Data Science Capstone Project:

The Best Neighborhood in Montreal for Opening a Cafe`/ Coffee Shop

IBM Data Science

Amir Hossein Fattahi

Table of Contents

1. INTRODUCTION 3

1.1. Business Problem 3
1.2. Target Audience 3

2. DATA 3

Administrative boundary of the Montréal agglomeration Data 3
Montreal's Census Profile 4
Foursquare API 4

3. METHODOLOGY 4

Registered Business Data 4
Montreal, Boroughs and Neighborhoods data 5
Census individual Income 6
Merging Datasets¶ 7
MAP Montreal District 8
Foursquare Data Analysis¶ 10
The most common Venues 11
Neighborhood Clustering¶ 12
Coffee shop venue frequency per neighborhood 12
Map Clustered Neighborhoods¶ 13
Clusters Analysis 13

4. RESULTS AND DISCUSSION¶ 15

5. CONCLUSION 16

1. Introduction

1.1. Business Problem

In a metropolitan-multicultural city like Montreal, there are a number of places full of students and businesses which are suitable to open a new Coffee Shop close there. This study is going to help people planning to open a new Coffee Shop in Montreal. This idea has a heated debate over the optimal neighborhood for setting up shop, but the people would like to be located in an appropriate location of the city and in a spot where people already tend to drink Coffee. Assuming, demography of population and income are not an issue of each neighborhood in accordance with competitors already exist on the same regions.

1.2. Target Audience

The target audience of this report would be anyone who wants to buy or establish a Coffee Shop in Montreal, or anyone in Montreal just looking for a nice area to drink a coffee.

2. Data

In order to provide the stakeholders, the necessary information to best make this decision, some data would be needed. For the city of Montreal has some public datasets that describe various aspects of the city, and Foursquare API allows access to collect competitors data on the same neighborhoods. Altogether, there are three sets of data for our analysis:

Administrative boundary of the Montréal agglomeration Data

This is going to help us by providing Polygons delimiting the boroughs of the City of Montreal, boroughs and related cities constituting the agglomeration of Montreal and allowing us to select one of the suitable areas for new Coffee Shop

http://donnees.ville.montreal.qc.ca/dataset/polygones-arrondissements

Montreal's Census Profile

This is going to provide data contains Montreal's 2016 census Aboriginal peoples; Education, Ethnic origin, Families, households and marital status, Housing; Immigration and citizenship, Income; Journey to work, Labour, Language; Language of work, Mobility, Population, Visible minority

https://www12.statcan.gc.ca/census-recensement/2016

Montreal income census 2016 per borough can be downloaded as MS. Excel in the following URL:

http://ville.montreal.qc.ca/pls/portal/url/ITEM/55637C4923B8B03EE0530A930132B03E

Foursquare API

Foursquare data is robust and provides location data allows you to retrieve information about the most popular spots in each neighborhood in Montreal.

3. Methodology

We need to access business data sets including Montreal Census income, Montreal Boroughs and Neighborhood Geo data to merge them altogether. The merged data let's explore appropriate boroughs in Montreal which has more individual number of people with average income to be potential customers. These data should be merged with Foursquare data to find which neighborhoods in these boroughs, Coffee Shop are popular venue to find suitable place to open a new Coffee Shop.

Registered Business Data

- There are two options to access census data for Montreal:
    - **Geographies:** Montréal [Population center], Quebec [Province] Topic(s): Aboriginal peoples; Education; Ethnic origin; Families, households, and marital status; Housing; Immigration and

citizenship; Income; Journey to work; Labour; Language; Language of work; Mobility; Population; Visible minority

https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/download-telecharger/currentactuelle.cfm?Lang=E&Geo1=POPC&Code1=0547&Geo2=PR&Code2=24&B1=All&type=0&FILETYPE=CSV (https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/download-telecharger/currentactuelle.cfm?Lang=E&Geo1=POPC&Code1=0547&Geo2=PR&Code2=24&B1=All&type=0&FILETYPE=CSV)

- **Montreal income census 2016**
  The second option is easier to find relating census data and in this project second option was used.

  https://montreal.ca/

- **Montreal boroughs and neighborhoods**
  - To explore with geographical information to create a map, The GeoJSON file contains the boundaries of the boroughs of the City of Montreal and its cities linked in JSON format could be download from following URL:

http://donnees.ville.montreal.qc.ca/dataset/f38c91a1-e33f-4475-a112-3b84b1c60c1e/resource/a80e611f-5336-4306-ba2a-fd657f0f00fa/download/quartierreferencehabitation.geojson (http://donnees.ville.montreal.qc.ca/dataset/f38c91a1-e33f-4475-a112-3b84b1c60c1e/resource/a80e611f-5336-4306-ba2a-fd657f0f00fa/download/quartierreferencehabitation.geojson)'

Montreal, Boroughs and Neighborhoods data

Examine the results for Montreal Neighborhoods and Boroughs by cleaning data for Montreal boroughs and coordinated neighborhoods is as following table:

	Neighborhood	Borough
0	Rivière-des-Prairies	Rivière-des-Prairies–Pointe-aux-Trembles
1	Beaurivage	Mercier–Hochelaga-Maisonneuve
2	Tétreaultville	Mercier–Hochelaga-Maisonneuve
3	Dupéré	Mercier–Hochelaga-Maisonneuve
4	Guybourg	Mercier–Hochelaga-Maisonneuve
5	Longue-Pointe	Mercier–Hochelaga-Maisonneuve
6	Louis-Riel	Mercier–Hochelaga-Maisonneuve
7	Marie-Victorin	Rosemont–La Petite-Patrie

Census individual Income

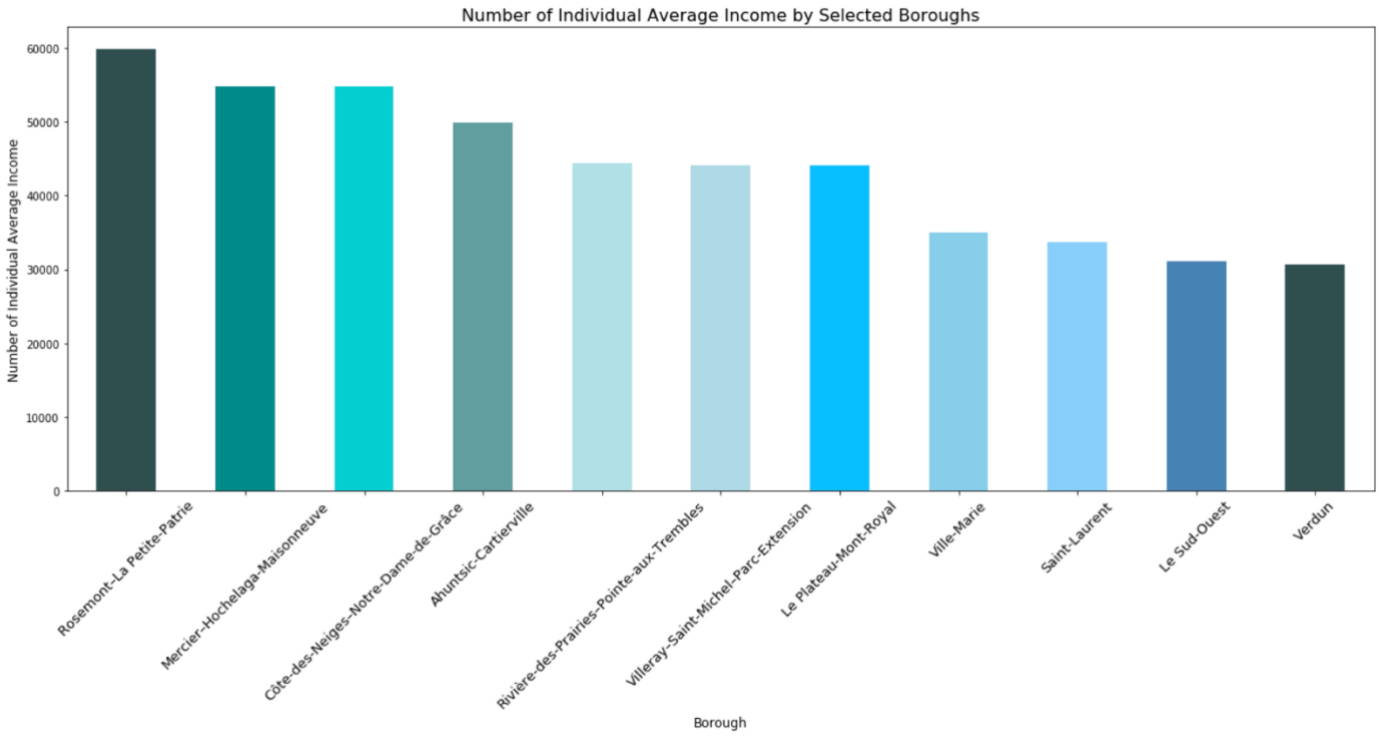
Examine the results for individual income revenue per Montreal boroughs. Let’s Cleaning data for Montreal census income data. Now let's calculate the individual number of average total income persons per borough. We can see people with more than 30,000 $ annually could be categorized as average income.

	Borough	Total Average income
0	Ahuntsic-Cartierville	49820
1	Anjou	17205
2	Baie-D'Urfé	1775
3	Beaconsfield	9170
4	Côte-Saint-Luc	12780
5	Côte-des-Neiges–Notre-Dame-de-Grâce	54745
6	Dollard-Des Ormeaux	20025
7	Dorval	8900
8	Hampstead	3385

Merging Datasets¶

Now, we can merge Montreal data with census data. Let's find boroughs with more than average number total income compared in contrast with other boroughs. We selected these boroughs because there are a greater number of people with more than average income that purchasing Coffee is not an issue for them.

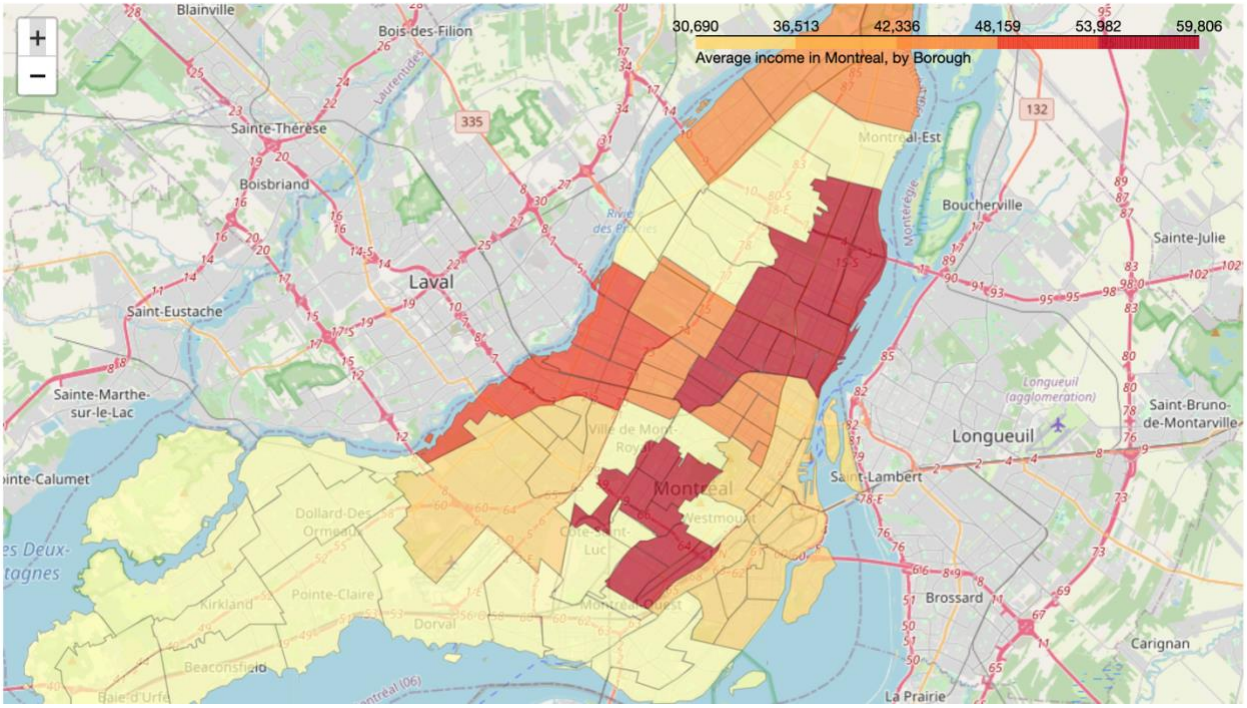
From 33 boroughs in Montreal, there are 11 boroughs with more than average number of people with average income. We focus on these boroughs with inside neighborhoods because there are more people with average income that interested to be potential customer for coffee shop. Now let's get the coordinates of each of our selected neighborhoods. It looks like **geopy** had them all over the place. Almost coordinates for borough seem to be right, but let's plug in values from Google Maps for the rest and clean up the coordinates a little.



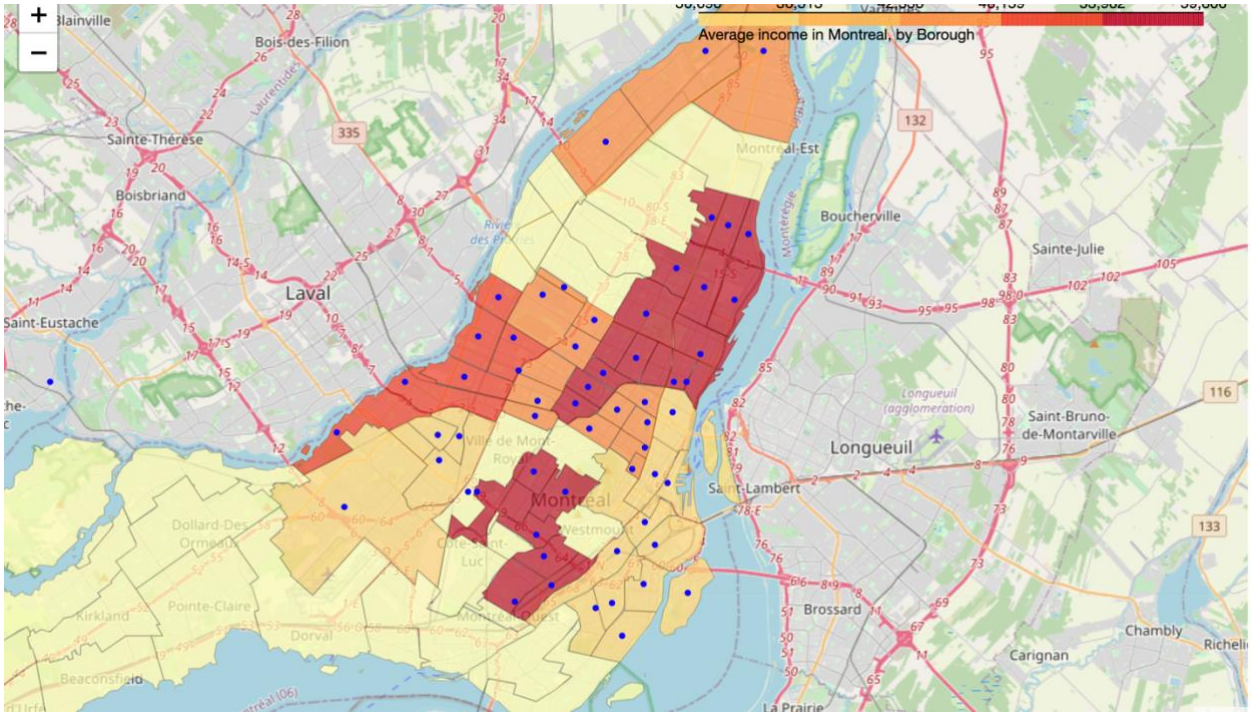
MAP Montreal District

First, Montreal district and show income average number per borough. If Borough has less than about 36,000 individuals with average income we will not explore that borough because there might be fewer potential customers in that region.

Therefore, in our investigation we focus just on 11 boroughs instead of 33 boroughs.



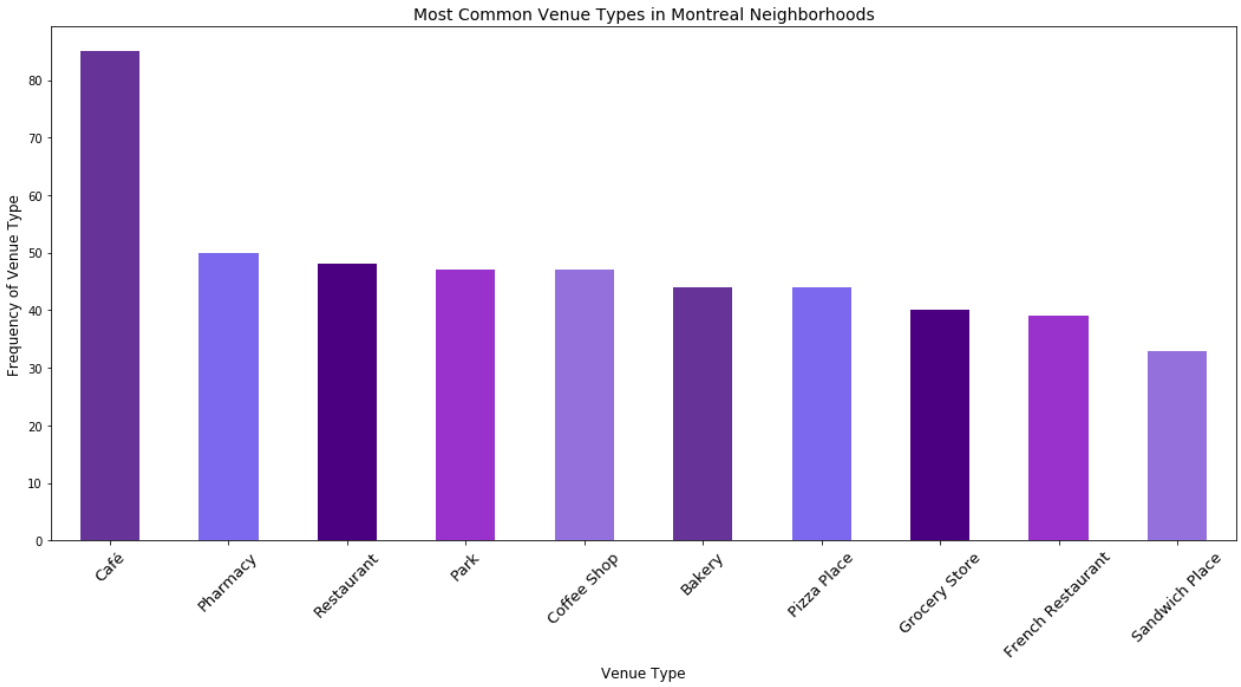
Now, we show the label of boroughs with more than average number of people in contrast with total.



Foursquare Data Analysis¶

This function helps us to search the most popular venues within an 800 meters radius of our neighborhoods. We found, there are in 249 unique categories in these 11 boroughs. More frequent venue categories are listed below:

	Venue Category	Frequency
39	Café	85
176	Pharmacy	50
195	Restaurant	48
170	Park	47
51	Coffee Shop	47
14	Bakery	44
178	Pizza Place	44
103	Grocery Store	40
92	French Restaurant	39
199	Sandwich Place	33



It looks like Cafe and Coffee shops are the most common venue type, followed by Pharmacy, Park, Bakery, and some restaurants.

The most common Venues

Let's do some one hot encoding to further analyze our results. We found the frequency of the top 5 types of venues for each neighborhood to explore the frequency of Coffee shop in each neighborhood.

	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
0	Beaurivage	Pharmacy	Fast Food Restaurant	Liquor Store	Pizza Place	Diner	Park	Baseball Field	Supermarket
1	Bois-Francs	Coffee Shop	Ski Area	Home Service	Restaurant	Women's Store	Donut Shop	Dumpling Restaurant	Eastern European Restaurant
2	Cartierville	Construction & Landscaping	Park	Tailor Shop	Dongbei Restaurant	Fish & Chips Shop	Filipino Restaurant	Fast Food Restaurant	Farmers Market
3	Chameran/Montpellier	Middle Eastern Restaurant	Train Station	Liquor Store	Falafel Restaurant	Pharmacy	Cosmetics Shop	Vietnamese Restaurant	Filipino Restaurant
4	Crémazie	Coffee Shop	Café	Mediterranean Restaurant	Sushi Restaurant	Park	Pub	Optical Shop	Restaurant

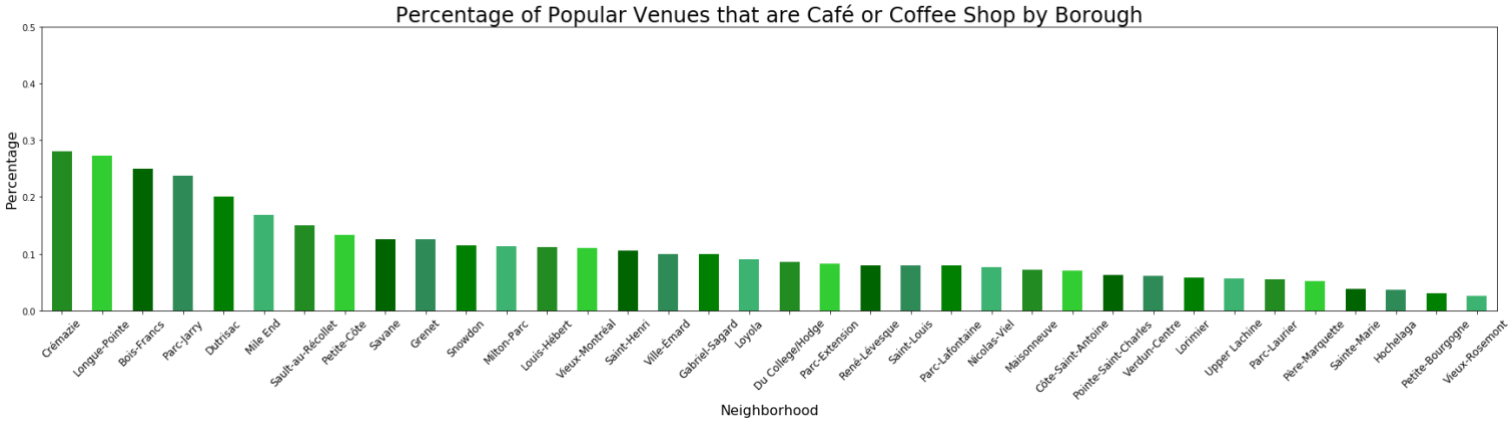
Neighborhood Clustering¶

Let's use K means clustering to cluster our neighborhoods for more insights. Now create a data frame that includes the cluster as well as the top 10 venues for each neighborhood.

	Cluster Labels	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 M Com Venu
0	1	Beaurivage	Pharmacy	Fast Food Restaurant	Liquor Store	Pizza Place	Diner	Park	Baseball Field	Supe
1	1	Bois-Francs	Coffee Shop	Ski Area	Home Service	Restaurant	Women's Store	Donut Shop	Dumpling Restaurant	East Euro Rest
2	2	Cartierville	Construction & Landscaping	Park	Tailor Shop	Dongbei Restaurant	Fish & Chips Shop	Filipino Restaurant	Fast Food Restaurant	Farm Mark
3	1	Chameran/Montpellier	Middle Eastern Restaurant	Train Station	Liquor Store	Falafel Restaurant	Pharmacy	Cosmetics Shop	Vietnamese Restaurant	Filipi Rest
4	1	Crémazie	Coffee Shop	Café	Mediterranean Restaurant	Sushi Restaurant	Park	Pub	Optical Shop	Rest

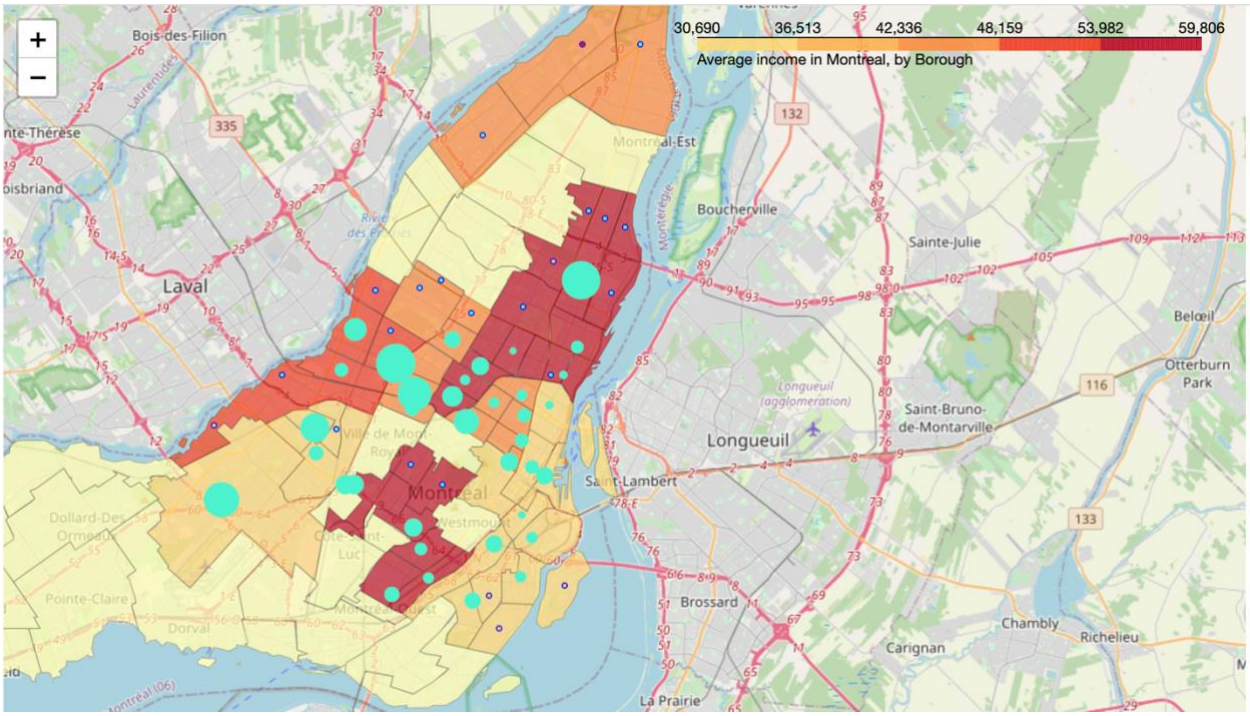
Coffee shop venue frequency per neighborhood

Following figure show the percentage of popular venues in each neighborhood that are Café or Coffee Shop. this will help us see which neighborhoods are more popular.



Map Clustered Neighborhoods¶

Let's pull the coordinates for Montreal and build a map.



Clusters Analysis

Cluster 1

There are two neighborhoods located in cluster 1

	Neighborhood	Borough	Cluster Labels	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
0	Rivière-des-Prairies	Rivière-des-Prairies–Pointe-aux-Trembles	0	Bus Stop	Women's Store	Donut Shop	Flower Shop	Fish & Chips Shop	Filipino Restaurant	Fast Food Restaurant	Farmers Market
17	Saint-Édouard	Rosemont–La Petite-Patrie	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Cluster 2

The majority of neighborhoods with popularity in Coffee shop venue are located in this cluster

57	Bois-Francs	Saint-Laurent	1	Coffee Shop	Ski Area	Home Service	Restaurant	Women's Store	Donut Shop
58	Du College/Hodge	Saint-Laurent	1	Sandwich Place	Pharmacy	Pizza Place	Sushi Restaurant	Italian Restaurant	Fast Fo Restau
60	Verdun-Centre	Verdun	1	Grocery Store	Restaurant	Pizza Place	Café	Bakery	Italian Restau

Cluster 3

There are 5 neighborhoods are located in this cluster

	Neighborhood	**Borough**	**Cluster Labels**	**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 M Com Venu**
8	Louis-Riel	Mercier–Hochelaga-Maisonneuve	2	Skating Rink	Park	Women's Store	Dog Run	Filipino Restaurant	Fast Food Restaurant	Farm Mark
31	Cartierville	Ahuntsic-Cartierville	2	Construction & Landscaping	Park	Tailor Shop	Dongbei Restaurant	Fish & Chips Shop	Filipino Restaurant	Fast Resta
41	Côte-Saint-Paul	Le Sud-Ouest	2	Park	Chinese Restaurant	Ice Cream Shop	Ethiopian Restaurant	Donut Shop	Dumpling Restaurant	Easte Europ Resta
59	Desmarchais-Crawford	Verdun	2	Park	Pool	Botanical Garden	Women's Store	Ethiopian Restaurant	Dumpling Restaurant	Easte Europ Resta
61	Ile-des-Soeurs	Verdun	2	Park	Convenience Store	Playground	Mediterranean Restaurant	Dongbei Restaurant	Filipino Restaurant	Fast Resta

4. Results and Discussion¶

We have pulled data on census individual average income for each borough in Montreal and used this information to narrow down boroughs' options from 33 boroughs to 11 boroughs with 61 neighborhoods. Our analysis has informed us that: Café & Coffee Shop, Restaurant, Pharmacy, Park, Bakery, Pizza Place, French Restaurant, Grocery Store, Sandwich Place are the most common venues in our 11 preferred boroughs.

Clustering neighborhoods based on their most popular venues grouped all 11 preferred boroughs with majority of their neighborhoods categorized into cluster 2. There are eleven neighborhoods in boroughs such Rivière-des-Prairies– Pointe-aux-Trembles, Mercier–Hochelaga-Maisonneuve, Rosemont–La Petite-Patrie, Ahuntsic-Cartierville, Le SudOuest, Saint-Laurent, and verdun are categorized into Cluster 1 and 3 that Café & Coffee Shop is not popular venue in these neighborhoods. There are Places such in Mercier–Hochelaga-Maisonneuve, Villeray–Saint-Michel–Parc-Extension, Saint-Laurent, Le Plateau-Mont-Royal, Ville-Marie have majority Café & Coffee shops as popular venues, whereas most of the popular venues in cluster 1 such as Dutrisac/Saint-Laurent, Louis-Riel/Mercier–Hochelaga-Maisonneuve,SaintSulpice/Ahuntsic-Cartierville, Côte-Saint-Paul/Le Sud-Ouest, Café & Coffee shop is not popular venue, but locations like Park, Gym, and restaurant are more popular.

Based on this analysis, Longue-Pointe/Mercier–Hochelaga-Maisonneuve, Parc-Jarry and Crémazie/ Villeray–SaintMichel–Parc-Extension, Bois-Francs/Saint-Laurent, and Mile End / Le Plateau-Mont-Royal are hot places for Café & Coffee Shops and may not suitable to open a new one. Saint-Laurent, Rosemont–La Petite-Patrie seems to offer a good balance between foot traffic, popularity for Café & Coffee Shop, and might be rent prices. Le Plateau-Mont-Royal and Ville-Marie seems to be a hot spot for Café & Coffee Shop, but also comes with the high cost of rent.

Loyola / Côte-des-Neiges–Notre-Dame-de-Grâce, René-Lévesque/Ville-Marie, Parc-Extension/Villeray–Saint-Michel– Parc-Extension, Saint-Louis/ Le Plateau-Mont-Royal seem to have similar feel on same clustering but is not hotspot and there is the medium busy neighborhood. Places Rosemont–La Petite-Patrie including four neighborhoods Marie-Victorin, Petite-Côte, Vieux-Rosemont, Louis-Hébert, and Père-Marquette have the most population of average individual income might be better place to open a new Café & Coffee Shop which is popular venue but not much in contrast with other places and has more potential customers too.

Ultimately, the optimal Café & Coffee shop spot depends on what type of Coffee Shop you would like to open. An upscale and trendy Coffee Shop might fare better against competition in an expensive and bustling area like Loyola/Côte-des-Neiges–Notre-Dame-de-Grâce, whereas a university may be the spot in an area, which likely receives most of its foot traffic exclusively from its residents.

A major drawback of this analysis is that the clustering was completely based on Foursquare' s data for popular venues. There are other ways to assess popularity of neighborhoods and the spots inside them, venue popularity is just one of them. It may also be helpful to look exclusively at Café & Coffee in an area, how many there are, and how popular they are on weekdays and weekends.

5. Conclusion

We have executed a project using common python libraries to manipulate data sets, Foursquare API, Montreal Open Data portal, StatCan portal to explore the Montreal Agglomeration, census income data and find popular venues of Montreal, and Folium leaflet map to cluster and segment neighborhoods. This project could be further developed with more census data such as rent price, traffic information, and crime data for each borough for more accurate investigation and have better decision.

This was just one use-case. These analytical tools open a world of possibilities for strategic decision making across the various realms of business.