Capstone Project
The Battle of Neighborhoods

Introduction/Business Problems

Project Description

Migration is a tough task. Leaving a place to which a person has accustomed itself to brings in a lot of challenges.Many people migrating to various states of Canada require search of a good housing prices as well as good rating schools for their children.

The projects aim to create an analysis of features for a neighborhood as a comparative analysis between neighborhoods.

The features include median house price and school ratings, crime rates, weather conditions, recreational facilities.

This would help people to get awareness of the places before moving to a new country, state, city or place for their work or to start a new life

Selection criteria

For the purposes of this project, the definition of a good neighborhood is one that has an appreciable commercial presence within a given community as well as:

1. Compare median housing prices

2. Compare school ratings

The Location:

Scarborough is a popular destination for new immigrants in Canada to reside. As a result, it is one of the most diverse and multicultural areas in the Greater Toronto Area, being home to various religious groups and places of worship.

Foursquare API:
This project would use Four-square API as its prime data gathering source as it has a database of millions of places, especially their places API which provides the ability to perform location search, location sharing and details about a business.

Work Flow:

Using credentials of Foursquare API features of near-by places of the neighborhoods would be mined. Due to http request limitations the number of places per neighborhood parameter would

reasonably be set to 100 and the radius parameter would be set to 500.

Clustering Approach

To compare the similarities of two cities, we decided to explore neighborhoods, segment them, and group them into clusters to find similar neighborhoods in a big city like New York and Toronto. To be able to do that, we need to cluster data which is a form of unsupervised machine learning: k-means clustering algorithm

Libraries

Pandas,Folium,Scikit Learn,JSON,Geopy,Requests,Matplotlib