

Exploring the City of Calgary

Chukwuka Nwakanma

June 15, 2020

1. Introduction

Calgary is the largest city in Alberta, Canada with a land mass of about 826km² and a population of approximately 1.3 million people as of 2017. It is still a growing city and the capital of oil and gas in Canada. As a growing city with a diverse population, people tend to move/immigrate to the city in doves.

Newcomers don't have a playbook when they come and as such it takes them a while to blend in. I move to the city in 2018 and I know the difficulties I had finding fun things, places to go eat. It was also difficult knowing the best neighborhood to live in that will have everything I needed without making long trips across the city. This makes it important to explore the different neighborhood in the city of Calgary.

The data can act as a playbook for newcomers to the city. Helping them in deciding the best neighborhood to live, where they can have access to everything they will need. The aim of this project is to explore, segment and cluster different neighborhoods in the city and in turn will help people make better choices.

I believe a lot of newcomers will be interested in having this playbook and I also think some residence too will want something that will help improve their choices in case they plan on moving to a different neighborhood in future.

2. Data Description and Data Cleaning

Unlike larger cities like Toronto and New York that have available dataset on the web, Calgary doesn't have that much dataset that can be used to achieve the task. But I was able to find something on Wikipedia that shows the list of neighborhoods in Calgary. To be able to extract the table that was used as my data, I did some web scrapping using the bs4 tool in the python library.

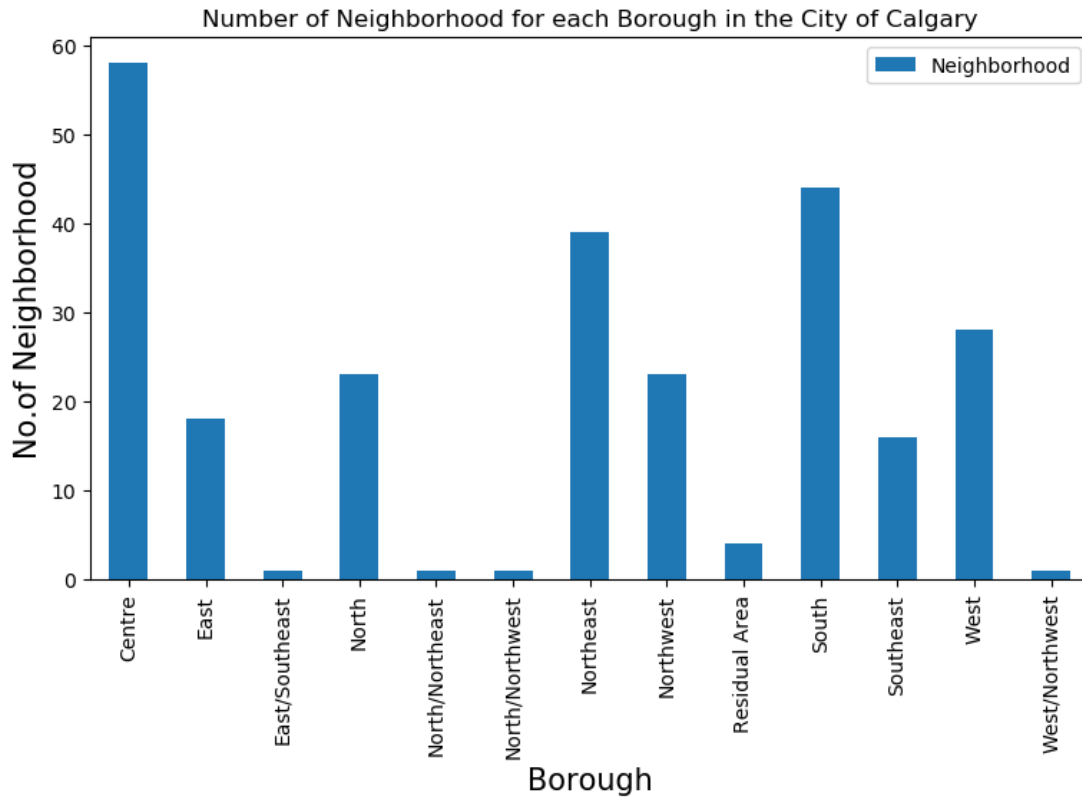
The table contained a lot of columns I didn't need. Since I just needed 2 important columns. I did some data cleaning by dropping the columns that wasn't necessary. I also renamed the important columns so I could read my data properly. My dataset consisted of 2 columns Neighborhood and Borough.

I used the python geocoder library to get the longitude and latitude of the neighborhoods.

I also used the tool folium to get the map of the city of Calgary. To get the most common venues I used a third-party tool called Foursquare API.

3. Methodology

There are a total of 257 neighborhoods in Calgary, Alberta. Represented in the chart below is the number of neighborhoods for each Borough.

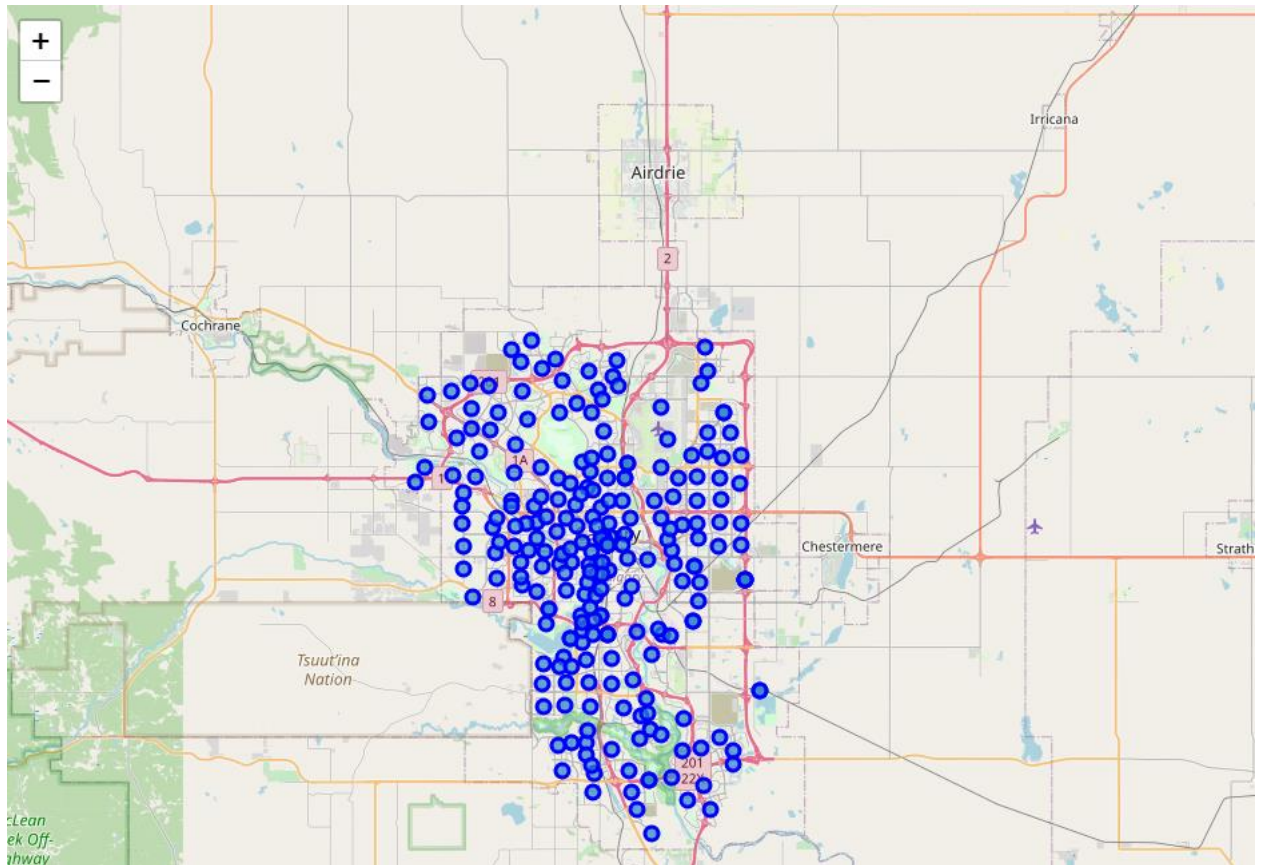


Centre borough has the highest number of neighborhoods with 58.

As noted, Geocoder library was used to get the longitude and latitude of the neighborhoods. I merged both the longitude and latitude information into my data frame.

	Neighborhood	Borough	Latitude	Longitude
0	Abbeydale	Northeast	51.05976	-113.92546
1	Acadia	South	50.97227	-114.05843
2	Albert Park/Radisson Heights	East	51.04200	-113.99683
3	Altadore	Centre	51.01601	-114.10558
4	Alyth/Bonnybrook	Centre	51.01155	-114.04481

After getting the longitude and latitude of the neighborhoods. I created a map of the city of Calgary using the visualization tool folium library as shown below:



I used the Foursquare API to explore the neighborhoods and segment. I realized that the city of Calgary has a total of 2416 venues with categories ranging from Health & Beauty Services, Gym/Fitness Centers, Recreation Centers etc. There are 225 unique categories.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Abbeydale	51.05976	-113.92546	Vanity Fitness	51.062982	-113.926238	Health & Beauty Service
1	Acadia	50.97227	-114.05843	Acadia Aquatic & Fitness Centre	50.973297	-114.059984	Gym / Fitness Center
2	Acadia	50.97227	-114.05843	Acadia Recreational Center	50.972256	-114.065077	Recreation Center
3	Albert Park/Radisson Heights	51.04200	-113.99683	Sugar Tree Cleaning Company	51.044662	-113.997207	Home Service
4	Altadore	51.01601	-114.10558	Garrison Pub	51.016131	-114.108674	Pub

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
2411	Woodbine	50.94076	-114.12887	Wildman BBQ & Juke Joint	50.940522	-114.135654	BBQ Joint
2412	Woodlands	50.94207	-114.10683	3 Crowns	50.940765	-114.109430	Pub
2413	Woodlands	50.94207	-114.10683	Russian Store	50.941063	-114.109452	Food & Drink Shop
2414	Woodlands	50.94207	-114.10683	Mac's	50.940876	-114.109816	Food & Drink Shop
2415	Woodlands	50.94207	-114.10683	A Buyer's Choice Home Inspections	50.941854	-114.100715	Shop & Service

I sorted the top 5 venue category for each neighborhood. Below is a data frame showing the top 5 venues for each neighborhood

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Abbeydale	Health & Beauty Service	Women's Store	Eastern European Restaurant	Food Truck	Food Court
1	Acadia	Recreation Center	Gym / Fitness Center	Women's Store	Donut Shop	Food Court
2	Albert Park/Radisson Heights	Home Service	Women's Store	Eastern European Restaurant	Food Truck	Food Court
3	Altdore	Pub	Massage Studio	Ice Cream Shop	Park	Women's Store
4	Alyth/Bonnybrook	Construction & Landscaping	Restaurant	Motorcycle Shop	Miscellaneous Shop	Deli / Bodega

I used the unsupervised machine learning algorithm k-clustering to cluster the neighborhood into 5 clusters, I choose k-cluster algorithm is because it is efficient in categorizing data into groups.

Then I created a new data frame which included the cluster and the top 5 neighborhoods. Shown below:

	Neighborhood	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Abbeydale	Northeast	51.05976	-113.92546	2.0	Health & Beauty Service	Women's Store	Eastern European Restaurant	Food Truck	Food Court
1	Acadia	South	50.97227	-114.05843	2.0	Recreation Center	Gym / Fitness Center	Women's Store	Donut Shop	Food Court
2	Albert Park/Radisson Heights	East	51.04200	-113.99683	4.0	Home Service	Women's Store	Eastern European Restaurant	Food Truck	Food Court
3	Altdore	Centre	51.01601	-114.10558	0.0	Pub	Massage Studio	Ice Cream Shop	Park	Women's Store
4	Alyth/Bonnybrook	Centre	51.01155	-114.04481	2.0	Construction & Landscaping	Restaurant	Motorcycle Shop	Miscellaneous Shop	Deli / Bodega

4. Result

We can examine each cluster to see which neighborhoods were assigned to the 5 clusters created with K-means clustering.

Analyzing each cluster to identify the characteristics of each cluster and the number of neighborhoods in the cluster

Cluster 1

199	Scenic Acres	Park	Women's Store	Eastern European Restaurant	Food Truck	Food Court
206	Sherwood	Park	Women's Store	Eastern European Restaurant	Food Truck	Food Court
208	Silver Springs	Home Service	Baseball Field	Park	Women's Store	Electronics Store
231	Sunnyside	Park	Scenic Lookout	Trail	Café	Skating Rink
240	Upper Mount Royal	Park	River	Athletics & Sports	Women's Store	Eastern European Restaurant
249	Westwinds	Lighting Store	Park	Women's Store	Electronics Store	Food Truck
251	Wildwood	Home Service	Trail	Park	Dog Run	Women's Store

```
[115]: calgary_merged[calgary_merged['Cluster Labels'] == 0].shape
```

```
[115]: (26, 10)
```

Cluster one has 26 neighborhoods and the most common venues consist of Parks, Eastern European Restaurant, Athletics & Sports, Dog run and Food Court.

Cluster 2

109	Killarney/Glengarry	Convenience Store	Hobby Shop	Pharmacy	School	Donut Shop
136	Monterey Park	Convenience Store	Construction & Landscaping	Adult Boutique	Health & Beauty Service	Eastern European Restaurant
148	Ogden Shops	Convenience Store	Pizza Place	Clothing Store	Bar	Electronics Store
154	Patterson	Gas Station	Health & Beauty Service	Pizza Place	Convenience Store	Vietnamese Restaurant
163	Ranchlands	Convenience Store	Pizza Place	Park	Sandwich Place	Donut Shop
181	Richmond	Pizza Place	Playground	Coffee Shop	Women's Store	Eastern European Restaurant
212	Skyview Ranch	Pizza Place	Gas Station	Film Studio	Women's Store	Eastern European Restaurant
213	Somerset	Convenience Store	Flower Shop	Bus Station	Plaza	Hockey Arena
222	Starfield	Convenience Store	Bus Station	Cosmetics Shop	French Restaurant	Food Truck
241	Valley Ridge	Convenience Store	Greek Restaurant	Restaurant	Golf Course	Food Court
243	Varsity	Pizza Place	Yoga Studio	Thrift / Vintage Store	Chinese Restaurant	Café
250	Whitehorn	Construction & Landscaping	Pizza Place	Pet Store	Women's Store	Eastern European Restaurant

```
[117]: calgary_merged[calgary_merged['Cluster Labels'] == 1].shape
```

```
[117]: (26, 10)
```

Cluster 2 also has 26 neighborhoods and the most common venues consist of Convenience Store, Pizza Place, Bus Stations, Thrift/Vintage Store and Pharmacy.

Cluster 3

237	Tuxedo Park	Italian Restaurant	Chinese Restaurant	Grocery Store	Pizza Place	Dim Sum Restaurant
238	University Heights	Sandwich Place	Women's Store	Eastern European Restaurant	Food Truck	Food Court
239	University of Calgary	Coffee Shop	Concert Hall	Skating Rink	Sandwich Place	Hotel
242	Valleyfield	Hotel	Restaurant	Coffee Shop	Pub	Steakhouse
244	Vista Heights	Vietnamese Restaurant	Café	Brewery	Women's Store	Food Truck
245	Walden	Bank	Coffee Shop	Mexican Restaurant	Sandwich Place	Salon / Barbershop
246	West Hillhurst	Convenience Store	Furniture / Home Store	Liquor Store	Italian Restaurant	Dessert Shop
247	West Springs	Coffee Shop	Pizza Place	Restaurant	Tapas Restaurant	Grocery Store
248	Westgate	Gym	Fried Chicken Joint	Food Truck	Food Court	Food & Drink Shop
252	Willow Park	Pool	Grocery Store	Jewelry Store	Food Truck	Food Court
253	Windsor Park	Restaurant	Women's Store	Eastern European Restaurant	Food Truck	Food Court
255	Woodbine	Home Service	BBQ Joint	Music Venue	Spa	Playground
256	Woodlands	Food & Drink Shop	Pub	Shop & Service	Women's Store	Eastern European Restaurant

```
[119]: calgary_merged[calgary_merged['Cluster Labels'] == 2].shape
```

```
[119]: (165, 10)
```

Cluster 3 has 165 neighborhoods and the most common venues consist of Coffee Shops, Liquor Stores, Vietnamese Restaurant and Hotel.

Cluster 4

Cluster 4

```
[120]: calgary_merged.loc[calgary_merged['Cluster Labels'] == 3, calgary_merged.columns[[0] + list(range(5, calgary_merged.shape[1]))]]
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
53	Deer Ridge	Dog Run	Women's Store	Electronics Store	French Restaurant	Food Truck
54	Deer Run	Dog Run	Women's Store	Electronics Store	French Restaurant	Food Truck
254	Winston Heights/Mountview	Golf Course	Dog Run	Women's Store	Electronics Store	Food Truck

```
[121]: calgary_merged[calgary_merged['Cluster Labels'] == 3].shape
```

```
[121]: (3, 10)
```

Cluster 4 has 3 neighborhoods and the most common venues consist of Dog Run, Electronics Stores. Women's stores and Food Trucks

Cluster 5

```
[122]:
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
2	Albert Park/Radisson Heights	Home Service	Women's Store	Eastern European Restaurant	Food Truck	Food Court
38	Citadel	Home Service	Big Box Store	Women's Store	Electronics Store	French Restaurant
39	Cityscape[2]	Home Service	Ice Cream Shop	Business Service	Women's Store	Electronics Store
127	Mayland Heights	Home Service	Light Rail Station	Thrift / Vintage Store	Women's Store	Eastern European Restaurant
149	Palliser	Home Service	Playground	Women's Store	Donut Shop	Food Court
150	Panorama Hills	Home Service	Pizza Place	Women's Store	Donut Shop	Food Court
165	Redstone	Home Service	Women's Store	Eastern European Restaurant	Food Truck	Food Court
195	Sage Hill	Asian Restaurant	Home Service	Eastern European Restaurant	French Restaurant	Food Truck
209	Silverado	Hotel	Home Service	Bus Station	Electronics Store	Food Truck
234	Temple	Construction & Landscaping	Home Service	Eastern European Restaurant	Food Truck	Food Court

```
[123]: calgary_merged[calgary_merged['Cluster Labels'] == 4].shape
```

```
[123]: (10, 10)
```

Cluster 5 has 10 neighborhoods and the most common venues consist of Light Rail Station, Asian Restaurant, Donut Shops and Home Service.

5. Discussion

The aim of this project was to act as a playbook for newcomers and visitors who want to explore the city of Calgary and get the most out of what the city can offer.

From examining the clusters, I realized that neighborhoods in cluster 3 would be the best playbook because it has almost every common venue. From Pubs to hotels, gym/fitness centers, malls, spas, golf courses, coffee shops, convenience store, restaurants and liquor store. So, I would highly recommend neighborhoods in cluster 3 to anyone planning to visit or move to the city of Calgary.

6. Conclusion

As early stated, this project helps newcomers and visitors to understand the neighborhoods in the city of Calgary. It will help people to make better choices when deciding to explore the city and get the best out of the beautiful city. The future of this project is to add crime data so people will also know the safest neighborhood in Calgary.