

The Battle of the Neighborhoods - Week 1

Table of contents

- [Introduction](#)
- [Business Problem](#)
- [Data](#)

Introduction

The city of Calgary is one of the largest municipalities Canada located in the province of Alberta. The city had a population of 1,285,711 in 2019, making it Alberta's largest city and Canada's third-largest municipality.

Calgary's economy includes activity in the energy, financial services, film and television, transportation and logistics, technology, manufacturing, aerospace, health and wellness, retail, and tourism sectors.

The Calgary Metropolitan Area (CMA) is home to Canada's second-highest number of corporate head offices among the country's 800 largest corporations.

With a thriving population and rich economy like this, there is no doubt a restaurant might be a good business venture in the city of Calgary.

However, an investor needs to be confident that they are making the right considerations before setting up a restaurant business in the city Calgary.

Business Problem

With the purpose in mind, finding the right location to open a thriving restaurant is one of the crucial factors to guarantee success in this business venture.

In this Capstone project, I will be focussing on different types of restaurants which are opened or closed in a particular location and then decide if it is good place to open a new restaurant based on the popular cuisine around that place. By using location analytics and machine learning algorithms such as clustering, this project aims to provide solutions to answer these business questions.

Source of Data

For this analysis, I will be using the "List of neighbourhoods in Calgary" data scraped from Wikipedia (https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Calgary).

From the scraped data, there are total 257 neighbourhoods are in Calgary. The data will be trimmed down to two features ("Name" and "Sector") to remove irrelevant data for this analysis.

There are total 257 neighbourhoods in Calgary. We will find latitude and longitude of each neighbourhood and cluster them according to the restaurants present in each neighbourhood fetched from foursquare location data. Then we will make a decision examining each cluster of neighbourhoods.

Sample records from Wikipedia.

	Name[9]	Quadrant	Sector[10]	Ward[11]	Type[10]	2012 PopulationRank	Population(2012) [9]	Population(2011) [9]	% change	Dwellings(2012) [9]	Area(km2) [10]	Populationdensity
0	Abbeydale	NE/SE	Northeast	10	Residential	82	5917.0	5700.0	3.8	2023.0	1.7	3480.6
1	Acadia	SE	South	9	Residential	27	10705.0	10615.0	0.8	5053.0	3.9	2744.9
2	Albert Park/Radisson Heights	SE	East	10	Residential	75	6234.0	6217.0	0.3	2709.0	2.5	2493.6
3	Altadore	SW	Centre	11	Residential	39	9116.0	8907.0	2.3	4486.0	2.9	3143.4
4	Alyth/Bonnybrook	SE	Centre	9	Industrial	208	16.0	17.0	-5.9	14.0	3.8	4.2

The dataframe will be enriched by writing a function to append "Calgary" to each neighborhood to enhance the chances of looking-up the coordinates of each neighborhood.

```
calgary_df.head()
```

```

      Neighborhood      Location
1             Acadia  South, Calgary
2  Albert Park/Radisson Heights  East, Calgary
3              Altadore  Centre, Calgary
4    Alyth/Bonnybrook  Centre, Calgary
5    Applewood Park    East, Calgary

```

Further enrichment of the dataframe will be performed to include the coordinates (latitude and longitude) for each neighbourhood using the geopy library.

```
[14]: calgary_df.head()
```

```

[14]:
      Neighborhood      Location  Latitude  Longitude
1             Acadia  South, Calgary  50.968655  -114.055587
2  Albert Park/Radisson Heights  East, Calgary  51.044845  -113.990195
3              Altadore  Centre, Calgary  51.015104  -114.100756
4    Alyth/Bonnybrook  Centre, Calgary  51.016669  -114.024294
5    Applewood Park    East, Calgary  51.044658  -113.928931

```

The Calgary demographic data will then be used as an input to the foursquare location data to fetch top 100 restaurants nearby to each neighbourhood within 4000 meters radius.

Finally, we will make a decision on the most promising neighborhood to open a restaurant in Calgary by examining each cluster of neighbourhoods using the k-means clustering algorithm.