**Proposal and Introduction**

The purpose of this analysis is to help people exploring better facilities around their neighborhood. It will assist families in making better and decisions on selecting the right neighborhood in Scarborough to move to.

My project aims to create an analysis of housing prices, school ratings and crime rates. The major purpose of this project is to recommend better neighborhoods in Toronto for a family to move to with better connectivity to the airport, transportation, and markets.

The information will be sorted by:

1. housing prices in an ascending or descending order
2. schools in terms of location, fees, rating and reviews

Scarborough is a popular destination for young adults to reside in. As a result, it is a diverse and multicultural area and is home to various religious groups and places of worship.

I plan on using Four-square API as a prime data gathering source as it can perform location search, location sharing and details about businesses. Due to the request limitations the number of places per neighborhood parameter will be set to 100 and the radius parameter would be set to 500.

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

During this project I will use the following libraries:

* Pandas: For creating and manipulating data frames.
* 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.
* XML: To separate data from presentation and XML stores data in plain text format.
* Geocoder: To retrieve Location Data.
* Beautiful Soup and Requests: To scrap and library to handle http requests.
* Matplotlib: Python Plotting Module.