmosphere to expect and what possible business outlets are favourable

- The most frequent outlets:-



- The fifth most frequent outlets:-



DISCUSSION

From this project it can be observed that for a person moving to a specific area in Toronto,ON, he can expect a specific amount of competetion, for example:-

A person opening a 'Pizza joint' will have to face significantly more competetion than a 'Japanese Restaurant' and thea simillar case can be expected in case of a 'Birger Joint'

For a person expecting a simillar place to stay to his existing place he can obtain that information by using thi KMeans model.

CONCLUSION

From this project it can be concluded that as far as unsupervised learning is concerned KMeans algorithm is a safe option for finding accurate predictions