Introduction and Discussion of Problem

People migrate for a variety of reasons. People move in search of better career and work opportunities or they may move for personal reasons. This migration may be between different cities or within the same city. It may be domestic or international. According to the World Migration Report 2020, there are 164 million migrant workers in the world.

Whatever be the cause of relocation, finding suitable accommodation in a suitable neighborhood is a crucial decision that people who move have to make. The choice of locality is determined by various factors: housing rent prices, safety of neighborhood, availability of preferred cuisine, parks and fitness clubs, leisure opportunities etc. Depending on the individual, some of these factors will carry more weight over the others. However a good decision requires individuals to balance these criteria wisely.

This project aims to help migrants with this decision. Based on above parameters, this project will try and classify neighborhoods in the city of Toronto into distinct categories. Based on those categories, people can make data aware decisions in choosing the right neighborhood for them according to their preferences.

Whom will this analysis be beneficial for?

- Housing service providers
- Rentals/ leasing agencies
- Individuals/ families moving to a new place

Data Collection and Usage

In this project, I will use the following variables to cluster neighborhoods in the city of Toronto: average rent, safety index, cuisines variety index, recreational opportunities index, and entertainment options index. Hence, I have chosen the following data sources:

1. https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M, http://cocl.us/Geospatial_data (Coordinates of all neighborhoods in Toronto)

   These links contain information about - names of all the neighborhoods in Toronto and their geospatial coordinates. These will be used to set up the neighborhoods dataframe.

2. Foursquare API

   This API will be used to get information about different venues in a neighborhood in Toronto. This will provide data for three of the above variables. Number of different types of restaurants will be used to set up the cuisine variety index; number of parks, gyms and fitness studios will be used to set up the recreational opportunities index; number of theaters, nightclubs and bars will be used to set up the entertainment options index.

3. https://data.torontopolice.on.ca/datasets/neighbourhood-crime-rates-boundary-file-

   This website of the Toronto Police department contains incident counts for assault, auto theft, break and enter, robbery, theft over and homicide neighborhood-wise for the city of Toronto. These numbers will be used to set up the safety index.

4. https://rentals.ca/toronto

   This website contains listings of over 6000 properties advertised for rent in Toronto. Using web scraping, information about rent, address and neighborhood of the properties will be extracted. This will be used to evaluate average rent for a given neighborhood.