

# Studying the proximity of schools and other venues in Albertan Communities

---

# There are many neighborhoods to choose from in Alberta

---

- 40,000 new immigrants make Alberta Canada their home each year
  - from other parts of Canada or from other countries
- This study
  - determined the venues and schools in close proximity to Albertan neighborhoods
  - grouped neighborhoods by clustering with similar venues and schools
- Find a neighborhood that meets their requirements using this study

# Description of Data and Analysis Tools

---

Alberta postal codes from [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada: T](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_T)

- Wiki page with 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 locations, not the driving distance

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

- API provides the venues and schools in close 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.”

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

# Methodology

## 1. Scrape the wiki page

---

This first step was to scrape the Wikipedia page into a pandas dataframe, including the latitude and longitude, then cleaned and processed into the header below.

	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**

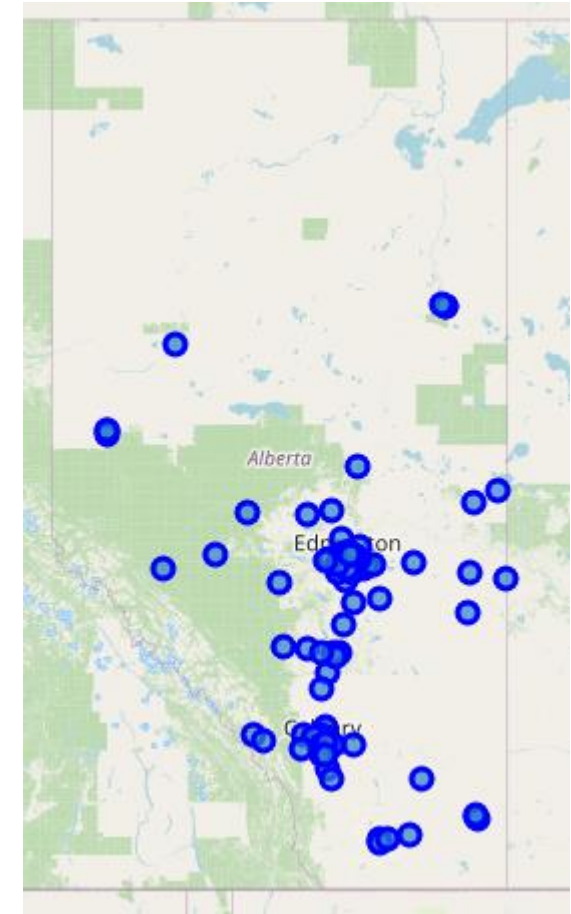
```
alberta_boroughs.shape
```

(138, 5)

# Methodology

## 2. 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.



# Methodology

## 3. Foursquare for venues and 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, etc.
- 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

	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

# Methodology

## 4. 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.

# Methodology

## 4. Seven clusters

---

**Red** – Neighborhoods with high frequency of fast food restaurants.

**Purple** – Neighborhoods with high frequency of specifically pizza places.

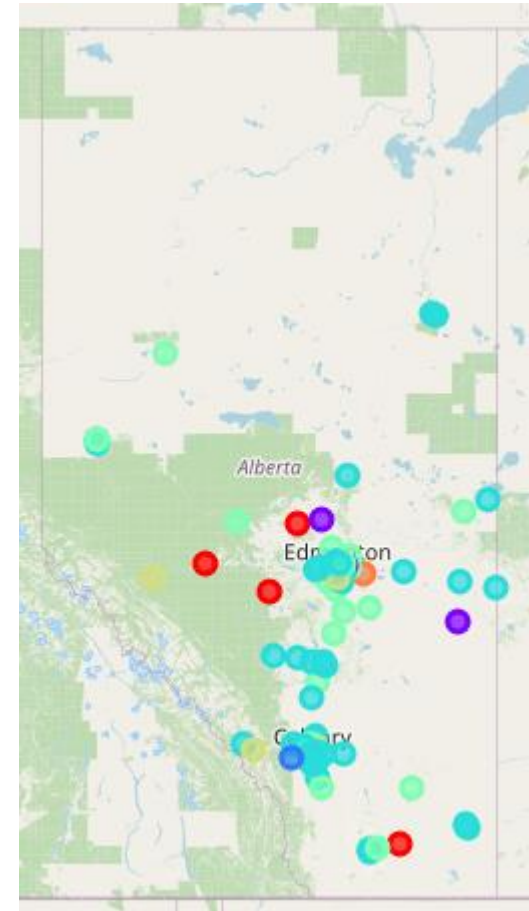
**Blue** – Neighborhoods with high frequency of golf courses.

**Teal** – Neighborhoods with high frequency of coffee shops, fast food, restaurants and other venues generally associated with larger city down town

**Light Green** - Neighborhoods with high frequency of convenience stores, fast food, restaurants and other venues but generally more rural

**Brown/Green** - Neighborhoods with high frequency of Hotels and guest services

**Orange** - Neighborhoods with high frequency of Construction & Landscaping





# Methodology

## 5. 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
- 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

# Methodology

## 5. School specific information

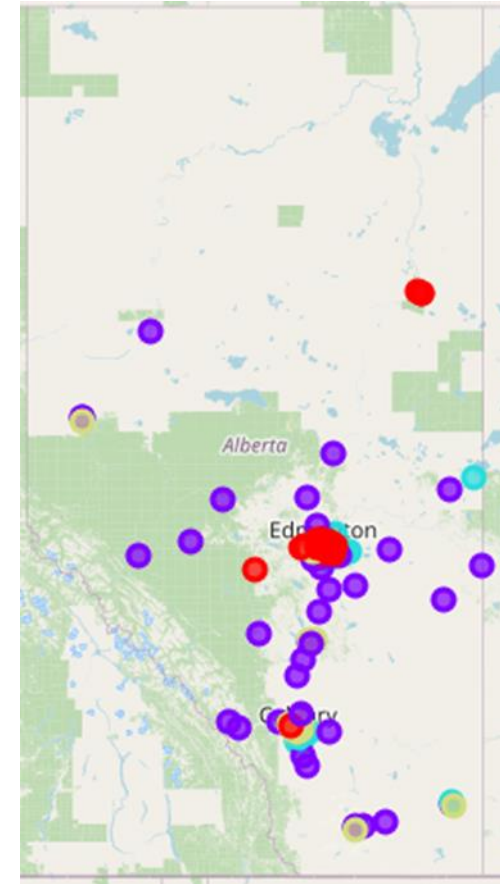
---

**Red** – Neighborhoods with Elementary and Middle Schools in close proximity (<1km) but High schools 4-6km away.

**Purple** - Neighborhoods with Elementary, Middle and High schools within close proximity to the neighborhood. Usually within 1km

**Blue** – Neighborhoods with Elementary and Middle Schools further from the neighborhoods. Around 2-6km.

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



# 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, seven groups of neighborhoods can be made
- Two of these groups represent the majority of the neighborhoods
  - the difference between them being the density of these venues between rural and urban
  - most frequent venues are observed in both groups
- The other most notable neighborhoods are 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

# Results and Discussion Continued

---

- 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 usually within 3km of schools
- 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
  - determine the size of the schools.
  - Edmonton may have fewer and larger schools than Calgary
  - Take into account driving distance

# 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 comfort 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, could consider medium sized towns like Red Deer and Medicine Hat.