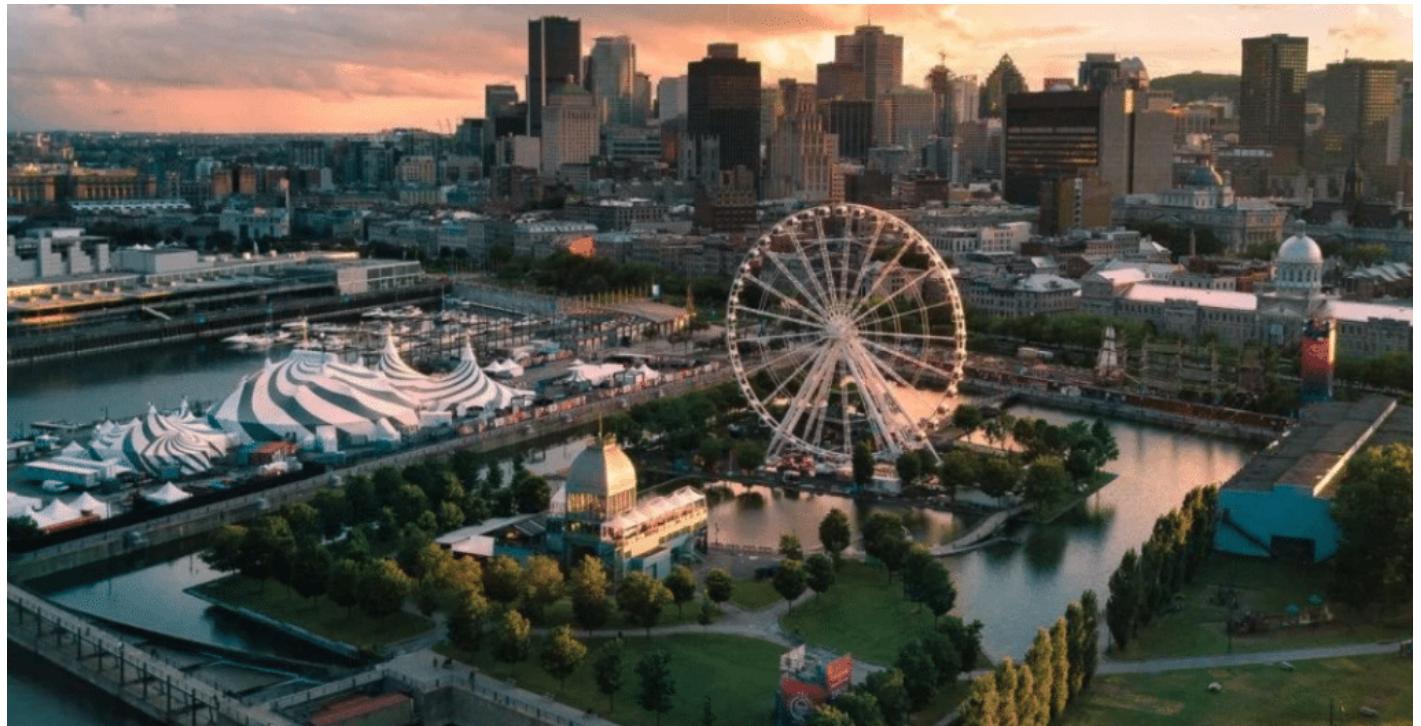


# Business Problem Opening a Cafe in Montreal

by: Cristiane de Souza da Silva

June 2021



*Montreal, Canada*

# Introduction

Montreal is the second-most populous city in Canada and most populous city in the Canadian province of Quebec.

The city is a tourist destination with several recreational areas, historic monuments, numerous bars, restaurants and cafes. An activity that is touristy but also for its residents is drinking coffee. The activity can be done alone, accompanied by friends, and anytime at this city.

Out of 20 countries in the world that drink more coffee per capita, Canadians use about 6.5 kg /year per capita, more than the US and the UK.<sup>1</sup>

On average, 72% of Canadians between 18 and 79 drink coffee on a daily basis but among regular coffee drinkers, the average consumption is 2.8 cups per day (Coffee Association of Canada).<sup>2</sup>

# Business Problem

The objective of this project is analyse and explore an optimal location for a coffee shop. This report will be targeted at stakeholders who want to start from scratch, to buy an existing business, or anybody interested in a good coffee in Montreal, Canada.

Since there are lots of coffee shops in Montreal we will try to detect locations that are not already crowded with coffee shops. We are also particularly interested in areas with no coffee shops in the vicinity. We would also prefer locations as close to the city center(McGill University neighbourhood) as possible, assuming that the first two conditions are met.

---

<sup>1</sup> <https://coffeebi.com/2020/03/16/coffee-in-canada-habits-and-innovation/>

<sup>2</sup> <https://coffeebi.com/2020/03/16/coffee-in-canada-habits-and-innovation/>

# Target Audience

Contractor who wants to start from scratch , buy an existing business or anybody interested in a good coffee.

# Data

For the Montreal neighbourhood data, a [Wikipedia](#) page exists that has all the information we need to explore and cluster the neighbourhoods in the city.

To obtain data on the neighbourhoods and postal codes of Montreal, the Wikipedia page was scrapped. The library used for this task was BeautifulSoup. With the scraping, data were obtained from 124 postal codes, districts, and districts. It was still necessary to add the latitudes and longitudes of each zip code and for that, I used the **pgeocode** library.

Data cleaning was performed after verifying that a postal code (HOP) had no information about its neighbourhood, district, and consequently neither latitude nor longitude. I researched Montreal's zip codes and found the H0 prefix an anomaly. The zero indicates a delivery rural area and therefore it is practically empty. ([postal codes in Canada](#) ). With that, that single zip code was erased.

Through the *boxplot* plot of the postal codes, an outlier of latitude and longitude was observed. The coordinates did not belong to Canada, so they were removed from the dataset.

After that, the data frame had a total of 122 neighbourhoods, and it was stored in such a way that it could be read into a pandas data frame so that it is in a structured format like the Montreal, Quebec dataset.

	PostalCode	Neighborhood
0	H1A	Pointe-aux-Trembles
1	H2A	Saint-Michel,East
2	H3A	McGill University
3	H4A	Notre-Dame-de-GrâceNortheast
4	H5A	Place Bonaventure

Figure 1 - Image by Author. **Montreal neighbourhoods**

## Methodology

In this session, the methodology used for the analysis of the project will be shown.

First, we checked the candidate neighbourhoods. It was created the latitude & longitude coordinates for centroids of them, McGill university neighbourhood.

The visualisation library **Folium** was used to create a map of Montreal with the neighbourhoods superimposed on top, and a grid of cells covering the area of interest which is approx. 20x20 kilometers centered around the *McGill University* neighbourhood.

## IBM Data Science Capstone

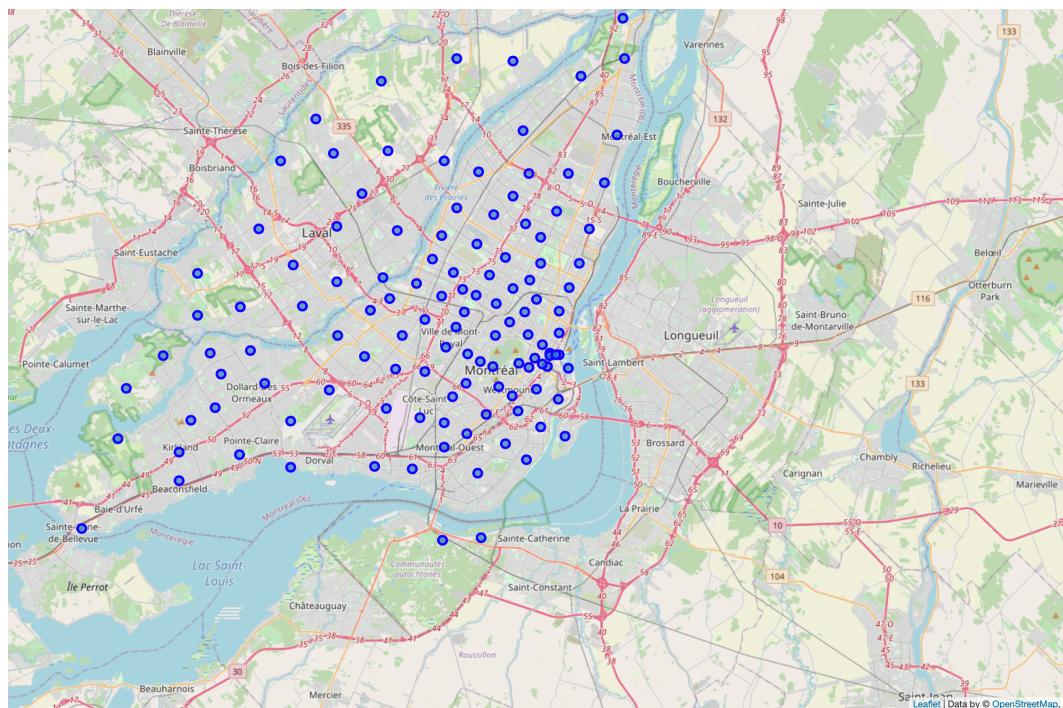


Figure 2 - Image by Author. **Candidate neighbourhood**

To explore the neighbourhoods of the city of Montreal, the **Foursquare API** was used. The interface was used to get the most common categories of locations in each neighbourhood (cafes, restaurants, ice cream parlors, art galleries, etc) and then use this feature to group the neighbourhoods. It's mandatory to create an account in the app to obtain your client credentials.

First, the API was used to get the top 100 venues that are in *McGill university neighborhood* in a radius of 2.5 kilometers.

Second, the analysis verified the frequency of 280 categories over the Montreal area, and after the coffee shop category was filtered

## IBM Data Science Capstone

	Neighborhood	Adult Boutique	Airport	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	Arepa Restaurant	Art Gallery	Art Museum	...	Tunnel	Vegetarian / Vegan Restaurant	Video Game Store	Video Store	Vietnamese Restaurant
0	Laval-sur-le-Lac	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	...	0.0	0.00	0.0	0.0	0.0
1	AhuntsicCentral	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	...	0.0	0.00	0.0	0.0	0.0
2	AhuntsicEast	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	...	0.0	0.00	0.0	0.0	0.0
3	AhuntsicNorth	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	...	0.0	0.00	0.0	0.0	0.0
4	AhuntsicSoutheast	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	...	0.0	0.00	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
117	Vimont	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	...	0.0	0.00	0.0	0.0	0.0
118	WestmountNorth	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	...	0.0	0.01	0.0	0.0	0.0
119	WestmountSouth	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	...	0.0	0.00	0.0	0.0	0.0
120	YUL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	...	0.0	0.00	0.0	0.0	0.0
121	Îles-Laval	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	...	0.0	0.00	0.0	0.0	0.0

122 rows × 281 columns

Figure 3 - Image by Author. **Frequency of each category**

	Neighborhood	Coffee Shop
0	Laval-sur-le-Lac	0.047619
1	AhuntsicCentral	0.040000
2	AhuntsicEast	0.030000
3	AhuntsicNorth	0.034884
4	AhuntsicSoutheast	0.024691

Figure 4 - Image by Author. **Coffee shop frequency**

Then, the k-means clustering algorithm was performed to cluster the neighbourhood into 4 clusters. Again, the Folium library was used to visualize the distribution of the clusters over the city. Each cluster was examined and discriminated based on the frequency of coffee shops in each area, followed by the requirements of this project: areas with few or no coffee shops in the vicinity.

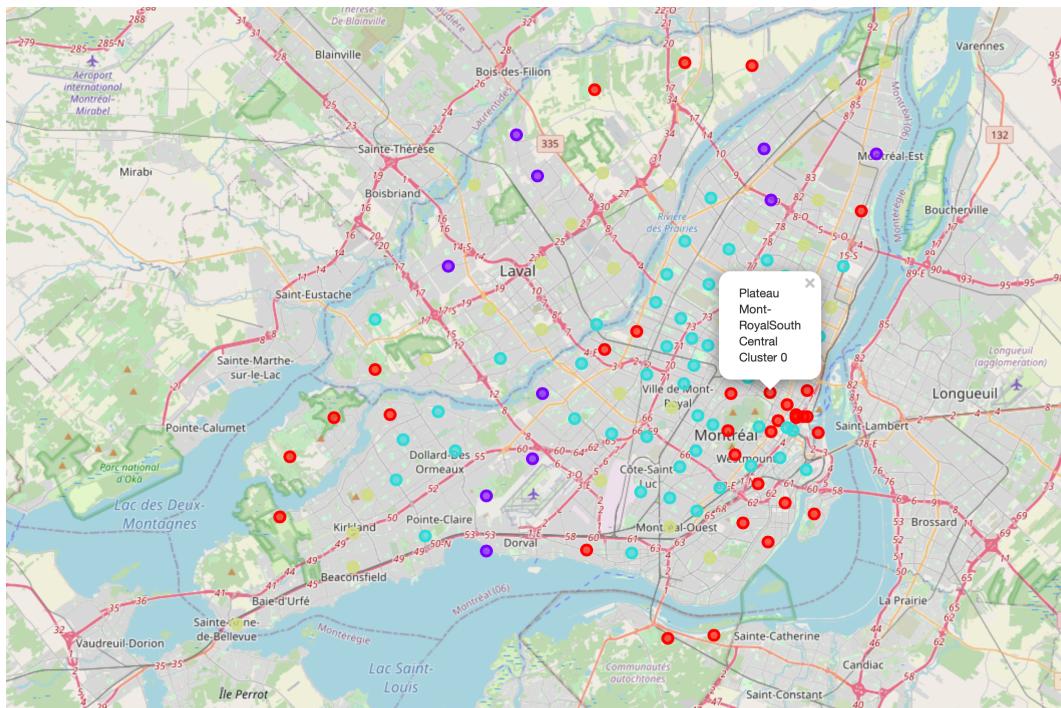


Figure 2 - Image by Author. **Clusters**

Finally, it was calculated and added to the analysis the distance of neighbourhoods where there are few or no coffee shops to the McGill university neighbourhood.

	Neighborhood	Coffee Shop	Cluster Labels	PostalCode	Latitude	Longitude	Distance from McGill University (km)
0	Îles-Laval	0.00	0	H7Y	45.5284	-73.8509	30.702174
1	Saint-François	0.00	0	H7B	45.6757	-73.6388	8.929536
2	Old Montreal	0.01	0	H2Y	45.5057	-73.5550	2.189817
3	RosemontSouth	0.02	0	H1Y	45.5486	-73.5788	1.473474
4	Downtown MontrealNortheast	0.01	0	H2Z	45.5052	-73.5622	1.389576

Figure 5 - Image by Author. **Top 5 neighbourhoods closest to center city and with a little coffee shop frequency**

# Results

This work analysed 122 neighbourhoods in Montreal with the objective of detecting places that are no longer full of coffee shops.

The area with the lowest density of cafeterias in Montreal is concentrated in the peripheral region of the analysed area. This region is related to cluster 0 (red dots on the map). In this cluster, 25 neighbourhoods have a low density.

The neighbourhoods where there is a high incidence of coffee shops are located in cluster 1, many of them close to McGill university, our central coordinate.

However, the neighbourhoods with the highest coffee frequencies are located in cluster 2, and they are not that close to McGill university. They are regions close to the airport, industrial park and more distant schools. The region with the highest incidence of coffee shops is located in the Dorval / L'Île-Dorval neighbourhood and is 22 km away from the central coordinate.

Verdun South is the closest neighbourhood to the McGill university and has low density of coffee shop. It is 1.97 km away from the central coordinate.

There are also places closer and not crowded with coffee shops: Plateau Mont-Royal Southeast ( 0.75km ), Verdun North ( 1.36km ), Place Desjardins, Downtown Montreal Northeast ( both 1.38km ). Conclusion

The neighbourhoods where there is a high incidence of coffee shops are located in cluster 1, many of them close to *McGill university*, our central coordinate.

## Conclusions and Recommendations

The neighbourhoods with the highest coffee frequencies are located in regions close to the airport, industrial park and more distant schools. This makes sense as they are regions of workers who usually consume coffee during their workday. About schools, students' parents usually drink coffee near these places, and if it is college, students have a large consumption of coffee during the year.

Project requirements include detecting places that are not already crowded with cafes, or areas with low density of coffee shop nearby, and as close as possible to the city center (McGill University). The following five options are suggested that meet the project requirements : *Verdun South*, the closest neighbourhood to the *McGill university* with the lowest density of coffee shop with 1.97 km away from the central coordinate; *Plateau Mont-Royal/Southeast*( 0.75km), *Verdun North* (1.36km), *Place Desjardins* (1.38km), and *Downtown Montreal Northeast* (1.38km).

This project only considers the frequency of coffee shops in the neighbourhoods. Other contributions are important for a final decision on the location of the cafeteria, such as other businesses close to it (which would be the other relevant categories that would complement it), cost of living in the neighbourhood, safety zones, among others. In addition, all analyses were performed using the free APIs or their open access option. The analysis can be done with paid resources so that more result options are achieved