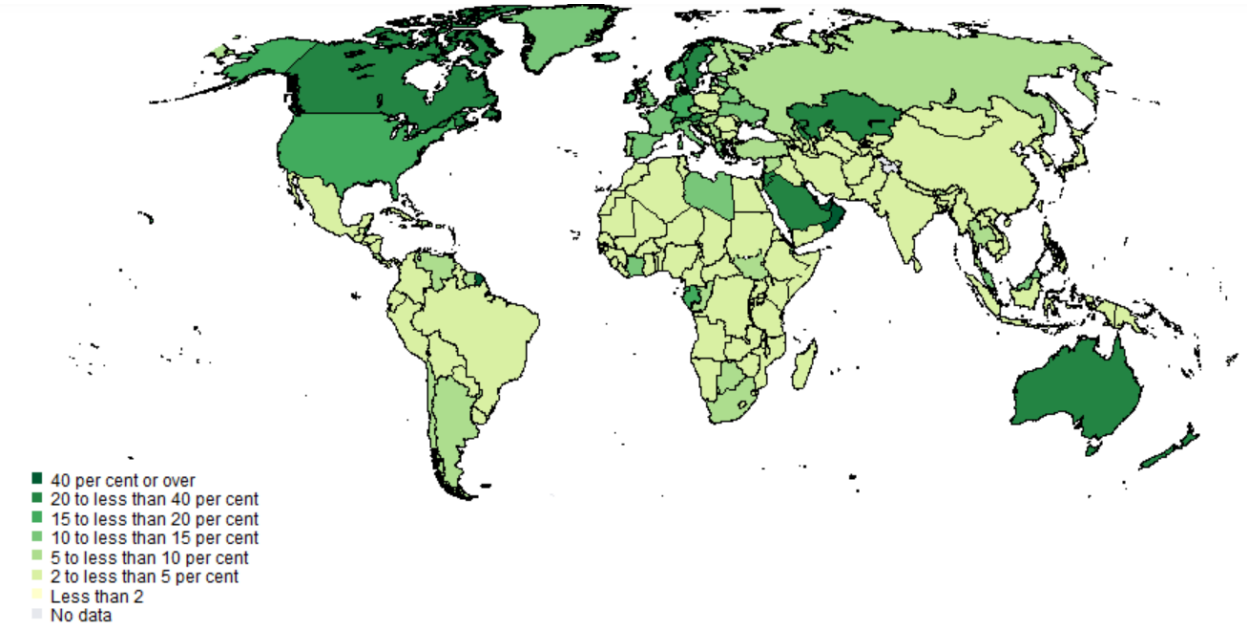
Capstone Project - Battle of Neighborhoods

North York, Toronto - Popular Neighbourhood Guidance to Immigrants

The Business Problem

Background:

According to US News, Canada is a top major destination for immigrants across the world. As per the latest statistics by the United Nations, between 20% to 40% of the Canadian Population in 2019 were immigrants.



Every immigrant looks for a better quality of life in the new country than her/his country of origin. A comfortable living with basic facilities are the needs of each immigrant families. Happy immigrants result in better and prosper.

Project Description:

As mentioned in the background, we are trying to solve a real-life macro problem of thousands of immigrants flocking to Canada. Immigrants coming to Canada needs a livelihood and hence stays near the economic centre of the country. Toronto becomes the natural choice for immigrants as per Wikipedia the city 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.

However, an immigrant moving to a large city like Toronto needs to decide the neighbourhood to stay out of the various choices.

The major determinable factors to settle in would be **housing prices** and **ratings of schools**. The end result provides users an overview of the places before they are moving to Toronto as their new country or city as their place for work or to began a new life.

In summary, this Project would assist an immigrant or users have more informed decision on choosing the best neighborhood out of many neighborhoods to move into Tornoto city based on the housing prices and school ratings.

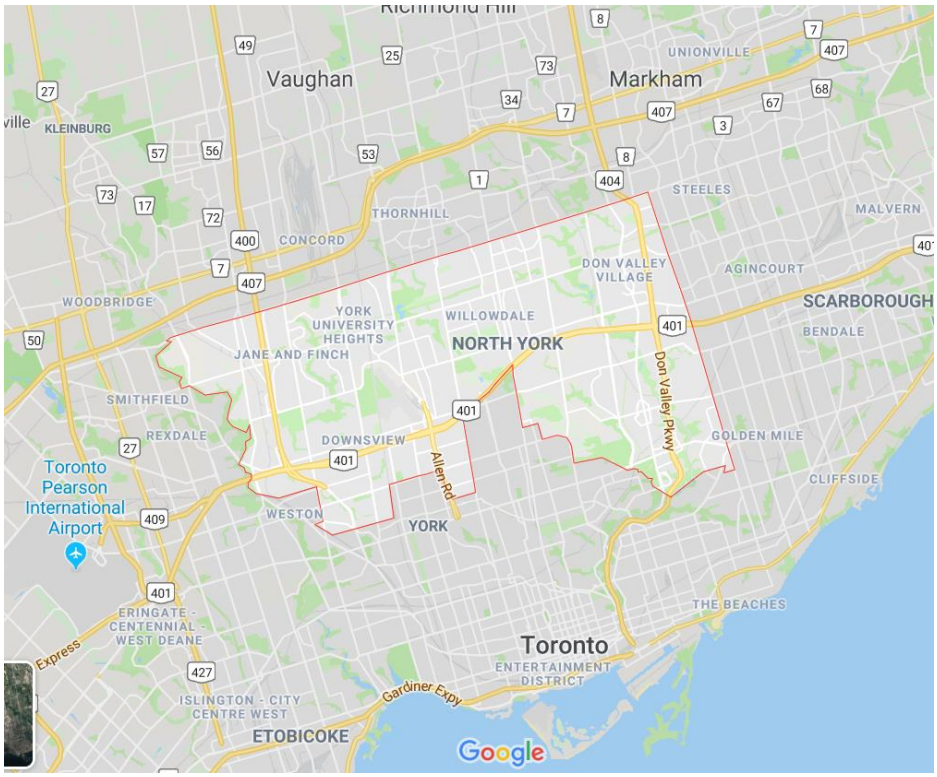
Selection criteria

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:

North York 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. Again referring to Wikipedia over the late 20th century and early 21st century, North York City Centre have emerged as secondary business districts outside Downtown Toronto. High-rise development in these areas has given the former municipalities distinguishable skylines of their own with high-density transit corridors serving them.
Thus, this projects aim to create an analysis of features for **North York neighborhood**. The features include like median house price, school ratings, population rate, crime rates, recreational facilities, etc.



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 cities 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:

List of Python libraries in alphabetical order with a brief overview

Folium: Python visualization library would be used to visualize the neighborhoods cluster distribution of using interactive leaflet map.

Geocoder: To locate the coordinates of a given addresses

Geopy: To retrieve Location Data

JSON: Library to handle JSON files

Matplotlib: Python Plotting Module

Pandas: For creating and manipulating dataframes

Requests: Library to handle http requests

Scikit Learn: For importing k-means clustering

XML: To parse and modify XML document