1.Introduction

1.1 Description of the Problem

Calgary, a cosmopolitan Alberta city with numerous skyscrapers, owes its rapid growth to its status as the centre of Canada's oil industry. However, it's still steeped in the western culture that earned it the nickname "Cowtown," evident in the Calgary Stampede, its massive July rodeo and festival that grew out of the farming exhibitions once presented here.

1.2 Discussion of the Background

My client runs a successful restaurant chain in the Eastern Region of Canada with 5 restaurants spread between Ottawa, Toronto and a Montreal now wants to explore the western restaurant market and has decided to start with the City of Calgary.
As a restaurant owner, it is important that your business is not too close to competitors while ensuring that it is also close to major attractions in the city. Therefore, my client wants to know the best are to locate the new restaurant that will be have the least competition and the most attractions.

1.3 Target Audience

This project is to anyone that wants to set up business in Calgary and will like to know what the competition hotspot is like around the city.

2. Data

This project relies on public data from Wikipedia and Foursquare.

The neighbourhoods in Calgary is gotten from the Wikipedia page

"https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Calgary" which serves as the

foundation for the analysis.

Data about venues in Calgary were gotten from the FourSquare API. Additional analysis was

performed by my client to provide a weighted rank of how much competition or attract a

venue type is on a scale of 1-10.

For example, an hotel has a higher rank of attraction (10) and a minimal rank of competition

(1) giving it a total weight of 9 (10-1).

However, a Pizza Place is a direct competitor and has a lower attraction rank of (3) and a

medium competition rank of (6), giving it a total weight of -3 (3-6).

The expectation is that for a cluster of venues, we take the average of all venue weights and

the cluster with the highest average weight is the best location for our restaurant.

3. Analysis

Extracted the list of neighbourhoods from Wikipedia using the panda's read_html method

to obtain a dataframe containing Calgary's neighbourhoods.

Used the geocoder library to search for the GPS coordinates of the each neighbourhood.

	Name	Quadrant	Sector	Ward	Type	Area	Latitude	Longitude
0	Abbeydale	NE/SE	Northeast	10	Residential	1.7	51.05976	-113.92546
1	Acadia	SE	South	9	Residential	3.9	50.97227	-114.05843
2	Albert Park/Radisson Heights	SE	East	10	Residential	2.5	51.04200	-113.99683
3	Altadore	SW	Centre	11	Residential	2.9	51.01601	-114.10558
4	Alyth/Bonnybrook	SE	Centre	9	Industrial	3.8	51.01155	-114.04481

Looped though each neighbourhood to obtain the first 200 venues within a 2000m radius.

This is enough to cover the whole city.

Note that there are duplicated venues that fall within 2000m of multiple neighbourhoods,

therefore when the listing was merged, the duplicates were removed.

Obtained a listing of the categories of all the venues obtained and assigned Attraction and

Competition weights to them.

Kmeans was used to create 20 cluster

	Cluster Labels	name	categories	lat	lng	Attraction	Competition	Weight
0	0	Atlas Pizza and Sports Bar	Pizza Place	51.052481	-113.941859	3	6	-3
1	0	A&W Canada	Fast Food Restaurant	51.068291	-113.933571	4	8	-4
2	0	Costco Wholesale	Warehouse Store	51.043141	-113.917518	7	2	5
3	0	Subway	Sandwich Place	51.059215	-113.934836	8	4	4
4	0	Calgary Co-op	Grocery Store	51.068719	-113.934014	6	2	4

The Clusters are sorted for the highest average weight which is identified as cluster 14

Cluster Labels	Weight	Count	avg
14	207	66	3.136364
13	259	85	3.047059
1	8	3	2.666667
10	497	190	2.615789
0	434	172	2.523256

List of 5 venues in Cluster 14

	Cluster Labels	name	categories	lat	lng	Attraction	Competition	Weight
1093	14	Yakima Social Kitchen + Bar	Gastropub	51.130619	-114.007792	8	4	4
1094	14	Homewood Suites by Hilton™ - Calgary-Airport	Hotel	51.142683	-114.012082	10	1	9
1095	14	Wyndham Garden Calgary Airport	Hotel	51.140607	-114.026205	10	1	9
1096	14	Starbucks	Coffee Shop	51.132087	-114.011475	6	2	4
1097	14	International Baggage Claim 1	Airport Service	51.130297	-114.004035	7	2	5

Plotted cluster 14 on the map and it turns out to be around the Airport