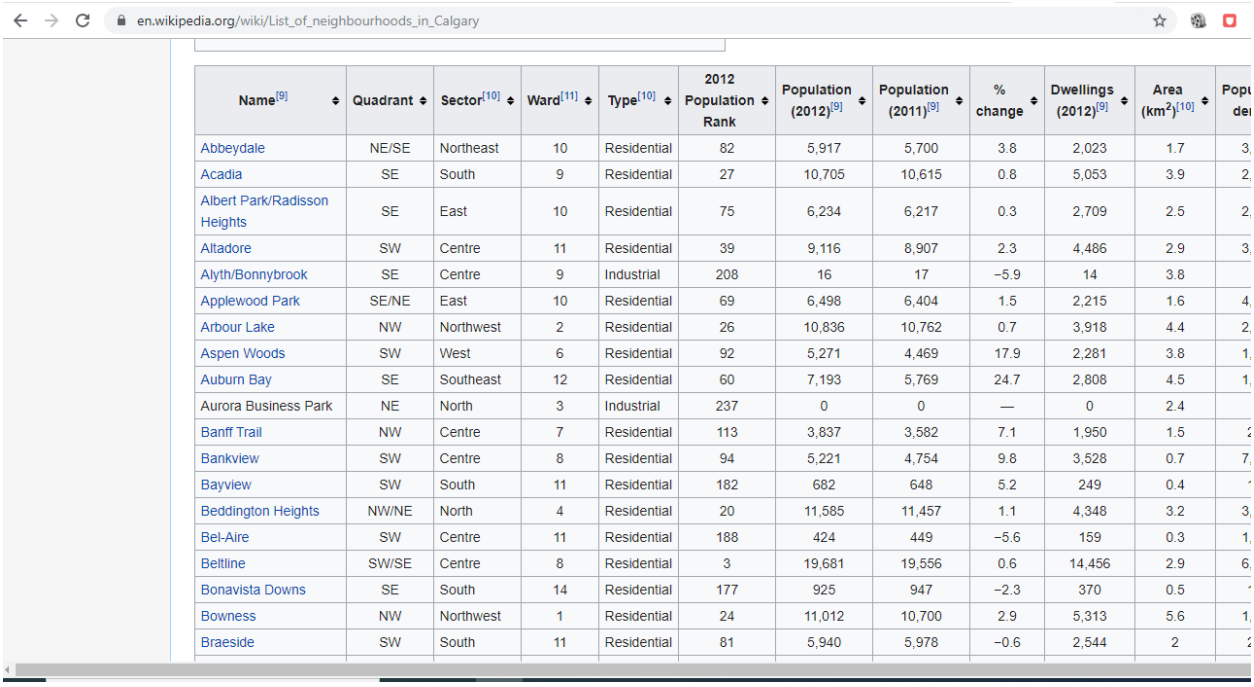
Capstone Final Project – IBM Data Science

Introduction:

In the Final Capstone project, we are going to explore neighborhoods in Calgary, Alberta Canada. We will explore how restaurants are distributed across different neighborhoods of Calgary. Data visualization for different restaurants across different parts of Calgary will help determine potential entrepreneurs to decide which part(s) of Calgary could be considered to for opening a restaurant. This project brings together solid data analysis and visualization to serves as a good initial basis for any entrepreneur in the food and services industry.

Data:

The neighborhood data that we are going to use will be retrieved from Wikipedia. Using the neighborhood data from wiki, we the used it to get the latitude and longitude of the neighborhoods.

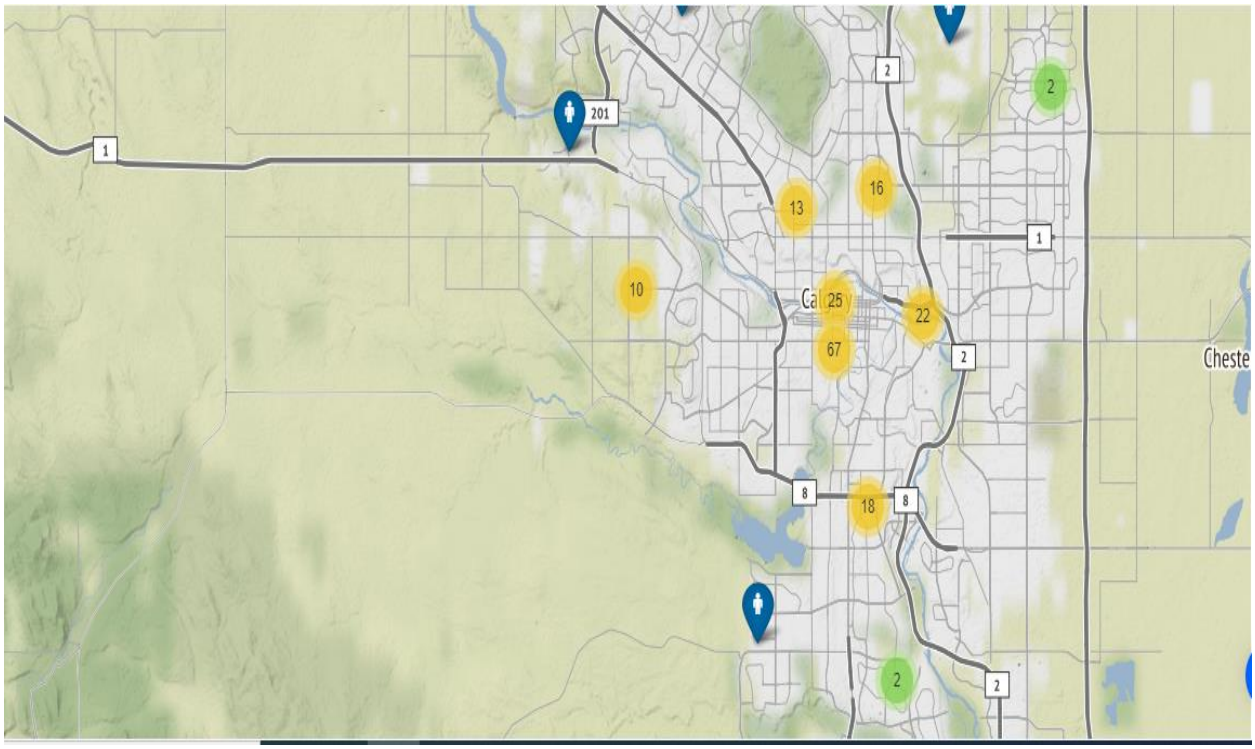


Name[9]	Quadrant	Sector[10]	Ward[11]	Type[10]	2012 Population Rank	Population (2012)[9]	Population (2011)[9]	% change	Dwellings (2012)[9]	Area (km²)[10]	Popu de
Abbeydale	NE/SE	Northeast	10	Residential	82	5,917	5,700	3.8	2,023	1.7	3,
Acadia	SE	South	9	Residential	27	10,705	10,615	0.8	5,053	3.9	2,
Albert Park/Radisson Heights	SE	East	10	Residential	75	6,234	6,217	0.3	2,709	2.5	2,
Altadore	SW	Centre	11	Residential	39	9,116	8,907	2.3	4,486	2.9	3,
Alyth/Bonnybrook	SE	Centre	9	Industrial	208	16	17	−5.9	14	3.8	
Applewood Park	SE/NE	East	10	Residential	69	6,498	6,404	1.5	2,215	1.6	4,
Arbour Lake	NW	Northwest	2	Residential	26	10,836	10,762	0.7	3,918	4.4	2,
Aspen Woods	SW	West	6	Residential	92	5,271	4,469	17.9	2,281	3.8	1,
Auburn Bay	SE	Southeast	12	Residential	60	7,193	5,769	24.7	2,808	4.5	1,
Aurora Business Park	NE	North	3	Industrial	237	0	0	—	0	2.4	
Banff Trail	NW	Centre	7	Residential	113	3,837	3,582	7.1	1,950	1.5	2
Bankview	SW	Centre	8	Residential	94	5,221	4,754	9.8	3,528	0.7	7,
Bayview	SW	South	11	Residential	182	682	648	5.2	249	0.4	1
Beddington Heights	NW/NE	North	4	Residential	20	11,585	11,457	1.1	4,348	3.2	3,
Bel-Aire	SW	Centre	11	Residential	188	424	449	−5.6	159	0.3	1,
Beltline	SW/SE	Centre	8	Residential	3	19,681	19,556	0.6	14,456	2.9	6,
Bonavista Downs	SE	South	14	Residential	177	925	947	−2.3	370	0.5	1
Bowness	NW	Northwest	1	Residential	24	11,012	10,700	2.9	5,313	5.6	1,
Braeside	SW	South	11	Residential	81	5,940	5,978	−0.6	2,544	2	2

```
import requests
from bs4 import BeautifulSoup
import pandas as pd

WIKI_URL = "https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Calgary"

req = requests.get(WIKI_URL)
soup = BeautifulSoup(req.content, 'lxml')
table_classes = {"class": ["sortable", "plainrowheaders"]}
wikitables = soup.findAll("table", table_classes)
right_table=soup.find('table', class_='wikitable sortable')
#print(right_table)
A=[]
B=[]

for row in right_table.findAll('tr'):
    cells=row.findAll('td')
    if len(cells)==12:
        A.append(cells[0].find(text=True))
        B.append(cells[4].find(text=True))


df=pd.DataFrame(A,columns=['Neighborhood'])
df['Type']=B
df.head()
```

Methodology:

- List of Calgary Neighborhoods was collected from Wikipedia using web scraping technique
- Using geopy package, latitude and longitude were extracted from the data
- The latitude and longitude were used in foursquare API to get the venues' information
- Restaurants' information was extracted from the venue's information
- Data analysis and visuals provided to better understand the distributed of restaurants in Calgary. Folium with enhanced features were used to create map



Results and Conclusion:

Results show that there are many areas of Calgary that there are very few restaurants. Entrepreneurs should take the advantage of this fact and start to research the amount of cost, populations, best suited cuisine and start something. This will not only help them but also to the people of the community who look to go out to different places for dinner.