Battle of the Neighbourhoods

NEEL RASTOGI

Introduction

Business Problem

The purpose of this project is to aim to inform people of the facilities in a given neighbourhood, helping them making decisions to determine if this neighbourhood is the right one for them as compared to others in Scarborough in Toronto.

Internal migrants of Canada would take into account specific factors such as whether housing is affordable, school ratings, access to facilities such as cafes, supermarkets, medical clinics, shopping centres etc. Other features which will be included in the analysis for people potentially migrating to Scarborough are crime rates, access by road, weather, emergency services, sanitation services and social or recreational venues. A comparative analysis with other neighbourhoods will also be provided to create awareness for those planning to move here.

The problems the project will aim to solve will be to sort housing prices in ascending order and sorting schools based on their locations, ratings and fee prices.

Data

Data Description

The Data used will be the same Scarborough dataset used previously, found on Wikipedia and consists of the latitude, longitude and postal codes of neighbourhoods in Scarborough. The data link is as follows: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

Using the Data

The Foursquare API will be used to help identify venues and/or facilities in the vicinity of the neighbourhood. All possible Foursquare location information will be requested and used to do so and will be the sole source of data and information in this project. Data collected from the API include names of the venue, its location, its menu (if applicable) and photos taken of the venue.

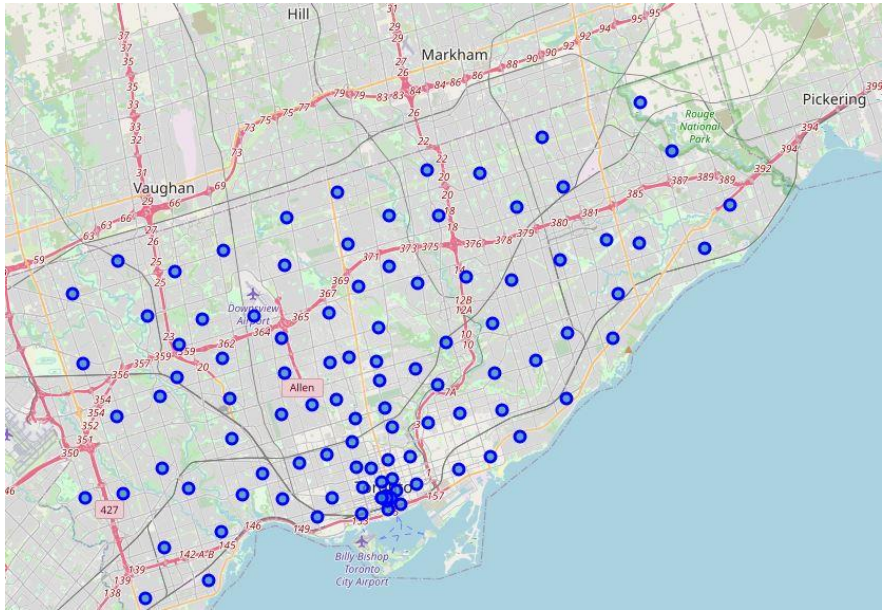
As per the data collected from Foursquare's location information, for neighbourhoods it will contain: Neighbourhood name Latitude of neighbourhood Longitude of neighbourhood

For the venue details, the information will contain: Name of the venue Latitude of the venue Longitude of the venue Type of venue

Data contd.

Location

Scarborough is a largely multicultural neighbourhood of Toronto and is home to diverse cultures and religions. Being a neighbourhood that is rising in popularity as well as rates of immigration, Scarborough is the perfect location to analyse.



Foursquare API

The Foursquare API will be used to gather resources and information in order to provide details about a business as well as performing location analysis. Neighbourhoods will be analysed and due to restrictions, neighbourhood parameters will be determined as well as a radius parameters.

Analysis Approach

Comparable analysis will be done using K-Means clustering in order to explore neighbourhoods and cluster them according to certain criteria.

Methodology

Clustering

In order to compare the two cities, neighbourhoods were clustered according to similar factors. To do so, an unsupervised machine learning methodology was used: k-means clustering

K-means Clustering:

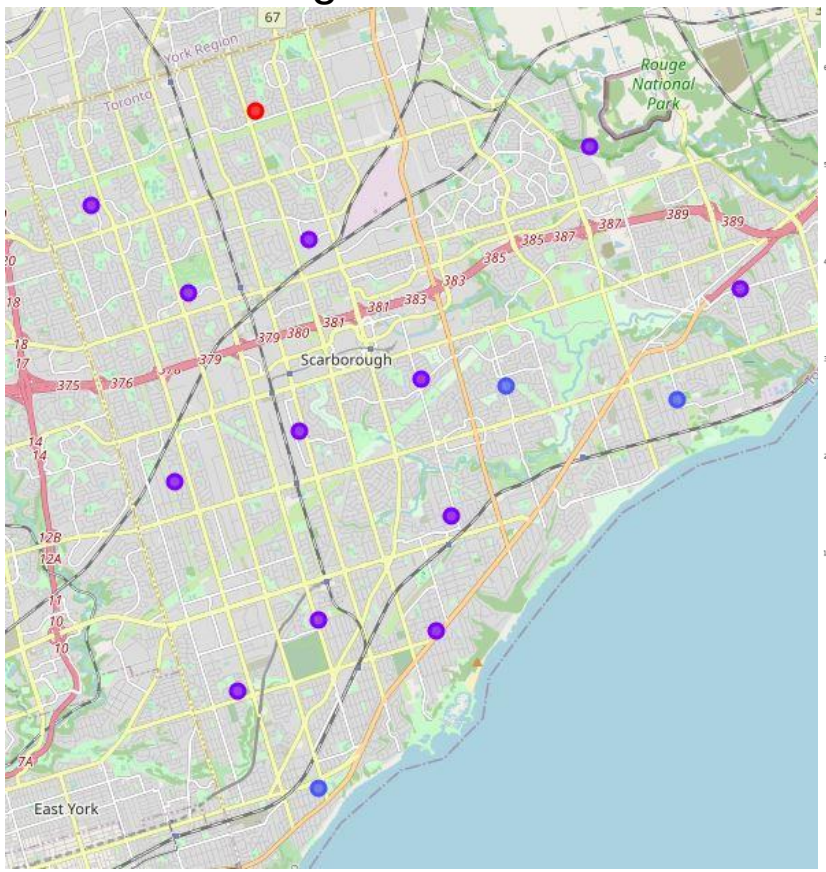
	Postalcode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	M1B	Scarborough	Malvern, Rouge	43.81139	-79.19662	1	Zoo Exhibit	Financial or Legal Service	Business Service	Fast Food Restaurant	Falafel Restaurant	Electronics Store	Elementary School	Escape Room	Ethiopian Restaurant	Event Space
1	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek	43.78574	-79.15875	1	Bar	Fish & Chips Shop	Yoga Studio	Farmers Market	Elementary School	Escape Room	Ethiopian Restaurant	Event Space	Falafel Restaurant	Farm
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.76575	-79.17470	2	Park	Gymnastics Gym	Gym / Fitness Center	Athletics & Sports	Donut Shop	Eastern European Restaurant	Electronics Store	Elementary School	Escape Room	Ethiopian Restaurant
3	M1G	Scarborough	Woburn	43.76812	-79.21761	2	Coffee Shop	Park	Chinese Restaurant	Fast Food Restaurant	Falafel Restaurant	Eastern European Restaurant	Electronics Store	Elementary School	Escape Room	Ethiopian Restaurant
4	M1H	Scarborough	Cedarbrae	43.76944	-79.23892	1	Bakery	Hakka Restaurant	Caribbean Restaurant	Gas Station	Athletics & Sports	Thai Restaurant	Bank	Fast Food Restaurant	Farmers Market	Farm

Most common venues:

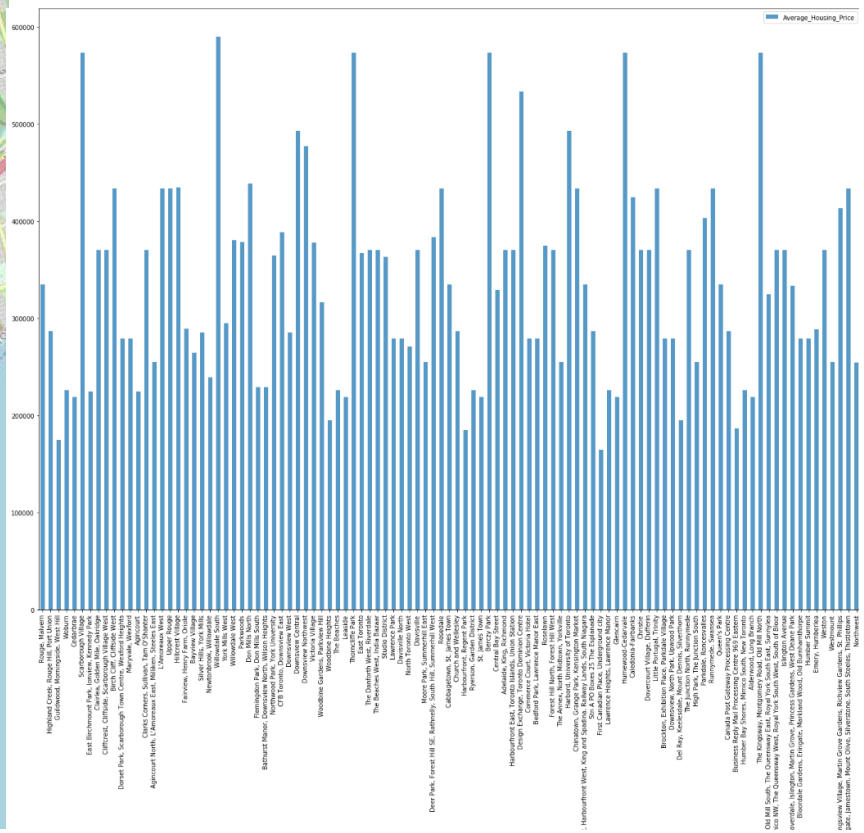
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Agincourt	Shopping Mall	Pool	Japanese Restaurant	Bank	Latin American Restaurant	Sushi Restaurant	Supermarket	Bubble Tea Shop	Skating Rink	Café
1	Alderwood, Long Branch	Gym	Gas Station	Print Shop	Pub	Pizza Place	Pharmacy	Coffee Shop	Sandwich Place	Dance Studio	Escape Room
2	Bathurst Manor, Wilson Heights, Downsview North	Pizza Place	Park	Coffee Shop	Mediterranean Restaurant	Restaurant	Men's Store	Mobile Phone Shop	Sushi Restaurant	Fried Chicken Joint	Sandwich Place
3	Bayview Village	Flower Shop	Dog Run	Park	Trail	Gas Station	Asian Restaurant	Yoga Studio	Eastern European Restaurant	Elementary School	Escape Room
4	Bedford Park, Lawrence Manor East	Sandwich Place	Coffee Shop	Italian Restaurant	Pizza Place	Indian Restaurant	Butcher	Café	Sports Club	Liquor Store	Sushi Restaurant

Results

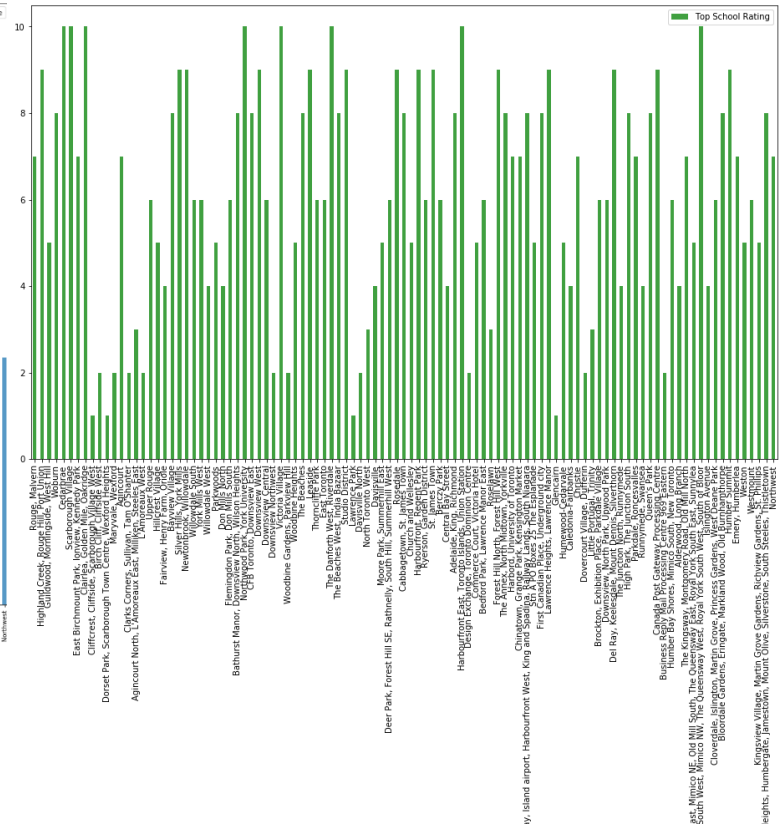
Scarborough Clusters:



Average housing prices:



Average school ratings:



Discussion

The purpose of this project was to provide insights into the different neighbourhoods of Scarborough that would appeal to those who are planning to migrate to this borough as well as providing important details as to what they can look forward to. The project aims to inform readers of social venues that would appeal to certain types of people, connectivity via roads, shopping venues and other venues that are nearby.

The histograms provide visualisations of the average price of housing for each neighbourhood in Scarborough as well as the ratings of schools in the neighbourhoods

Conclusion

In this project, k-means clustering was used in separate neighbourhoods to create 10 clusters from the dataset chosen. Using the above histograms, the results show average housing prices in the neighbourhoods and the school ratings accordingly.

This project has provided me the opportunity to showcase a practical application of a real world problem that can be answered using data science tools and methodology. Creating maps with the python package Folium is a very powerful tool which can analyse and visualise to aid in making better decisions with confidence, back by facts.

This project can be improved upon in different locations to provide the same or similar insights on a cost-effective basis.