The Battle of Neighborhoods

Introduction:

The purpose of this Project is to help people in exploring better facilities around their neighborhood. It will help people making smart decision on selecting a great neighborhood in Scarborough, Toronto.

Lot of people migrate to various parts of Canada for jobs and other business reasons. They need to do a lot of research for finding good housing price and reputed school for their children. This project is for people who are looking for better neighborhoods with easy access to Cafe, School, Super market, medical shop & grocery shops, hospital, etc.

This Project aims to create an analysis of feature for a people migrating to Scarborough to search a best neighborhood as a comparative analysis between neighborhoods. The features include median housing price and better school according to ratings, crime rates of that particular area , weather conditions, good management for emergency and recreational facilities.

It will help people to get awareness of the area and neighborhood before moving to a new city, state, country or place for their work or to start a new fresh life.

Problems:

The main purpose of this project is to suggest a better neighborhood in a new city for the people who are planning to move. It will be easy for them to find the Connectivity to the airport, bus stand, city center, markets and other daily needs things nearby. In this project I have

1. Sorted list of house in terms of housing prices in a ascending or descending order
2. Sorted list of schools in terms of location, fees, rating and reviews

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:

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 used:

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.

XML: To separate data from presentation and XML stores data in plain text format.

Geocoder: To retrieve Location Data.

Scikit Learn: For importing k-means clustering.

JSON: Library to handle JSON files.

Matplotlib: Python Plotting Module.