

Studying the proximity of schools and other venues to Albertan Communities

G. Standish

2020/08/03

1. Introduction

40,000 new immigrants make Alberta Canada their home each year. These immigrants can be from other parts of Canada or from other countries. This study determined the venues and schools in close proximity to Albertan neighborhoods. These neighborhoods were grouped by clustering neighborhoods with similar venues and schools. Immigrants can use this information to find a neighborhood that meets their requirements.

2. Description of Data and Analysis Tools

This study required the following sources of data:

1. Alberta postal codes from https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_T

This wiki page has all the postal codes for Alberta and includes latitude & longitude coordinates. This table will be web-scraped, stored, prepared and then used to define neighborhood locations and distances to schools and venues. Distances will be calculated based on the geodesic distance between geomatic locations and not the driving distance.

2. Foursquare API found at <https://developer.foursquare.com/>

This API provides the venues and schools in close proximity to the neighborhoods. This requires code to get the venue information in proximity to the defined neighborhoods. Stated on their website, this API will “Enable location discovery, venue search, and more in your app or website. Access firmographic and rich, user-generated content such as photos, ratings, and reviews—via flat file or API.”

This study primarily used python notebooks and the following libraries: numpy, pandas, json, matplotlib, sklearn, bs4, geopy, folium.

3. Methodology

Scraping the wiki page

This first step was to scrape the Wikipedia page into a pandas dataframe for all of the neighborhoods in Alberta, including the latitude and longitude. This dataframe was then cleaned and processed into the header below. This dataframe of neighborhood becomes the primary list of neighborhoods when calling the foursquare API for venue information.

Figure 1: alberta_boroughs – first 5 neighborhoods and locations

	Postal Code	Borough	Neighborhood	Latitude	Longitude
0	T1A	Medicine Hat	Central Medicine Hat	50.03646	-110.679250
1	T2A	Calgary	Penbrooke Meadows, Marlborough	51.04968	-113.964320
2	T3A	Calgary	Dalhousie, Edgemont, Hamptons, Hidden Valley	51.12606	-114.143158
3	T4A	Airdrie	East Airdrie	51.27245	-113.986980
4	T5A	Edmonton	West Clareview, East Londonderry	53.58990	-113.441300

Shape of dataframe

```
alberta_boroughs.shape
```

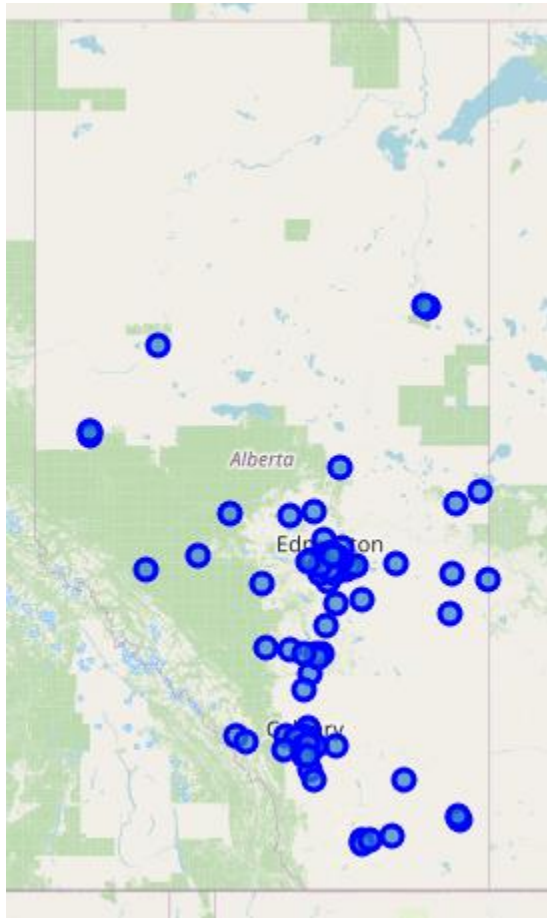
(138, 5)

Plotting Albertan neighborhoods on a map

Alberta is a province in Western Canada. Edmonton is located in central Alberta and is the provincial capital and second largest city in Alberta. Calgary is located in Southern Alberta and is the largest city in Alberta. Many neighborhoods are within Edmonton and Calgary but there are

others scattered throughout mostly central and southern Alberta. This is shown in the following figure.

Figure 2: Province of Alberta and Neighborhoods



■ Use Foursquare API to pull all venues and calculate venue frequency

The free version of foursquare developer was used to pull up to 100 venues within 1km of each neighborhood's designated location. The frequency of the venues categories within each neighborhood was calculated. Venue categories include but are not limited to restaurants, fast food, grocery stores, golf courses, schools, and many more. The neighborhood was defined by the top 10 most frequent venues. An example of the top 5 venues for 5 neighborhoods is provided in the figure below.

Figure 3: example of venues by neighborhood

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Athabasca	Gas Station	Automotive Shop	Sandwich Place	Burger Joint	Food & Drink Shop
1	Banff	Hotel	Steakhouse	Restaurant	Italian Restaurant	Pub
2	Barrhead	Grocery Store	Bank	Fast Food Restaurant	Financial or Legal Service	Falafel Restaurant
3	Beaumont	Coffee Shop	Fast Food Restaurant	Bank	Ice Cream Shop	Pizza Place
4	Bonnyville	Coffee Shop	Ice Cream Shop	Sandwich Place	Grocery Store	Gym / Fitness Center

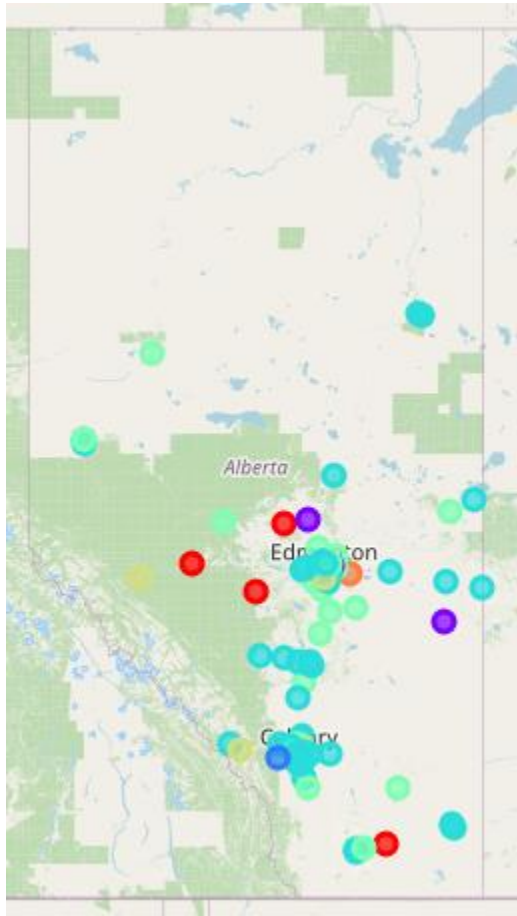
Clustering using kmeans machine learning

K means clustering was used to get a better understanding of the neighborhoods and the different venues available within them. 7 clusters were selected after observing the most meaningful categories formed. The clustering algorithm was run on the frequency of venues within neighborhoods. Neighborhoods with similar venues were grouped together.

These 7 clusters are defined as:

1. Red – Neighborhoods with high frequency of fast food restaurants.
2. Purple – Neighborhoods with high frequency of specifically pizza places.
3. Blue – Neighborhoods with high frequency of golf courses.
4. Teal – Neighborhoods with high frequency of coffee shops, fast food, restaurants and other venues generally associated with larger city down town
5. Light Green - Neighborhoods with high frequency of convenience stores, fast food, restaurants and other venues but generally more rural
6. Brown/Green - Neighborhoods with high frequency of Hotels and guest services
7. Orange - Neighborhoods with high frequency of Construction & Landscaping

Figure 4: Alberta neighborhoods grouped by venue frequency



School specific information

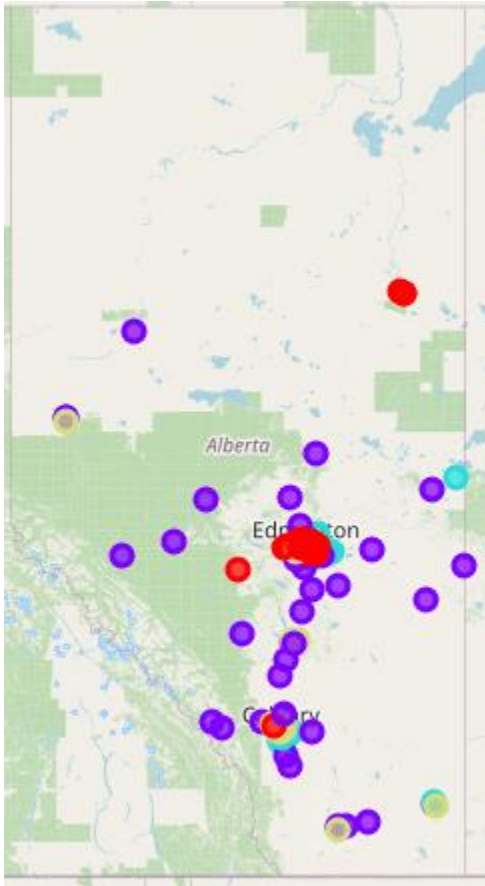
An important part of a young family's decision when deciding where to live is the proximity of schools within the neighborhood. This section looks further in the average distance between elementary, middle and high schools to the neighborhood's location.

Again, the Foursquare API was used to specifically collect school locations within the defined neighborhoods. The distance for each school was calculated from the neighborhood location. The neighborhoods were then compared and clustered into 4 groups.

1. Red – Neighborhoods with Elementary and Middle Schools in close proximity (<1km) but High schools 4-6km away.
2. Purple - Neighborhoods with Elementary, Middle and High schools within close proximity to the neighborhood. Usually within 1km
3. Blue – Neighborhoods with Elementary and Middle Schools further from the neighborhoods. Around 2-6km.

4. Green – Neighborhoods with Elementary, Middle and High schools within medium proximity to the neighborhood. Usually within 3km.

Figure 5: Neighborhoods grouped by proximity to schools



4. Results and Discussion

There are 130 neighborhoods in Alberta. Many of these neighborhoods are located in the large cities of Edmonton and Calgary, with other neighborhoods mostly scattered around southern Alberta.

Based on the venues described in section 3.4, seven groups of neighborhoods can be made. However, two of these groups represent the majority of the neighborhoods. The difference between them being the density of these venues between rural and urban. In these two groups many of the most frequent venues are observed in both groups. The other most notable neighborhoods are likely to be tourist locations or near airports with higher frequency of hotels.

The remaining four groups may appeal to golfers, pizza enthusiasts, convenience store regulars, or construction & landscape employees.

Generally, in Alberta, neighborhoods have similar access to venues.

When it comes to schools, Edmonton has many Elementary and Middle schools close to neighborhoods but high schools are further away. North and West Calgary neighborhoods are medium distance from schools, usually within 3km. Many of the neighborhoods in smaller cities around Alberta are less than 1km from schools.

For families looking for inner city venues and close proximity to schools, neighborhoods in some of the medium sized towns like Red Deer and Medicine Hat, provide this.

Further analysis may be able to determine the size of the schools. Edmonton may have fewer and larger schools than Calgary.

5. Conclusion

In this study, the frequency of venues and proximity to schools within neighborhoods was analyzed using k-means, a machine learning clustering algorithm. People trying to determine which neighborhood they would like to live in can use this information to narrow down their options.

Additionally, Albertans can take solace in knowing that many of the neighborhoods are similar in venues. New Albertans looking for the venue density offered by Calgary and Edmonton but smaller town and closer proximity of schools, should consider medium sized towns like Red Deer and Medicine Hat.