Capstone Project - Battle of Neighbourhoods: Neighbourhood recommendations to newcomers in Edmonton, Alberta, Canada

1 Introduction

Edmonton is Canada's fifth-largest municipality and the capital city of the Canadian province of Alberta, which has the world's third largest oil reserves and is one of the major oil & gas producers and exporters in the world. Edmonton is the primary supply and service hub for the province's crude oil industry and other northern resource industries. Because of its dynamic economy, Edmonton has been a very attractive immigration destination city for migrating workers and small business owners from other Canadian provinces and other countries [1, 2].

Fig. 1 City of Edmonton

Many newcomers come to Edmonton require information to choose a neighbourhood to live in, or to start a small business. The objective of this project is to provide a thorough analysis of the

residential neighbourhood property value, criminal occurrence and featured venues of neighbourhoods in Edmonton to help people to make better decisions on selecting the best neighbourhood in Edmonton according to the following criteria:

- ➢ Good housing value
- ➢ Low crime rate
- ➢ Convenience aka availability of venues

2 Data Sources and Collection

Datasets that is used for this project include Edmonton Property Assessment Data, Edmonton Neighbourhood Criminal Occurrence and Foursquare API data.

- ➢ **2019 Edmonton Property Assessment Data** was obtained from the City of Edmonton. In this dataset, the address, type, account number, assessed value, latitude and longitude of each property and the neighbourhood of each property belongs to were provided. The average residential housing value of each neighbourhood was calculated from the assessed value of each property in each neighbourhood. The location of each neighbourhood was identified with its latitude and longitude

- ➢ **Edmonton Neighbourhood Criminal Occurrence** was obtained from the Edmonton Police Service. The criminal occurrence in each neighbourhood was reported on monthly, quarterly and annual basis. From these data, the number of crimes each neighbourhood was calculated.

- ➢ **Foursquare API** will be used as a data source to gain information about near-by places per neighbourhood and to analyze venue presence in each neighbourhood within a radius of 100 meters. The detailed data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of a neighbourhood. The information obtained per venue are:
  - Neighbourhood
  - Neighbourhood Latitude
  - Neighbourhood Longitude
  - Name of the venue e.g. the name of a store or restaurant
  - Venue Latitude
  - Venue Longitude
  - Venue Category

3 Methodology

Due to the request limits of Foursquare API free developer account, only limited neighbourhoods that have good property value and low crime rates were selected for clustering analysis. Steps included data acquisition and cleansing, data preparation, feature selection, clustering.

3.1 Data acquisition and cleansing

Data acquisition and cleansing includes four sections/steps:

1. Obtain the property assessed values in Edmonton, drop non-residential property data and get the average assessed housing value in each neighbourhood;
2. Obtain the crime dataset in Edmonton, drop crime data in non-residential neighbourhood and get the total number of crimes in each neighbourhood from 2009 and 2019;
3. Select neighbourhoods that have great property value and lowest number of crimes for further analysis due to the request limitations of Foursquare API free developer account;
4. Obtain and analyze the venues within these neighbourhoods.

3.2 Clustering Approach

In order to compare the similarities of each neighbourhood or provide more information for newcomers to choose a neighbourhood to live in or start a business, neighbourhoods in Edmonton was explored, segmented and group into clusters. An unsupervised machine learning – k-means clustering algorithm was used.

4 Results

4.1 Neighbourhoods with Great Property Value and Low Crimes

There are about 340 residential neighbourhoods in Edmonton. Due to the request limit of Foursquare API free developer account, only 42 neighbourhoods with good property values and low number of crimes are selected for further analysis. These 42 neighbourhoods are also potential neighbourhoods that will be recommended to newcomers, as illustrated in Fig. 2.
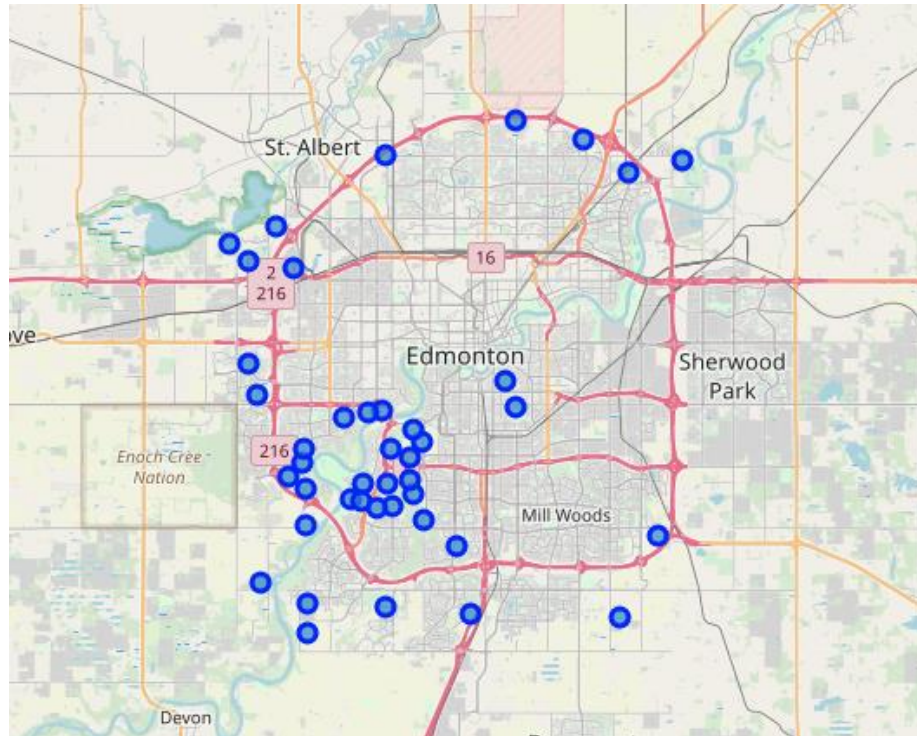
Fig. 2: Illustration of selected neighbourhoods in Edmonton

The following quadrant chart demonstrated the selected neighbourhoods with normalized housing value and total number of crimes. A full list with detailed neighbourhood information is available in Appendix A. As the Fig. 3 demonstrated, **Windermere Area** and **Rural West Big Lake** has the highest average property value (2.29 million Canadian Dollars and 1.9 million Canadian Dollars, respectively) and lowest number of crimes (5 and 58 occurrences respectively), because those two areas are newly developed neighbourhood. **Blackmud Creek Ravine, Anthony Henday Lake District, Anthony Henday Energy Park** and **River's Edge** have lowest number of crimes (4, 4, 13, and 11, respectively) and relatively cheap housing price (640, 620, 523, and 450 k Canadian Dollars, respectively).
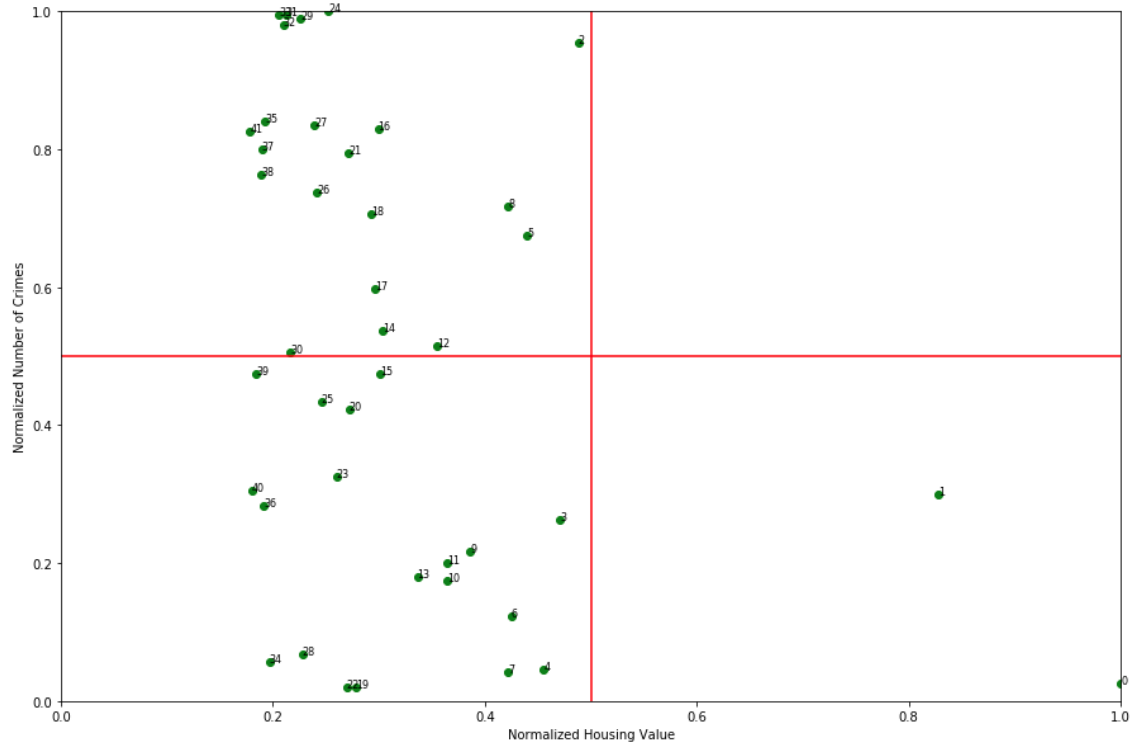
Fig. 3 Neighbourhoods' normalized housing value and total crimes

4.2 Clustering of Selected Neighbourhood

The venues information obtained from Foursqure API contains the information of 21 neighbourhoods, which were clustered, segmented and grouped by the categories of venues that are available in each neighbourhood. Five clustered neighbourhoods are presented. Cluster 1, cluster 4 and cluster 5 consist of less than two neighbourhoods. Cluster 2 consists of six neighbourhoods, featured with parks, sports fields and arts gallery. Cluster 3 consists of eleven neighbourhoods, featured with parks, stores and restaurants. The location of clustered neighbourhoods is demonstrated in the map of Fig. 4.

Cluster 1: discount store

	Neighbourhood	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
35	WEBBER GREENS	0.0	Discount Store	Electronics Store	Hockey Arena	Health & Beauty Service	Gym	Golf Course	Gas Station	Garden Center	Fast Food Restaurant	Event Service

Cluster 2: parks, sports area, arts gallery

	Neighbourhood	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
3	MILL CREEK RAVINE NORTH	1.0	Water Park	Home Service	Park	Trail	Butcher	Campground	Hockey Arena	Health & Beauty Service	Gym	Golf Course
15	HENDERSON ESTATES	1.0	Park	Performing Arts Venue	Water Park	Home Service	Health & Beauty Service	Gym	Golf Course	Gas Station	Garden Center	Fast Food Restaurant
21	BROOKSIDE	1.0	Soccer Field	Park	Water Park	Electronics Store	Health & Beauty Service	Gym	Golf Course	Gas Station	Garden Center	Fast Food Restaurant
25	WEDGEWOOD HEIGHTS	1.0	Park	Water Park	Indian Restaurant	Hockey Arena	Health & Beauty Service	Gym	Golf Course	Gas Station	Garden Center	Fast Food Restaurant
27	WESTRIDGE	1.0	Park	Event Service	Dance Studio	Water Park	Hockey Arena	Health & Beauty Service	Gym	Golf Course	Gas Station	Garden Center
29	BULYEA HEIGHTS	1.0	Playground	Pizza Place	Baby Store	Park	Health & Beauty Service	Gym	Golf Course	Gas Station	Garden Center	Fast Food Restaurant

Cluster 3: parks, stores, restaurants

	Neighbourhood	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
6	WHITEMUD CREEK RAVINE NORTH	2.0	Campground	Ski Chalet	Rest Area	Water Park	Electronics Store	Health & Beauty Service	Gym	Golf Course	Gas Station	Garden Center
13	MILL CREEK RAVINE SOUTH	2.0	Hockey Arena	Trail	Butcher	Liquor Store	Smoke Shop	Event Service	Health & Beauty Service	Gym	Golf Course	Gas Station
17	RURAL NORTH EAST SOUTH STURGEON	2.0	Golf Course	Garden Center	Water Park	Electronics Store	Hockey Arena	Health & Beauty Service	Gym	Gas Station	Fast Food Restaurant	Event Service
19	BLACKMUD CREEK RAVINE	2.0	Electronics Store	Movie Theater	Convenience Store	Event Service	Hockey Arena	Health & Beauty Service	Gym	Golf Course	Gas Station	Garden Center
20	DONSDALE	2.0	Italian Restaurant	Liquor Store	Coffee Shop	Pharmacy	Event Service	Health & Beauty Service	Gym	Golf Course	Gas Station	Garden Center
26	RIO TERRACE	2.0	Pub	Coffee Shop	Park	Convenience Store	Water Park	Event Service	Health & Beauty Service	Gym	Golf Course	Gas Station
31	POTTER GREENS	2.0	Steakhouse	Coffee Shop	Gas Station	Convenience Store	Water Park	Event Service	Hockey Arena	Health & Beauty Service	Gym	Golf Course
32	KINOKAMAU PLAINS AREA	2.0	Home Service	Electronics Store	Hockey Arena	Health & Beauty Service	Gym	Golf Course	Gas Station	Garden Center	Fast Food Restaurant	Event Service
33	BLUE QUILL ESTATES	2.0	Trail	Nature Preserve	Home Service	Hockey Arena	Health & Beauty Service	Gym	Golf Course	Gas Station	Garden Center	Fast Food Restaurant
38	FALCONER HEIGHTS	2.0	Asian Restaurant	Pizza Place	Coffee Shop	Fast Food Restaurant	Bakery	Gym	Indian Restaurant	Juice Bar	Liquor Store	Pet Store
40	HAWKS RIDGE	2.0	Pub	Water Park	Electronics Store	Hockey Arena	Health & Beauty Service	Gym	Golf Course	Gas Station	Garden Center	Fast Food Restaurant

Cluster 4: construction, landscaping, SPA

	Neighbourhood	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
12	GRANDVIEW HEIGHTS	3.0	Construction & Landscaping	Water Park	Electronics Store	Hockey Arena	Health & Beauty Service	Gym	Golf Course	Gas Station	Garden Center	Fast Food Restaurant
39	STARLING	3.0	Construction & Landscaping	Spa	Water Park	Electronics Store	Hockey Arena	Health & Beauty Service	Gym	Golf Course	Gas Station	Garden Center

Cluster 5: Health & Beauty Service

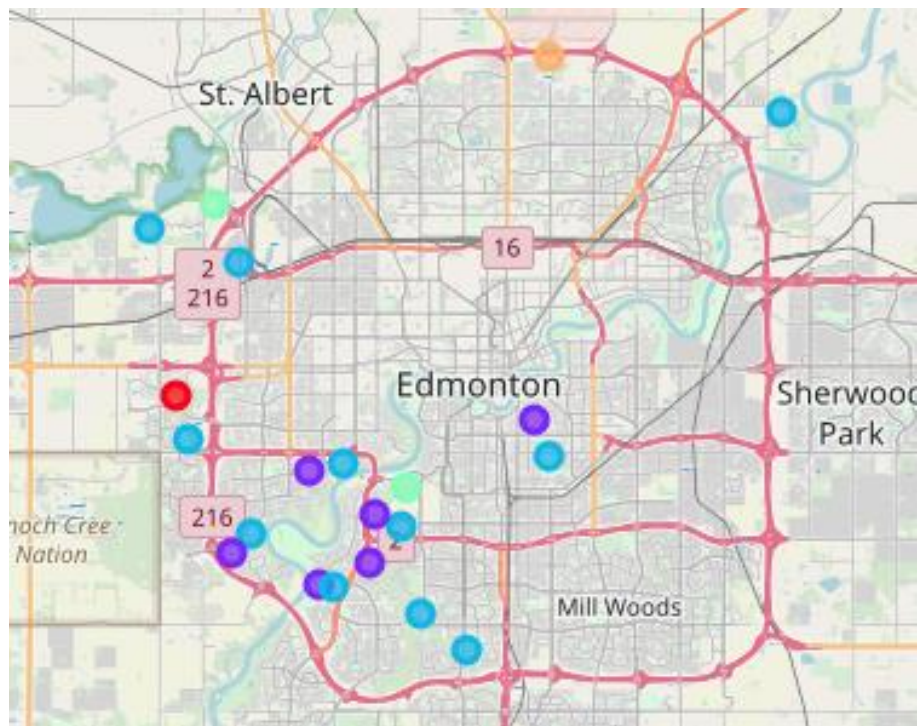|	Neighbourhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
22	ANTHONY HENDAY LAKE DISTRICT	4.0	Health & Beauty Service	Water Park	Electronics Store	Hockey Arena	Gym	Golf Course	Gas Station	Garden Center	Fast Food Restaurant	Event Service

Fig. 4 Locations of clustered neighbourhoods in Edmonton

5 Discussion

Through the analysis of this report, newcomers can gain information about housing value, criminal occurrence and the features of each neighbourhood's lifestyle and about what type of small business can be opened in each clustered neighbourhoods in Edmonton. For cluster 2 neighbourhoods, most

of the venues are parks, sports fields and art galleries. For cluster 3, most of the venues are parks, stores and restaurants. For some neighbourhoods, such as **Windermere Area,** venue information isn't available. This might because it's a newly developed neighbourhood. These areas are great choices for newcomers to start a small business.

6 Conclusion

In this project, through the analysis of the Edmonton Property Assessment Data and Neighbourhood Criminal Occurrence datasets, and a k-means cluster algorithm, the neighbourhoods in Edmonton were presented by the average housing value, numbers of crimes and features of venues. Newcomers can use this information to choose the neighbourhood that they like and want to start a life at.

Reference

[1] https://en.wikipedia.org/wiki/Edmonton

[2] https://en.wikipedia.org/wiki/Petroleum_industry_in_Canada

Appendix A:

Selected Neighbourhoods and their Average Housing Value and Total Number of Crimes

Index	Neighbourhood	Assessed Value (CAD)	Total Crimes
0	WINDERMERE AREA	2.298667e+06	5
1	RURAL WEST BIG LAKE	1.903969e+06	58
2	WESTBROOK ESTATES	1.122334e+06	185
3	MILL CREEK RAVINE NORTH	1.082500e+06	51
4	ASTER	1.047361e+06	9
5	ANTHONY HENDAY RAMPART	1.010370e+06	131
6	WHITEMUD CREEK RAVINE NORTH	9.767500e+05	24
7	WHITEMUD CREEK RAVINE SOUTH	9.693958e+05	8
8	UNIVERSITY OF ALBERTA FARM	9.693333e+05	139
9	QUESNELL HEIGHTS	8.881579e+05	42
10	HAYS RIDGE AREA	8.380394e+05	34
11	RIVERVIEW AREA	8.377312e+05	39
12	GRANDVIEW HEIGHTS	8.149694e+05	100
13	MILL CREEK RAVINE SOUTH	7.729167e+05	35

14	DECOTEAU	6.968318e+05	104
15	HENDERSON ESTATES	6.923065e+05	92
16	CASHMAN	6.881207e+05	161
17	RURAL NORTH EAST SOUTH STURGEON	6.804964e+05	116
18	CAMERON HEIGHTS	6.729217e+05	137
19	BLACKMUD CREEK RAVINE	6.400000e+05	4
20	DONSDALE	6.256041e+05	82
21	BROOKSIDE	6.229045e+05	154
22	ANTHONY HENDAY LAKE DISTRICT	6.194706e+05	4
23	OGILVIE RIDGE	5.972208e+05	63
24	RHATIGAN RIDGE	5.790437e+05	194
25	WEDGEWOOD HEIGHTS	5.646905e+05	84
26	RIO TERRACE	5.543425e+05	143
27	WESTRIDGE	5.494800e+05	162
28	ANTHONY HENDAY ENERGY PARK	5.231846e+05	13
29	BULYEA HEIGHTS	5.195562e+05	192
30	CARTER CREST	4.974992e+05	98
31	POTTER GREENS	4.876096e+05	193
32	KINOKAMAU PLAINS AREA	4.823500e+05	190
33	BLUE QUILL ESTATES	4.728476e+05	193
34	RIVER'S EDGE	4.523281e+05	11
35	WEBBER GREENS	4.432169e+05	163
36	KESWICK AREA	4.393652e+05	55
37	GORMAN	4.354038e+05	155
38	FALCONER HEIGHTS	4.344448e+05	148
39	STARLING	4.237402e+05	92
40	HAWKS RIDGE	4.138588e+05	59
41	DECHENE	4.100778e+05	160