Predicting Location to Open New Restaurant in Edmonton

(Coursera Capstone Project - The Battle of Neighborhoods – Week2)

Mary Ngembu

December 21 2019

1. Introduction

1.1 Background

Edmonton, capital of Canada's Alberta province, sits on the North Saskatchewan River. The city had a population of 932,546 in 2016, making it Alberta's second-largest city and Canada's fifth-largest municipality. Also, in 2016, Edmonton had a metropolitan population of 1,321,426, making it the sixth-largest census metropolitan area (CMA) in Canada. Edmonton is North America's northernmost metropolitan area with a present population over one million. Since I am a resident of Edmonton, I decided to use Edmonton as the research city in my project.

Due to cultural diversity that has been fueled by Canada's generous immigration policies, Edmonton is highly multicultural; more so because Edmonton is the major economic centre for the oil and gas industry. These provide favorable grounds for business opportunities, thus a highly competitive market. So, a careful analysis has to be carried out in order to make informed decisions on starting a new business here. Information derived from this analysis will provide deep knowledge of the neighborhoods that have a competitive edge.

1.2 Problem

Edmonton's cultural diversity also means the city is blessed with restaurants offering an array of international cuisines. In this very competitive market, a potential new French restaurant, ABC Ltd is faced with this problem: Which neighborhood will be the best location for a new French restaurant? The best locations that will be considered here will be the ones that have relatively fewer or no French restaurants, better accessibility to raw materials (in this case, ingredients used for cooking) and a potential for customer volumes.

1.3 Interest

The target audience for this analysis is an investor seeking to start a French restaurant business in Edmonton.

1.4 Success Criteria:

The success criteria of the project will be a good recommendation of neighborhood choice to ABC Company Ltd based on lack of such restaurants in that location, nearness to suppliers of ingredients, and nearness to offices for customers.

2. Data

Based on the problem, the following data was used to solve the problem:

- List of all neighborhoods in Edmonton. This was retrieved from Canada's geonames data available online on geonames.org. https://www.geonames.org/postalcode-search.html?q=Edmonton&country=CA&adminCode1=AB. This data was useful in this analysis in that it will give a big picture of all neighborhoods available to explore.
- The coordinates of the neighborhoods and venues which were retrieved from the same geonames data. This was used to map the location of the venues.
- The top 10 Venues of all Edmonton neighborhoods. This was retrieved form the FourSquare location data. This data was used to cluster neighborhoods into groups.

3. Methodology

In order to establish the targeted neighborhood(s), I explored the demographics of the neighborhoods in the city of Edmonton by segmenting the data and grouping them into clusters to find similar neighborhoods. For clustering the data, I used K-means clustering algorithm which is a form of unsupervised machine learning

I then conducted descriptive analysis using Pandas.

Additional data was gleaned by web scraping.

Using credentials of Foursquare API, near-by venues of the neighborhoods were mined. Due to http request limitations the number of places per neighborhood parameter was reasonably be set to 100 and the radius parameter was be set to 500.

3.1. Data 1:

This data was retrieved from Canada's geonames data and wrangled to create a Pandas dataframe. The dataframe contains all the 38 neighborhoods in the city of Edmonton, with their

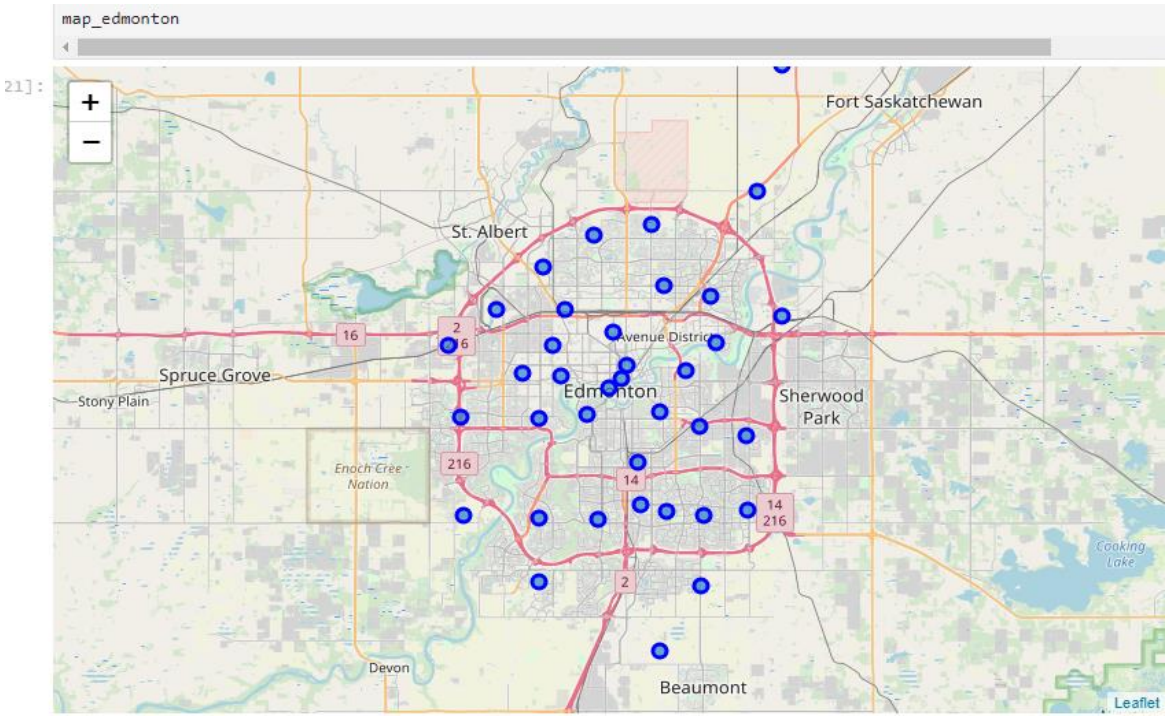
corresponding postal codes and latitude and longitude coordinates.

```
edmonton
```

6]:

	Neighborhood	Code	Latitude	Longitude
0	Edmonton (Kaskitayo)	T6J	53.456	-113.521
1	Edmonton Southwest	T6M	53.459	-113.655
2	Edmonton (Riverbend)	T6R	53.457	-113.58
3	Edmonton (Meadows)	T6T	53.462	-113.371
4	Edmonton (Ellerslie)	T6X	53.415	-113.418
5	Edmonton (Central Londonderry)	T5C	53.6	-113.455
6	Edmonton (North Downtown)	T5J	53.543	-113.497
7	Edmonton (Glenora / SW Downtown Fringe)	T5N	53.544	-113.557
8	Edmonton (North Jasper Place)	T5P	53.546	-113.596
9	Edmonton (West Northwest Industrial / Winterburn)	T5S	53.563	-113.67
10	Edmonton (Central Mistatim)	T5V	53.585	-113.622
11	Edmonton (Central Beverly)	T5W	53.565	-113.402
12	Edmonton (East Castledowns)	T5X	53.631	-113.524

I then visualized the neighborhoods by creating a map of Edmonton with neighborhoods superimposed, using the folium library:



3.2. Data 2:

I used the Foursquare API to collect detailed data on the venues of the neighborhoods. This revealed JSON files like this:

```
[24]: {'meta': {'code': 200, 'requestId': '5dff5c4daba297001b70a7ed'
       'response': {'headerLocation': 'Blue Quill',
        'headerFullLocation': 'Blue Quill, Edmonton',
        'headerLocationGranularity': 'neighborhood',
        'totalResults': 12,
        'suggestedBounds': {'ne': {'lat': 53.460500004500005,
          'lng': -113.51345666563903},
         'sw': {'lat': 53.4514999955, 'lng': -113.52854333436098}},
        'groups': [{'type': 'Recommended Places',
          'name': 'recommended',
          'items': [{'reasons': {'count': 0,
             'items': [{'summary': 'This spot is popular',
               'type': 'general',
               'reasonName': 'globalInteractionReason'}]},
            'venue': {'id': '4b40079cf964a52080b425e3',
             'name': 'Bistecca Italian Steakhouse Wine Bar',
             'location': {'address': '2345 111 St '
```

After some wrangling, I got a dataframe containing details of all Edmonton neighborhoods and details of their nearby venues. Actually, 104 unique venue categories were returned.

```
edmonton_venues.head()
```

```
(270, 7)
```

[31]:

	Neighborhood	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Edmonton (Kaskitayo)	53.456	-113.521	Bistecca Italian Steakhouse Wine Bar	53.454474	-113.514847	Steakhouse
1	Edmonton (Kaskitayo)	53.456	-113.521	Sicilian Pasta Kitchen South	53.453466	-113.525314	Italian Restaurant
2	Edmonton (Kaskitayo)	53.456	-113.521	Spring Roll Kitchen	53.455403	-113.514724	Vietnamese Restaurant
3	Edmonton (Kaskitayo)	53.456	-113.521	Vinny's Pub	53.453520	-113.524898	Pub
4	Edmonton (Kaskitayo)	53.456	-113.521	Brewsters Century Park	53.454629	-113.515185	Brewery

3.3. Analysis:

The one-hot coding was then used to analyze each neighborhood to get the 10 most common venues of each neighborhood. This gave us a picture of the 10 most common venues in each neighborhood. Since one of the criteria for this project is to choose a neighborhood that doesn't have any French restaurants, at this point, I could immediately eliminate all those with 'French Restaurant' or 'Creperie' in their list of 10 most common venues

```
----Edmonton (Central Bonnie Doon)----
                  venue  freq
0   American Restaurant  0.08
1         Breakfast Spot  0.08
2               Gym Pool  0.08
3          Grocery Store  0.08
4               Pharmacy  0.08
5  Fast Food Restaurant  0.08
6           Coffee Shop  0.08
7         Clothing Store  0.08
8          Dance Studio  0.08
9          Liquor Store  0.08


----Edmonton (Central Jasper Place / Buena Vista)----
                venue  freq
0   Sushi Restaurant   0.2
1             Bakery   0.2
2     Sandwich Place   0.2
3       Liquor Store   0.2
4        Pizza Place   0.2
5  Korean Restaurant   0.0
6          Juice Bar   0.0
7           Pharmacy   0.0
8          Pet Store   0.0
9               Park   0.0


----Edmonton (Central Londonderry)----
                venue  freq
0  Recreation Center   1.0
1              Plaza   0.0
2        Pizza Place   0.0
3           Pharmacy   0.0
4          Pet Store   0.0
5               Park   0.0
```

```
neighborhoods_venues_sorted
```

10]:

Neighborhood	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
Edmonton (Central Beverly)	Park	Wine Shop	Hotpot Restaurant	Creperie	Dance Studio	Diner	Discount Store	Eastern European Restaurant	Electronics Store	Farmers Market
Edmonton (Central Bonnie Doon)	American Restaurant	Breakfast Spot	Fast Food Restaurant	Coffee Shop	Clothing Store	Dance Studio	Grocery Store	Liquor Store	Pharmacy	Bank
Edmonton (Central Jasper Place / Buena Vista)	Pizza Place	Bakery	Sandwich Place	Liquor Store	Sushi Restaurant	Food & Drink Shop	Creperie	Dance Studio	Diner	Discount Store
Edmonton (Central Londonderry)	Recreation Center	Wine Shop	Convenience Store	Creperie	Dance Studio	Diner	Discount Store	Eastern European Restaurant	Electronics Store	Farmers Market
Edmonton (East Castledowns)	Convenience Store	Whisky Bar	Creperie	Dance Studio	Diner	Discount Store	Eastern European Restaurant	Electronics Store	Farmers Market	Fast Food Restaurant
Edmonton (East Southeast Industrial / South Cl...	Construction & Landscaping	Pet Store	Cosmetics Shop	Creperie	Dance Studio	Diner	Discount Store	Eastern European Restaurant	Electronics Store	Farmers Market

I then clustered the neighborhoods using K-Means, based on their 10 most common venues, and used folium to visualize the resulting 5 clusters.

Cluster 1

```
]: del list
```

```
]: edmonton_merged.loc[edmonton_merged['Cluster Labels'] == 0, edmonton_merged.columns[[0]+list(range(5,edmonton_merge
```

]:

	Neighborhood	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
11	Edmonton (Central Beverly)	Park	Wine Shop	Hotpot Restaurant	Creperie	Dance Studio	Diner	Discount Store	Eastern European Restaurant	Electronics Store	Farmers Market

Cluster 2

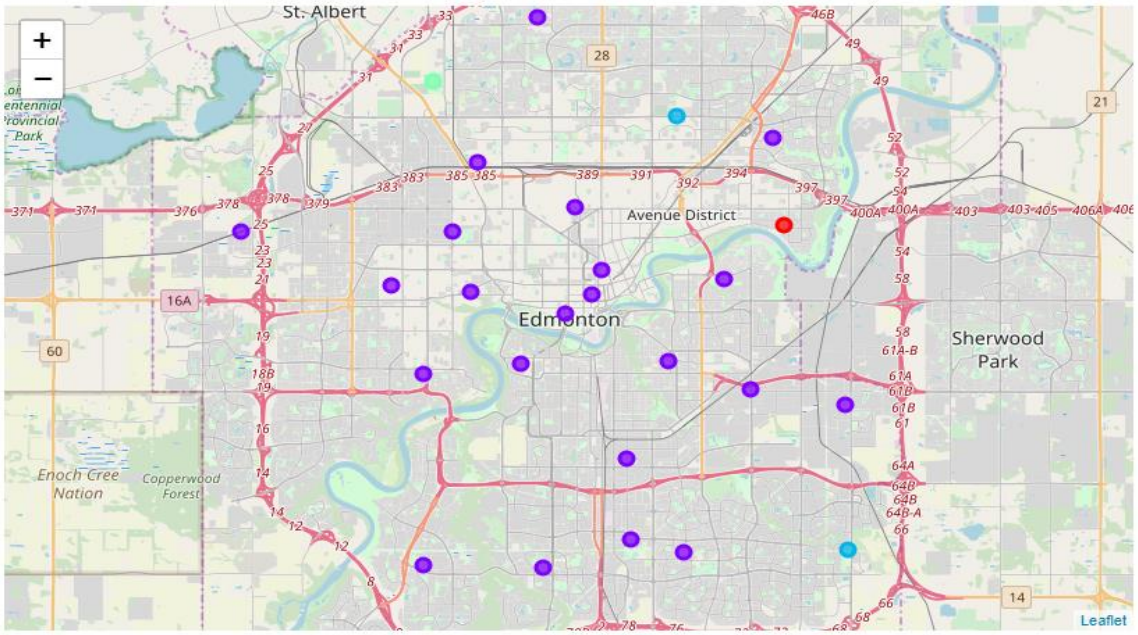
```
]: edmonton_merged.loc[edmonton_merged['Cluster Labels'] == 1, edmonton_merged.columns[[0]+list(range(5,edmonton_merge
```

]:

	Neighborhood	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
0	Edmonton (Kaskitayo)	Pharmacy	Italian Restaurant	Pizza Place	Pub	Coffee Shop	Shoe Store	Liquor Store	Steakhouse	Brewery	Bus Station
2	Edmonton (Riverbend)	Ice Cream Shop	Bus Station	Coffee Shop	Gym	Farmers Market	Dance Studio	Diner	Discount Store	Eastern European Restaurant	Electronics Store
6	Edmonton (North Downtown)	Coffee Shop	Sandwich Place	Italian Restaurant	Hotel	Café	Pub	Bar	Restaurant	Fast Food Restaurant	Vegetarian / Vegan Restaurant

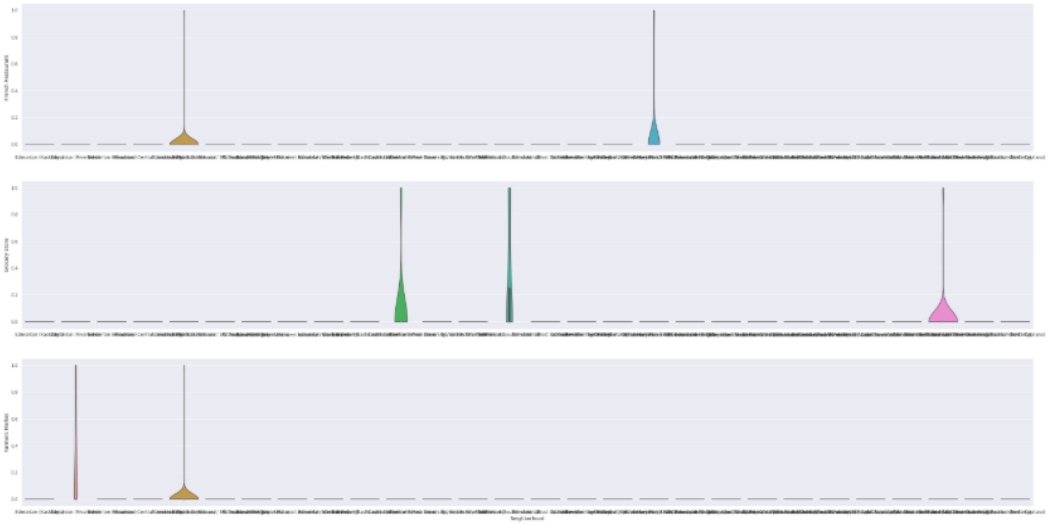
```
map_clusters
```

[45]:



In order to demonstrate the frequency of occurrence of French restaurants in all neighborhoods, I also used the violin plots from the seaborn library - this is a great way to visualize frequency distribution datasets, as they display a density estimation of the underlying distribution

Frequency distribution for the top 3 venue categories for each neighborhood (click to enlage)



The analysis went further to find every neighborhood from the previous result, that has at least 1 French restaurant or Creperie. This was done by specifying the French restaurant category ID of 4bf58dd8d48988d10c941735 when getting near-by venues. The resulting neighborhoods in the dataframe below are thus eliminated:

```
frenchR_grouped
```

[56]:

	0	1	2	4	5
0	Edmonton (Central Bonnie Doon)	53.522	-113.459	53.522588	-113.466897
1	Edmonton (North Downtown)	53.543	-113.497	53.540711	-113.496410
2	Edmonton (North Jasper Place)	53.546	-113.596	53.545543	-113.582920
3	Edmonton (North and East Downtown Fringe)	53.551	-113.492	53.542724	-113.490677
4	Edmonton (South Downtown / South Downtown Fringe)	53.537	-113.510	53.539725	-113.509504

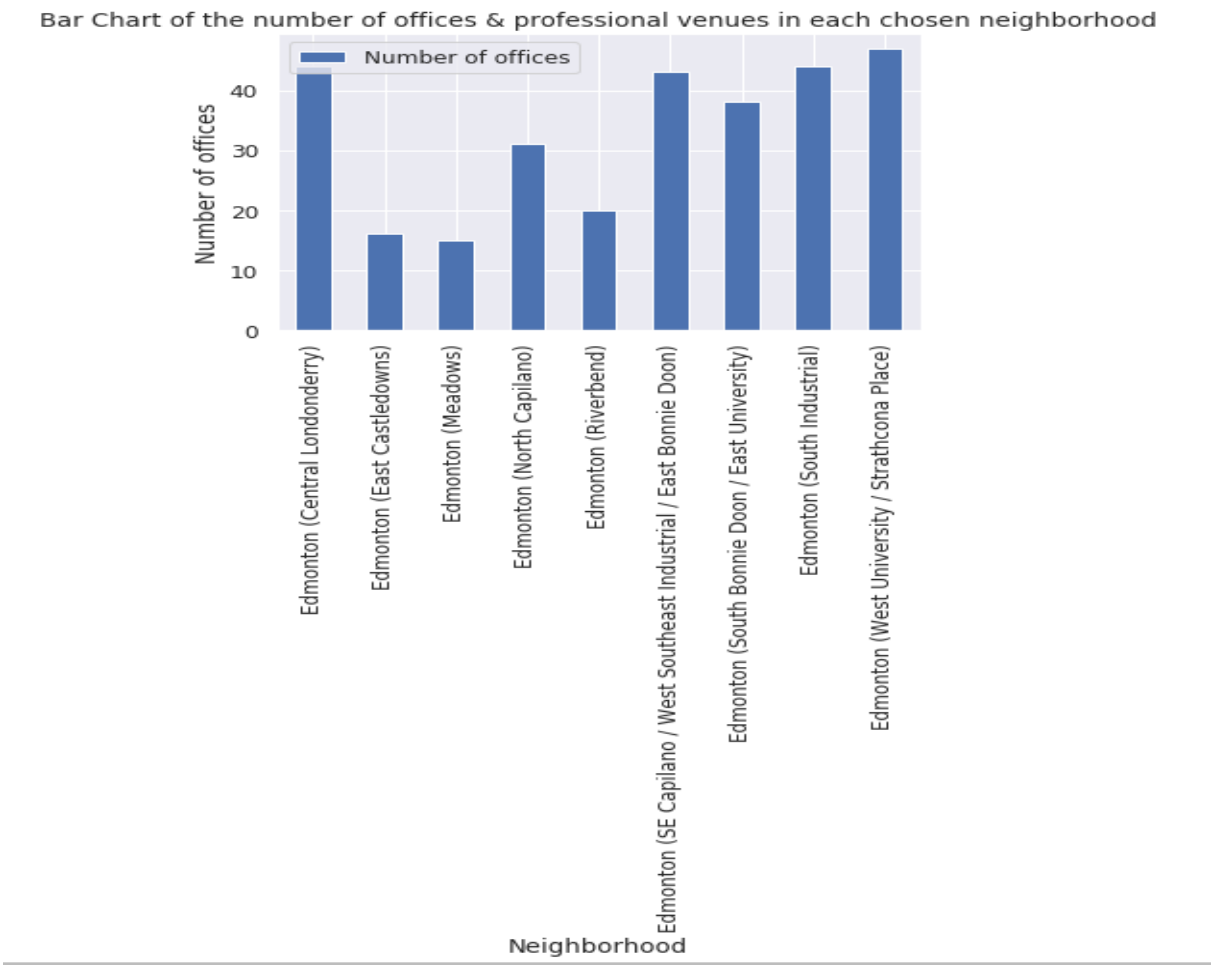
Finally, a further get-nearby-venues analysis was done on the 9 remaining neighborhoods we now have to choose from. Here, we used the criterion of nearness to offices and other professional venues. After much data wrangling, the following dataframe was produced showing a comparison of the number of offices available in each of our 9 chosen neighborhoods.

```
offices_grouped
```

[82]:

Neighborhood	Number of offices
Edmonton (Central Londonderry)	44
Edmonton (East Castledowns)	16
Edmonton (Meadows)	15
Edmonton (North Capilano)	31
Edmonton (Riverbend)	20
Edmonton (SE Capilano / West Southeast Industrial / East Bonnie Doon)	43
Edmonton (South Bonnie Doon / East University)	38
Edmonton (South Industrial)	44
Edmonton (West University / Strathcona Place)	47

I then used matplotlib's bar chart to visualize this data and pick the best neighborhood location for a new French restaurant in Edmonton.



4. **Results**

As can be seen from the analysis, based on the presence of French restaurants in the top 10 nearby venues, the following 12 neighborhoods could be removed from the list of possible locations as they each have French restaurant as one of their top 10 venues:

- Central Beverly
- East Southwest Industrial/South Clareview
- Ellerslie
- Glenora/SW Downtown
- Heritage Valley
- Landbank/Oliver/East Lake District
- North Westmount/West Calder
- South Downtown and Fringe

- West Castledowns
- West Clareview/East Londonderry
- West Lake District
- West Northwest Industries/Winterburn

From the rest of the neighborhoods, based on the factor of nearness to raw materials (suppliers) for the restaurant business, the analysis is based on the availability of Grocery Stores or farmers Markets.

Therefore, the following list of 12 neighborhoods will provide a better choice of location of a new French restaurant considering that there is presently very few (or no) French restaurants and because they each have a relatively higher concentration of Grocery Stores or Farmers Markets:

- Central BonnieDoon
- Central Londonderry
- East Castledowns
- Meadows
- North Capilano
- North Jasper Place
- North&East Downtown Fringe
- Riverbend
- Capilano/West Southeast Industrial
- South BonnieDoon/East University
- South Industrial
- West University/Stratcona

After analysing this list, it was revealed that 3 of the neighborhoods have atleats one French restaurant. These were: North Jasper Place, Central BonnieDoon, North&East Downtown Fringe. So our list of possible favourable neighborhoods was reduced to 9:

- Central Londonderry
- East Castledowns
- Meadows
- North Capilano
- Riverbend
- Capilano/West Southeast Industrial
- South BonnieDoon/East University
- South Industrial
- West University/Stratcona

I was finally able to choose the best neighborhood location after the nearness to offices analysis revealed that among the 9 neighborhoods above, the one surrounded by the highest number of offices and professional venues is West University / Strathcona Place.

5. Discussions

According to this analysis, West University / Strathcona Place area will provide least competition for an upcoming French restaurant as college gym is the most common venue in this area and, it has the highest number of potential customers compared to other comparable neighborhoods. So, definitely this region could potentially be a target for starting the restaurant.

Some drawbacks of this analysis are — the clustering is completely based on the most common venues obtained from Foursquare data. Since land price, distance of the venues from closest stations, could all play a major role and thus, this analysis is definitely far from being conclusory. However, it certainly gives us some very important preliminary information on possibilities of opening restaurants around Edmonton.

Furthermore, these results also could potentially vary if we use some other clustering techniques like DBSCAN.

From the clustering, it becomes clear that concentration radiates from the centre in a circular pattern outward in terms of popular venues. While it was not included in the research and analysis, commercial square footage rental rates tend to follow this trend as well. This should also be considered when choosing a location to open a new restaurant.

6. Conclusion

In conclusion, the scope of this of the analysis is somewhat limited, as not all factors that could influence a location decision of a new French restaurant were examined. Also, the demographics are ever changing, and the information afforded us may be outdated due to relying on FourSquare information. Overall though, the model created can easily be replicated with monitored data through the Foursquare API. All the same, with the data analyzed, I stand by the recommendations made.