BATTLING OF NEIGHBORHOODS-REPORT

INTRODUCTION

Toronto is the provincial capital of Ontario. It is the most populous city in Canada and the fourth most populous city in North America. Toronto is an international centre of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world.It attracts people from all over the world.

Scarborough is the region of Toronto hugging the city's eastern fringe. Known by Toronto people as a mosaic of different ethnic cultures, Scarborough is also known for its breathtaking natural scenery.It is one of the most popular neighborhood.

PROBLEM

This Project aim to create an analysis of features for a people migrating to Scarborough to search a best neighborhood as a comparative analysis between neighborhoods. The features include median housing price and better school according to ratings, crime rates of that particular area, road connectivity, weather conditions, good management for emergency, water resources both fresh and waste water and excrement conveyed in sewers and recreational facilities.

Data Section

•

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

Will use Scarborough dataset which we scrapped from wikipedia on Week 3. Dataset consisting of latitude and longitude, zip codes.

FOURSQUARE API

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.

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:Neighborhood,Neighborhood Latitude, Neighborhood Longitude, Venue,Name of the venue e.g. the name of a store or restaurant, Venue Latitude, Venue Longitude, Venue Category

METHEDOLOGY

To compare the similarities of two cities, we decided to explore neighborhoods, segment them, and group them into clusters to find similar neighborhoods.K-means clustering algorithm was used.

CONCLUSION

In this project, I separated the neighborhood into 10(Ten) different clusters using k-means cluster algorithm, and for 103 different latitude and loNGITUDE from the dataset, which have very similar neighborhoods around it. Using the above charts the results were presented to a specific neighborhood based on average house prices and school ranking.