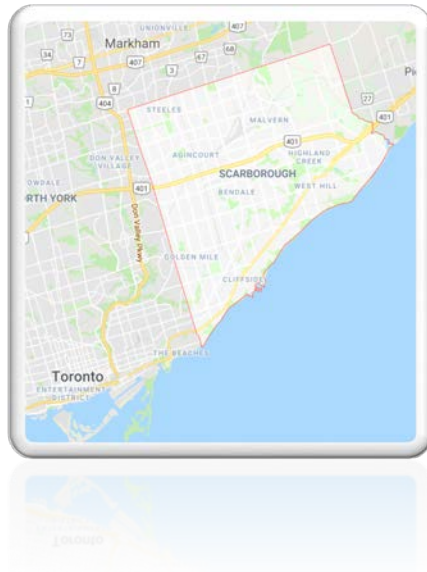


Capstone project – The Battle of Neighborhoods

Business Problems



Project Description

Every year, many people migrating to various states of Canada require to search for a great location with the good housing price; those families with kids are under even greater pressure to find an ideal neighborhood to suit their family's needs. Whether a neighborhood has great schools or is located in a coveted school district is the main consideration for 39 per cent of Canadians with aspirations to purchase a home.

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.

The project aims to create an analysis of features for 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. The aim of this project is to help the people explore different possibilities and make a better decision on choosing their best neighborhood out of many

neighborhoods in Scarborough city based on the distribution of various facilities in and around the neighborhood.

Selection criteria

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

- Compare median housing prices
- 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. Although immigration has become a hot topic over the past few years with more governments seeking more restrictions on immigrants and refugees, the general trend of immigration into Canada has been one of on the rise.

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: For creating and manipulating dataframes
- Folium: Python visualization library would be used to visualize the neighborhoods cluster distribution of using interactive leaflet map.
- Scikit Learn: For importing k-means clustering
- JSON: Library to handle JSON files
- Geopy: To retrieve Location Data
- Requests: Library to handle http requests
- Matplotlib: Python Plotting Module