Coursera Captone Project

THE BEST LOCATION FOR A HIGH-RISE MIXED-USE BUILDING

INTRODUCTION

Canadian migration has been on the rise since the 90s and popular areas such as Scarborough have become excellent hubs for both migrants and locals alike. In addition to the mass immigration across the regions, there has been an increased interest in mixed use buildings. These buildings have apartments, hotels, offices, public parking, medical facilities or retail included. These buildings are really attractive investments for developers as its approach is more modular rather than fixed such as apartment or retail only.

BACKGROUND

Sometimes it is just not enough to perform the groundwork by using qualitative research and surveys asking local people about their preferences in the targeted area as qualitative data entails much interpretation and is readily available for bias. In this case the aim of the project will be the utilization of foursquare data to ingest much of the local area as we can before proceeding to qualitative interviews such that we can target the respondents with the correct questions.

Problem Statement

How can we decide on which neighborhood to develop a new mixed use apartment based on data that is readily available?

Selection Criteria

The location should be both affordable, have fewer amenities than other areas and be positioned close to schools. Scarborough is there favorite amongst our investors however if there is another area with similar traits then a change in strategy is welcomed.

DATA FLOW

This project would use Foursquare 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.

Foursquare

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.

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
- ☑ **Geopy:** To retrieve Location Data
- ☑ **Requests:** Library to handle http requests
- ☑ **Matplotlib:** Python Plotting Module