The Battle of Neighborhood

A Capstone Project

Introduction

Background

Affordable housing has always been a problem in Toronto. Mr. X is very interested in buying a home in Etobicoke but is unsure about the neighbourhood. He has requested us to find and suggest the best neighbourhood suited as per his needs.

Business Problem

Mr. X is interested in a neighbourhood that meets the below criteria:

- price should be around $400000 to $500000
- nearby shopping centre
- nearby restaurants and eateries
- nearby park or green area

We need to find a neighbourhood in Etobicoke which fulfills the above conditions and make our suggestions.

Data Section

The data needed for our analysis will be collected from various sources.

1. **Toronto neighbourhood geo location and boundaries**
   This dataset will help us identify the latitude and longitude for all the neighbourhood. Further we will use this data with Foursquare APIs to find nearby information.
   Website url - https://open.toronto.ca/dataset/neighbourhoods/

2. **Toronto Housing data**
   The data came from various sources including Toronto Community Housing Corporation, City of Toronto's Shelter, Support and Housing Administration, City of Toronto Affordable Housing Office and Statistics Canada. Average Home Price data was taken from Realosophy.com.
   Website url - https://open.toronto.ca/dataset/wellbeing-toronto-housing/

3. **Toronto borough-neighbourhood data**
   This wiki page provides a list of neighbourhoods for each borough in Toronto City..
   Website url - https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Toronto

Methodology

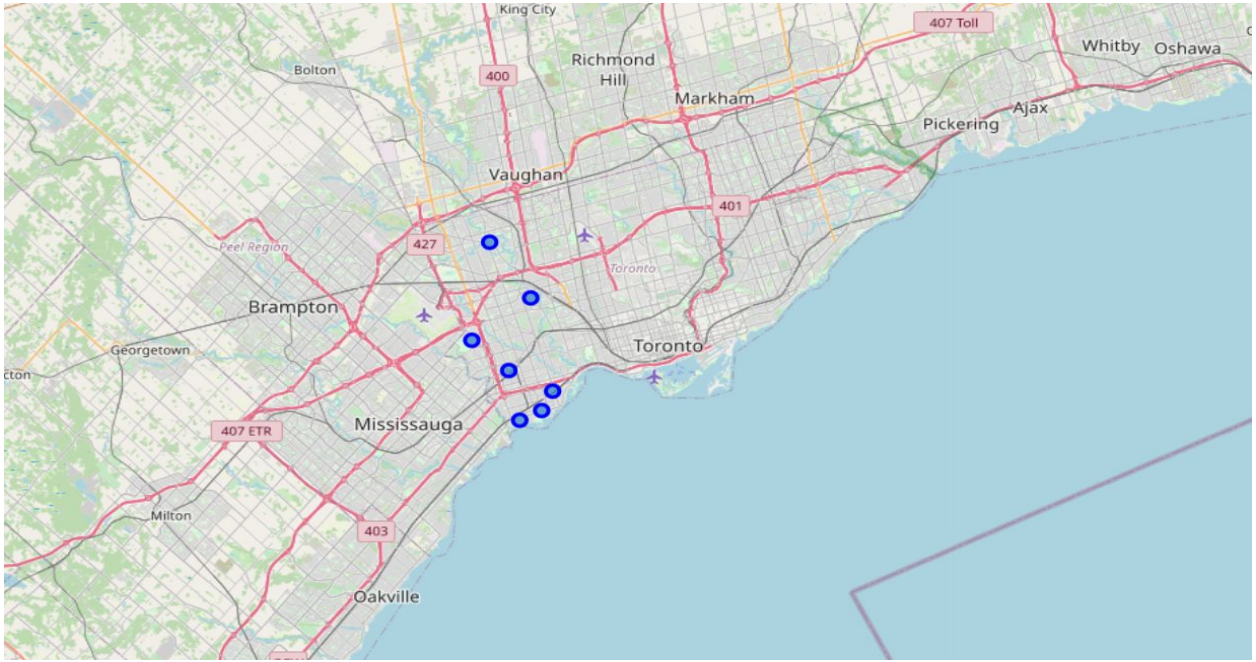
Exploratory data analysis - We analyse and select our desired neighborhoods for testing our model. Using the prepared dataset we have identified the below listed 5 neighborhoods which fulfill our pricing criteria.

- Eringate-Centennial-West Deane
- Humber Heights-Westmount
- Islington-City Centre West
- Long Branch
- Mimico

Neighborhood	Borough	Longitude	Latitude	Home Prices
Eringate-Centennial-West Deane	Etobicoke	-79.580445	43.658017	423034
Humber Heights-Westmount	Etobicoke	-79.522416	43.692233	491396
Islington-City Centre West	Etobicoke	-79.543317	43.633463	491678
Long Branch	Etobicoke	-79.533345	43.592362	459088
Mimico	Etobicoke	-79.500137	43.615924	429941

Methodology

Neighborhoods



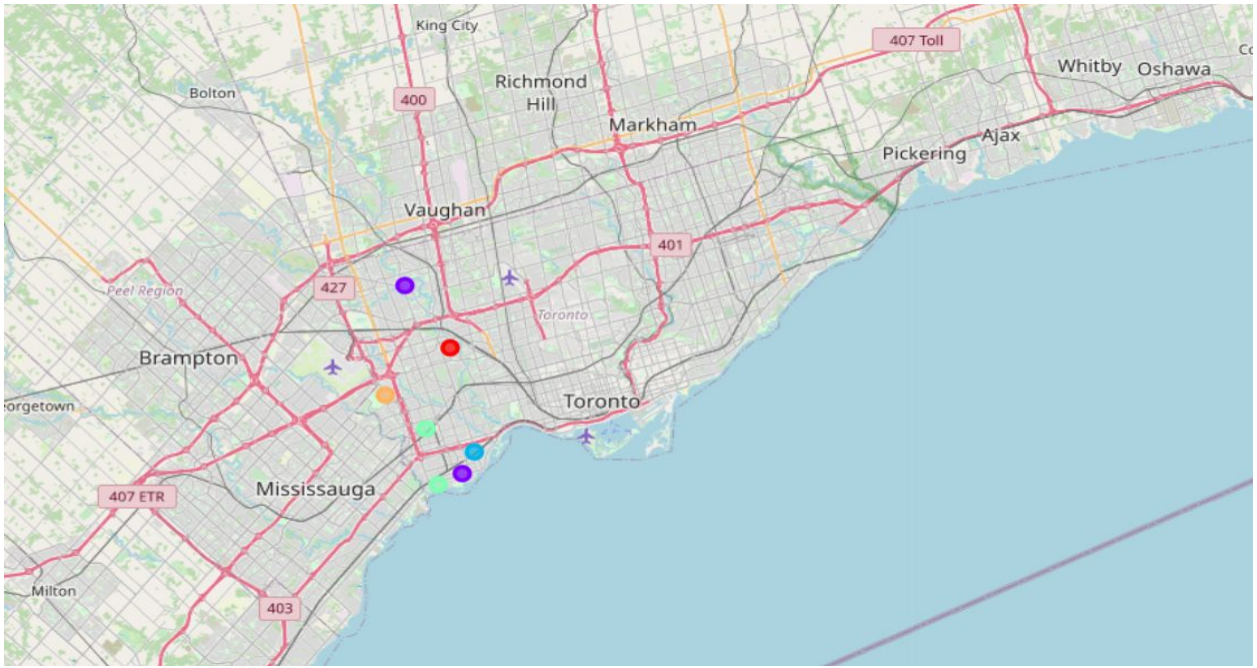
Methodology

Modelling - We will create a machine learning model using k-mean clustering mechanism and try to identify the most suitable neighbourhood amongst the selected ones. Clustering the neighborhood based on similar venues will let us compare and make better suggestions. We will be creating 5 clusters in our neighborhood.

Neighborhood	Borough	Longitude	Latitude	Home Prices	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	Com V
Eringate-Centennial-West Deane	Etobicoke	-79.580445	43.658017	423034	4	Chinese Restaurant	Hockey Arena	Convenience Store	Park	Coffee Shop	Pizza Place	Gas Station	Fried Chicken Joint	Fast Food Restaurant	Wor
Humber Heights-Westmount	Etobicoke	-79.522416	43.692233	491396	0	Gas Station	Park	Pizza Place	Greek Restaurant	Garden Center	Fried Chicken Joint	Fast Food Restaurant	Convenience Store	Coffee Shop	Wor
Islington-City Centre West	Etobicoke	-79.543317	43.633463	491678	3	Restaurant	Fast Food Restaurant	Women's Store	Garden Center	Greek Restaurant	Ice Cream Shop	Pizza Place	Bank	Fried Chicken Joint	Sanc
Long Branch	Etobicoke	-79.533345	43.592362	459088	3	Grocery Store	Bank	Coffee Shop	Restaurant	Greek Restaurant	Wings Joint	Italian Restaurant	Pharmacy	Pizza Place	
Mimico	Etobicoke	-79.500137	43.615924	429941	2	Grocery Store	Bakery	Bar	Skating Rink	Chinese Restaurant	Garden Center	Fried Chicken Joint	Fast Food Restaurant	Convenience Store	C

Methodology

Neighborhood clusters



Results and Discussions

The clusters are analysed to get a better understanding of the neighborhood.

Cluster 1

```
etobioke_merged.loc[etobioke_merged['Cluster Labels'] == 0, etobioke_merged.columns[[1] + list(range(5, etobioke_merged.shape[1]))]]
```

	Borough	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
49	Etobicoke	0	Gas Station	Park	Pizza Place	Greek Restaurant	Garden Center	Fried Chicken Joint	Fast Food Restaurant	Convenience Store	Coffee Shop	Women's Store

Cluster 2

```
etobioke_merged.loc[etobioke_merged['Cluster Labels'] == 1, etobioke_merged.columns[[1] + list(range(5, etobioke_merged.shape[1]))]]
```

	Borough	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
80	Etobicoke	1	Pub	Coffee Shop	Supermarket	Indian Restaurant	Italian Restaurant	Park	Caribbean Restaurant	Fast Food Restaurant	Convenience Store	Chinese Restaurant
110	Etobicoke	1	Indian Restaurant	Caribbean Restaurant	American Restaurant	Pharmacy	Coffee Shop	Ice Cream Shop	Pizza Place	Bank	Supermarket	Thai Restaurant

Results and Discussions

Cluster 3

```
etobioke_merged.loc[etobioke_merged['Cluster Labels'] == 2, etobioke_merged.columns[[1] + list(range(5, etobioke_merged.shape[1]))]]
```

	Borough	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
73	Etobicoke	2	Grocery Store	Bakery	Bar	Skating Rink	Chinese Restaurant	Garden Center	Fried Chicken Joint	Fast Food Restaurant	Convenience Store	Coffee Shop

Cluster 4

```
etobioke_merged.loc[etobioke_merged['Cluster Labels'] == 3, etobioke_merged.columns[[1] + list(range(5, etobioke_merged.shape[1]))]]
```

	Borough	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
54	Etobicoke	3	Restaurant	Fast Food Restaurant	Women's Store	Garden Center	Greek Restaurant	Ice Cream Shop	Pizza Place	Bank	Fried Chicken Joint	Sandwich Place
68	Etobicoke	3	Grocery Store	Bank	Coffee Shop	Restaurant	Greek Restaurant	Wings Joint	Italian Restaurant	Pharmacy	Pizza Place	Café

Cluster 5

```
etobioke_merged.loc[etobioke_merged['Cluster Labels'] == 4, etobioke_merged.columns[[1] + list(range(5, etobioke_merged.shape[1]))]]
```

	Borough	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
36	Etobicoke	4	Chinese Restaurant	Hockey Arena	Convenience Store	Park	Coffee Shop	Pizza Place	Gas Station	Fried Chicken Joint	Fast Food Restaurant	Women's Store

Results and Discussions

In this project

- We tried to identify the neighborhoods where houses were affordable enough.
- We created neighborhood clusters to compare them and get a better insight.
- We also used foursquare data to help us understand the surrounding of those neighborhood clusters.

In this process we realised we have cluster 1, 2 and 5 which look suitable for Mr. X preference, each having their own advantages.

Conclusion

- Cluster 1 and 5 satisfy all the criteria
- Cluster 3 and 4 failed to satisfy the criteria
- Cluster 2 has a variety of restaurants to choose from, has parks and stores. It also has an added advantage of having banks and pharmacies in the area.

Our suggestion would be to buy a home in cluster 2 neighborhood. Cluster 1 and 5 can also be considered if needed.