Introduction

There is lot of diversity in Canada with people coming from all across the world to stay and live here in Canada. People migrate from different parts of the world to settle here and make a better life for themselves. With diverse people comes diverse cuisine that every Canadian can taste in without having to trouble outside the country.

With this idea, I am planning to use my data science skills and python knowledge to build a recommendation engine that any new business owner can use to understand in part of the city in Calgary, it would be a good idea to open a new restaurant.

Business Problem

In this project I will try to find a location where it would be optimum to open a new restaurant in the city of Calgary, Canada. Since there are lots of restaurants in Calgary we will try to identify those locations where there are less number of restaurants but at the same time, location which is not too far away from the city boundaries. Another interesting view would to see if there are any other specific cuisine restaurants like Italian, Mexican, Indian etc. in the location where I will be recommending it.

Data Requirements

I will be making use of City of Calgary open source data set that provides list of all neighborhoods in the city. Here is an example of this data set.

Area	Type	Sector
SUNALTA	RESIDENTIAL	CENTER
WEST SPRINGS	RESIDENTIAL	WEST
HORIZON	INDUSTRIAL	NORTHEAST

Along with that I will be making use of geocoder to extract geographical information of each neighborhood. Finally, I will also be using FOURSQUARE API data to get venue details corresponding to each neighborhood to understand which location would be a good candidate for building a new restaurant.

For example, after getting the neighborhood information of Calgary outside Downtown Core area (where most of the restaurants usually are), FOURSQUARE API data will be used to extract nearby venues for each neighborhood such coffee shops, grocery stores, restaurants. Filtering on the restaurants category, we can get an idea of how many restaurants are in each neighborhood and what type of restaurants are situated. If FOURSQUARE API data is available only for a certain area of the city, I will limit my analysis to only part due to limited data availability.

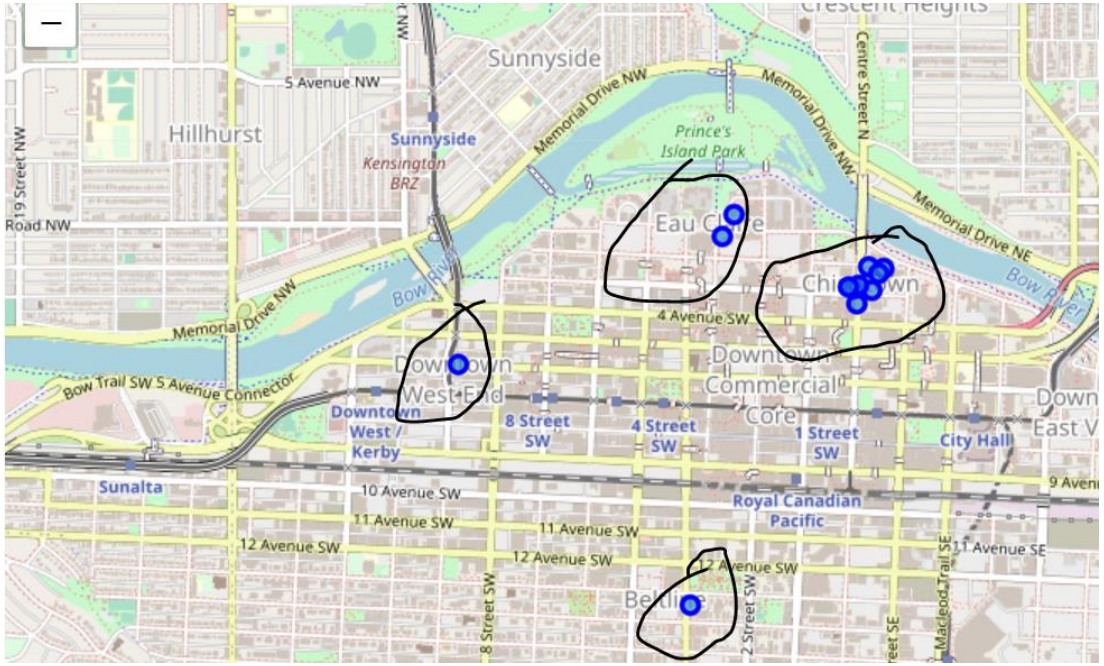
Methodology

Using City of Calgary open source data set and combining that with FOURSQUARE API data to get neighborhood venues, I found that most of the data from FOURSQUARE API was available for downtown area only. As a result, I filtered the data set to SECTOR = CENTRE and non-Industrial Class as shown in Fig (1).
Next, after extracting the results of FOURSQUARE API search with keyword “Restaurants”, four areas in downtown area were found as highlighted in Fig (2).

```
[61]: df3 = df3[df3['CLASS'] != 'Industrial']
      df3 = df3[df3['SECTOR'] == 'CENTRE']
      df3.head()
```

[61]:

	Address	City	Province	SECTOR	CLASS
0	SUNALTA	Calgary	AB	CENTRE	Residential
3	WINDSOR PARK	Calgary	AB	CENTRE	Residential
10	ALTADORE	Calgary	AB	CENTRE	Residential
11	HILLHURST	Calgary	AB	CENTRE	Residential
15	PARKDALE	Calgary	AB	CENTRE	Residential



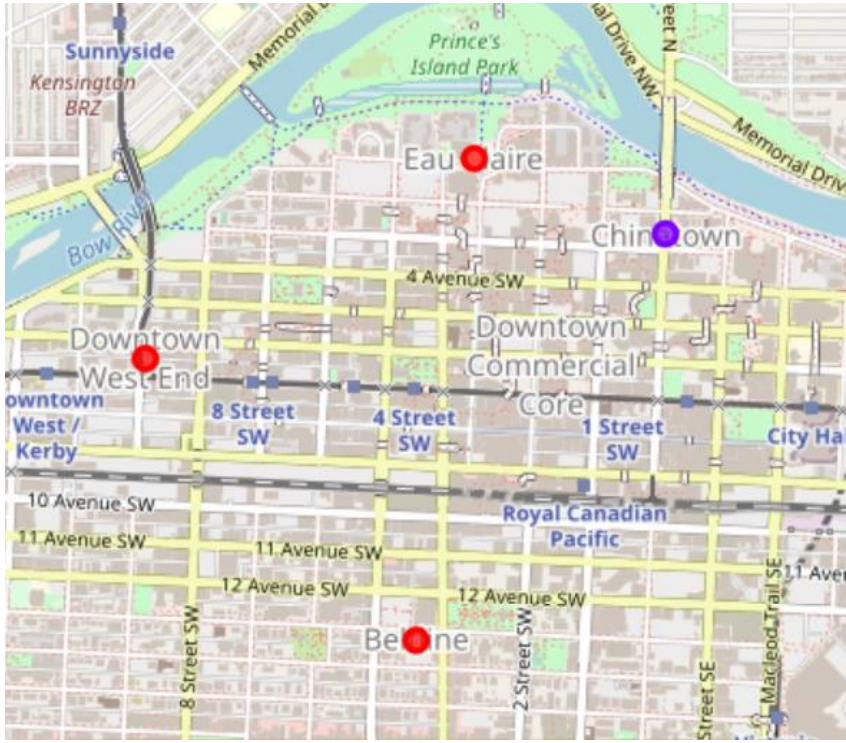
Results

The restaurants types that FOURSQUARE API identified were mainly Chinese, Middle Eastern, European and American restaurants as shown in Fig (3) below.

These areas within downtown were further clustered into two groups using K Means Clustering algorithm as shown in Fig (4). This helped in understanding which areas would or would not be a good candidate for a prospective new restaurant area.

[97]:

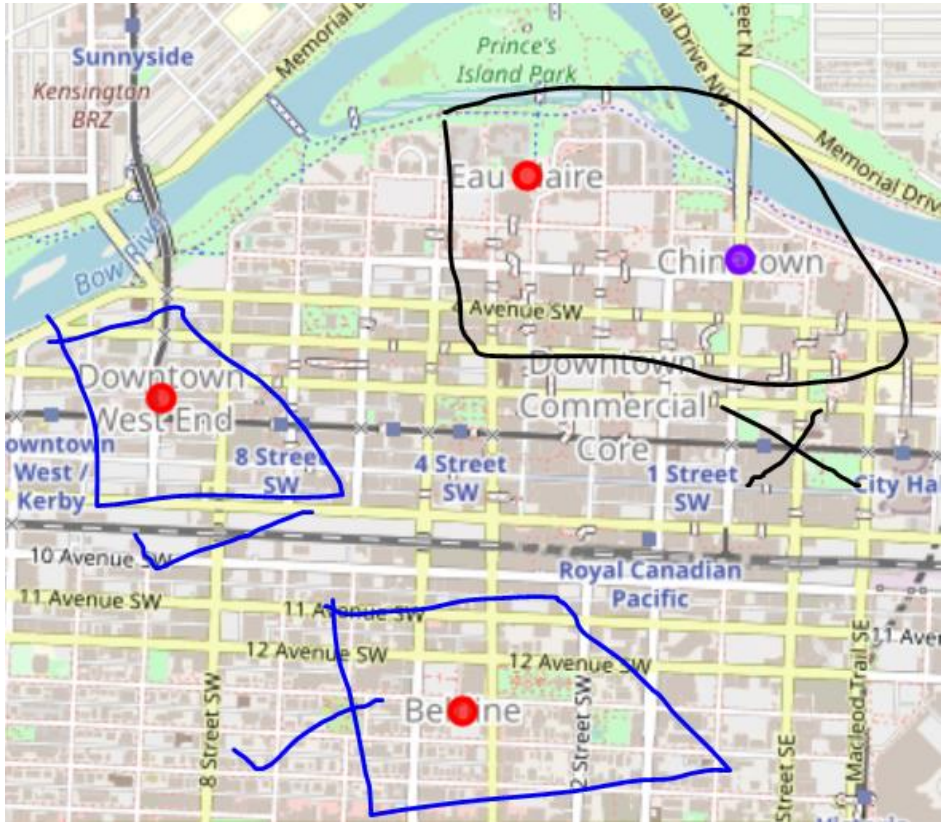
	Neighborhood	Asian Restaurant	Chinese Restaurant	Eastern European Restaurant	Middle Eastern Restaurant	New American Restaurant	Sushi Restaurant
3	CHINATOWN ,Calgary, AB	0	1	0	0	0	0
4	CHINATOWN ,Calgary, AB	1	0	0	0	0	0
5	CHINATOWN ,Calgary, AB	0	1	0	0	0	0
6	CHINATOWN ,Calgary, AB	0	1	0	0	0	0
7	CHINATOWN ,Calgary, AB	0	1	0	0	0	0
8	CHINATOWN ,Calgary, AB	0	1	0	0	0	0
10	CHINATOWN ,Calgary, AB	0	1	0	0	0	0
11	CHINATOWN ,Calgary, AB	0	1	0	0	0	0
15	DOWNTOWN WEST END ,Calgary, AB	0	0	1	0	0	0
24	EAU CLAIRE ,Calgary, AB	0	0	0	0	1	0
28	EAU CLAIRE ,Calgary, AB	0	0	0	0	0	1
32	BELTLINE ,Calgary, AB	0	0	0	1	0	0



Discussion

Based on the data available and assuming we are limited to Centre/downtown part of the city, I would recommend the business owners to open restaurants near Downtown West End or Beltline area where there is only one Eastern European and Middle Eastern restaurant respectively. Also, American, Italian or Indian cuisine might be a good option since there are very few or no such cuisines in the nearby neighborhood.

Eau Claire area would not be a good candidate as Chinatown is very close to this area with multiple cuisine options already available for customers.



Conclusion

Thus using various data points, using the power of Python to clean the data and applying data science methodologies such as Clustering methods on top of it, we can help new business owners recommend the best area to open new restaurants. In our case study, our area was restricted to Downtown or Centre part of the city and we identified two different areas which could be good candidate to open a new restaurant.

Also, based on the existing restaurant types available in the neighborhood, we could also recommend what type of restaurant (cuisine) might be a good option.