Capstone Project: Battle of the neighborhoods

Rahul Pramod Rote

Where to buy a house in Montreal?

Introduction

I currently reside in Montreal, Quebec, Canada. With more than 4 million inhabitants in its urban area, Montreal is the most populous city in the Canadian province of Quebec, and the second most populous in Canada, after Toronto [1]. The economy of Montreal has been booming in the recent times, especially in the tech sector and biotechnology. This has led to a steep rise in the number of people moving to Montreal, with as many as 179,000 immigrants moving to the metropolitan area between 2011-2016 (Census data). Because of a long history and being culturally distinct from the rest of Canada, Montreal has developed into a multicultural city with distinct neighborhoods. For someone wanting to settle in the city by buying a house, the choices can seem overwhelming. Thus, the aim of my project is to come up with a solution for this problem using the data available online and applying the data science and machine learning techniques I have learned in this course.

Montreal's housing market is considered to be one of the most affordable among major Canadian cities and knowing where to buy a house would maximize the return on the investment. The decision to choose where to buy the house would depend on two factors. First, the price of the house has to be affordable. Secondly, the neighborhood should have a wide selection of venues and restaurants. An ideal house would thus be the one from a neighbourhood with relatively lower prices, while at the same time, with a broader choice of venues and restaurants around.

Problem and interested audience

While Montreal offers a wide selection of neighborhoods and housing, the problem becomes choosing the house in a neighborhood that will strike a perfect balance between affordability and social life. The results from my analysis aim to help the people who are new to the city and are looking to buy property.

Data Description

I obtained coordinates of the neighborhoods of Montreal and their boundaries from the City of Montreal website. The Json file with this data will be used to generate the maps and to segment the neighborhoods so that they can be explored [2].

I will use the Foursquare API to explore the neighborhoods and to obtain the list of the venues in them. This data will be crucial in creating clusters of neighborhoods.

---

[1] https://en.wikipedia.org/wiki/Montreal
[2] https://github.com/blackmad/neighborhoods/blob/master/gn-montreal.geojson

The data for housing prices by neighborhoods in Montreal is available online. This data will be used to cluster the neighborhoods by average price of a home [3].

Methodology

Our aim in this project is to identify the neighborhoods with low property prices and with a wide selection of venues. I restricted the analysis to the limits of the Municipality of Montreal.

First, I downloaded the geojson file that was used to generate the map of Montreal with the neighborhoods of the city. Then, using the Foursquare location data for each neighborhood, I found the 10 most common venues in those neighborhoods. I applied the unsupervised machine learning algorithm (k-means clustering) on this dataset, with the purpose of clustering neighborhoods with similar characteristics.

Secondly, I created a dataframe containing average price of a house for a family of 4 by neighborhood. I cleaned up the data from the housing dataframe can be merged with the clustering dataframe. The resulting dataframe then was used to generate a choropleth map of the city of Montreal by house prices, overlaid with the cluster labels.

The final step was to observe the clusters and the house prices to identify whether the clusters contains neighborhoods with similar characteristics with affordable house prices.

Results

Generating a map of Montreal showing the center points of the neighborhoods

The Montreal geojson file contains coordinates for 19 neighborhoods of Montreal. Using this file, I generated a *pandas* dataframe with the names and coordinates of the city, named "neighborhoods".

```
In [12]: neighborhoods.head()
```

Out[12]:

	Neighborhood	Latitude	Longitude
0	Ahuntsic-Cartierville	45.576388	-73.662712
1	Pierrefonds--Roxboro	45.481054	-73.867003
2	Ahuntsic-Cartierville	45.543146	-73.680433
3	Cote-des-Neiges--Notre-Dame-de-Grace	45.484893	-73.631757
4	Outremont	45.515644	-73.608670

[3] https://news.shupilov.com/blog/montreal-boroughs-ranked-according-to-5-year-appreciation/

Figure 1: Dataframe containing the names and coordinates of the neighborhoods
Using this dataframe as an input and the Folium library code, I generated a map centered around Montreal showing the center points of neighborhoods. This map served as a template for the further data visualization using maps.

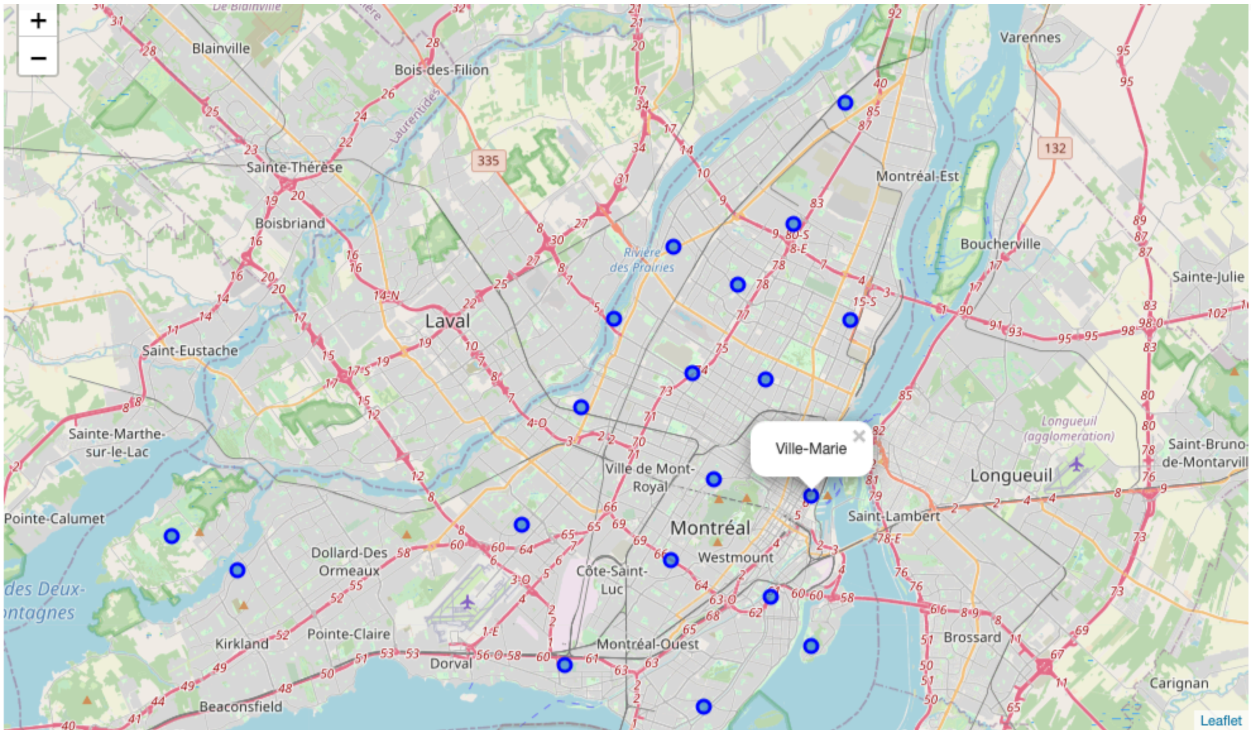

Figure 2: Map centered around Montreal showing the center points of neighborhoods as blue circles.

Obtaining lists of venues in the neighborhoods of Montreal using Foursqure API

Using my Client ID and Client secret details in the Foursquare API, I generated the URL that will be used to query the API with a GET request. As a test, I explored the neighborhood of Ahuntsic-Cartierville by obtaining the list of venues within 500 meters.

	name	categories	lat	lng
0	Parc-nature de l'Île-de-la-Visitation	Park	45.575632	-73.658867
1	Bistro Des Moulins	Café	45.574878	-73.661219
2	Cité historia, Musé Du sault-au-Récolet, Meunier	History Museum	45.574858	-73.661050
3	Mahmoud Sidibe	Business Service	45.573212	-73.661367

```
print('{} venues were returned by Foursquare.'.format(nearby_venues.shape[0]))
```

```
4 venues were returned by Foursquare.
```

Figure 3: List of venues within 500 meters of Ahuntsic-Cartierville in Foursquare

To generate a dataframe with venues from all the neighborhoods of Montreal, I looped through the “neighborhoods” dataframe and extracted all possible venues with Foursquare. There were 115 unique categories of venues. Plateau-Mont-Royal, Ville-Marie (downtown), and Côte-Des-Neiges-Notre-Dame-De-Grâce were the neighborhoods with the highest numbers of venues.

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Ahuntsic-Cartierville	10	10	10	10	10	10
Anjou	5	5	5	5	5	5
Cote-des-Neiges--Notre-Dame-de-Grace	37	37	37	37	37	37
L'Ile-Bizard--Sainte-Genevieve	1	1	1	1	1	1
LaSalle	16	16	16	16	16	16
Lachine	4	4	4	4	4	4
Mercier-Hochelaga-Maisonneuve	4	4	4	4	4	4
Montreal-Nord	6	6	6	6	6	6
Outremont	12	12	12	12	12	12
Pierrefonds--Roxboro	6	6	6	6	6	6
Plateau-Mont-Royal	100	100	100	100	100	100
Pointe-aux-Trembles-Rivieres-des-Prairies	2	2	2	2	2	2
Rosemont--La-Petite-Patrie	4	4	4	4	4	4
Saint-Laurent	9	9	9	9	9	9
St-Leonard	7	7	7	7	7	7
Sud-Ouest	4	4	4	4	4	4
Verdun--Ile-des-Soeurs	4	4	4	4	4	4
Ville-Marie	40	40	40	40	40	40
Villeray-Saint-Michel-Parc-Extension	8	8	8	8	8	8

Figure 4: List of neighborhoods with counts for the number of venues

Clustering analysis

Processing the dataframe for clustering

For the venues data to be processed by the machine learning algorithm, I generated dummy variables from the data through one-hot encoding. The data now has 279 rows and 116 columns.

	Neighborhood	Afghan Restaurant	American Restaurant	Asian Restaurant	BBQ Joint	Bagel Shop	Bakery	Bank	Bar	Baseball Field	Beer Bar	Bistro	Bookstore	Breakfast Spot	Burger Joint	Burrito Place	Busine Servi
0	Ahuntsic-Cartierville	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1	Ahuntsic-Cartierville	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
2	Ahuntsic-Cartierville	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
3	Ahuntsic-Cartierville	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
4	Pierrefonds--Roxboro	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

```
montreal_onehot.shape
```

```
(279, 116)
```

Figure 5: One hot encoding of the neighborhoods dataframe

To condense the data into the numbers of rows equal to the number of neighborhoods, I grouped the rows by neighborhood and by taking the mean of the frequency of occurrence of each category.

	Neighborhood	Afghan Restaurant	American Restaurant	Asian Restaurant	BBQ Joint	Bagel Shop	Bakery	Bank	Bar	Baseball Field	Beer Bar	Bistro	Bookstore	Breakfast Spot	Burger Joint
0	Ahuntsic-Cartierville	0.00	0.000000	0.000000	0.0	0.00	0.000000	0.000000	0.0000	0.00	0.00	0.000	0.00	0.100000	0.00
1	Anjou	0.00	0.000000	0.000000	0.2	0.00	0.000000	0.000000	0.0000	0.00	0.00	0.000	0.00	0.000000	0.00
2	Cote-des-Neiges--Notre-Dame-de-Grace	0.00	0.000000	0.027027	0.0	0.00	0.027027	0.027027	0.0000	0.00	0.00	0.000	0.00	0.000000	0.00
3	L'Ile-Bizard--Sainte-Genevieve	0.00	0.000000	0.000000	0.0	0.00	0.000000	0.000000	0.0000	0.00	0.00	0.000	0.00	0.000000	0.00
4	LaSalle	0.00	0.000000	0.000000	0.0	0.00	0.000000	0.000000	0.0625	0.00	0.00	0.000	0.00	0.000000	0.00
5	Lachine	0.00	0.000000	0.000000	0.0	0.00	0.000000	0.000000	0.0000	0.00	0.00	0.000	0.00	0.000000	0.00

Figure 6: Dataframe showing the mean of the frequency of occurrence of each category of venues by neighborhood

The dataframe was further condensed by keeping only the top 10 venues for 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	9th Most Common Venue	10th Most Common Venue
0	Ahuntsic-Cartierville	Hockey Arena	Breakfast Spot	Italian Restaurant	Liquor Store	Middle Eastern Restaurant	Park	Café	Business Service	History Museum	Train Station
1	Anjou	BBQ Joint	Convenience Store	Construction & Landscaping	Motorsports Shop	Sporting Goods Shop	Women's Store	Furniture / Home Store	Donut Shop	Drugstore	Farmers Market
2	Cote-des-Neiges--Notre-Dame-de-Grace	Coffee Shop	Pharmacy	Deli / Bodega	Fast Food Restaurant	Italian Restaurant	Liquor Store	Sandwich Place	Chinese Restaurant	Caribbean Restaurant	Café
3	L'Ile-Bizard--Sainte-Genevieve	Golf Course	Women's Store	Gaming Cafe	Discount Store	Donut Shop	Drugstore	Farmers Market	Fast Food Restaurant	Flower Shop	Food Truck
4	LaSalle	Department Store	Grocery Store	Pharmacy	Supermarket	Fast Food Restaurant	Gym / Fitness Center	Discount Store	Bar	Shopping Mall	Movie Theater

Figure 7: List of neighborhoods with 10 most common venue cateogories

Finding the optimal k for k-means clustering

To find the optimal numbers of clusters the *k*-means algorithm, I used the silhouette method. It is an iterative process that measures the similarity of a point is to the others in its own clusters compared to other clusters.

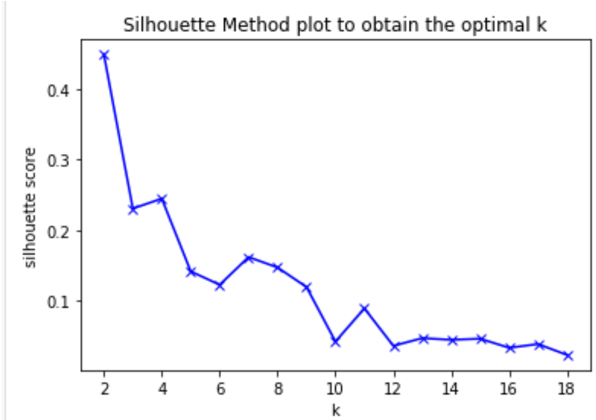

Figure 8: Graph showing the silhouette score for each of the k value for the neighborhood clustering dataset.

From the chart in Fig.8, the maximum silhouette score was at k=2. However, having only two clusters would not give a meaningful classification of the neighborhoods in Montreal. There was another peak in silhouette score at k=4. Thus, I set the number of clusters as 4 for further analysis.

	Neighborhood	Latitude	Longitude	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	9th Most Common Venue	C
0	Ahuntsic-Cartierville	45.576388	-73.662712	1	Hockey Arena	Breakfast Spot	Italian Restaurant	Liquor Store	Middle Eastern Restaurant	Park	Café	Business Service	History Museum	
1	Pierrefonds--Roxboro	45.481054	-73.867003	1	Pizza Place	Convenience Store	Performing Arts Venue	Grocery Store	Fast Food Restaurant	Diner	Women's Store	Frozen Yogurt Shop	Donut Shop	D
2	Ahuntsic-Cartierville	45.543146	-73.680433	1	Hockey Arena	Breakfast Spot	Italian Restaurant	Liquor Store	Middle Eastern Restaurant	Park	Café	Business Service	History Museum	
3	Cote-des-Neiges--Notre-Dame-de-Grace	45.484893	-73.631757	1	Coffee Shop	Pharmacy	Deli / Bodega	Fast Food Restaurant	Italian Restaurant	Liquor Store	Sandwich Place	Chinese Restaurant	Caribbean Restaurant	
4	Outremont	45.515644	-73.608670	1	Restaurant	Park	Bakery	Pizza Place	Metro Station	Italian Restaurant	Ice Cream Shop	Theater	American Restaurant	

Figure 9: Dataframe showing the first 5 entries from the list of neighborhoods with cluster labels.

The dataframe from Fig.9 was then used to generate a map of Montreal with centres of neighborhoods showing the cluster labels (Fig.10). While a large number of neighborhoods were contained within cluster 1, the remaining neighborhoods formed distinct clusters with one or two neighborhoods each. For example, the neighborhood of Pointe-aux-Trembles-Rivieres-des-Prairies was the sole constituent in cluster 0.

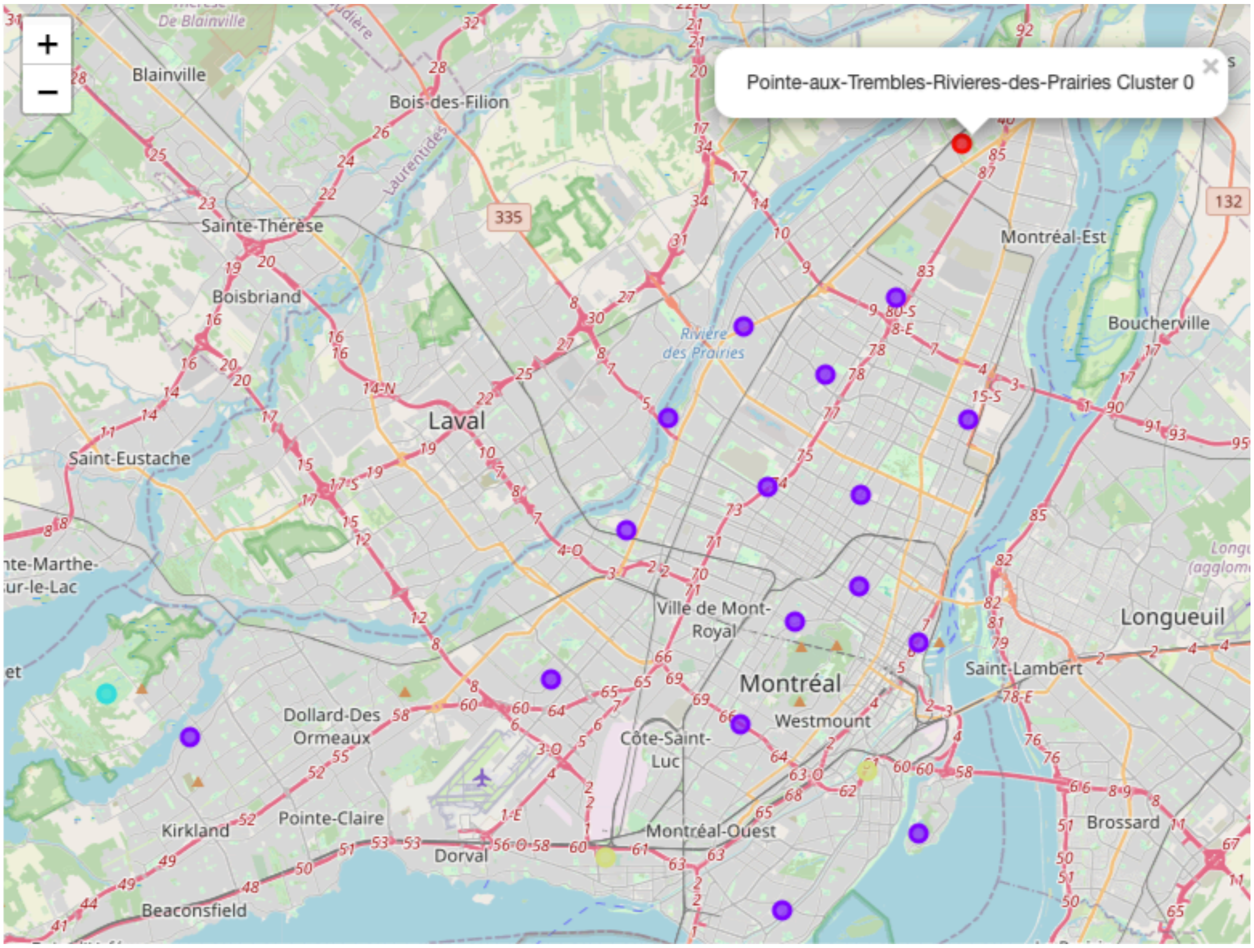

Figure 10: Map of Montreal showing the neighborhoods with their respective cluster labels

Processing the housing price data

The house price data contained the list of neighborhoods with average price of a house for a family of 4, and 5-year price growth for that neighborhood. This data was loaded into a pandas dataframe, sorted alphabetically, and merged with the clustering dataframe. The average price data was converted into float for further processing.

	Neighborhood	Average Price	5 Year Price Growth	Latitude	Longitude	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
0	Ahuntsic-Cartierville	565560.0	19.68%	45.576388	-73.662712	1	Hockey Arena	Breakfast Spot	Italian Restaurant	Liquor Store	Middle Eastern Restaurant	Park	Café
9	Anjou	396957.0	4.66%	45.612503	-73.565355	1	BBQ Joint	Convenience Store	Construction & Landscaping	Motorsports Shop	Sporting Goods Shop	Women's Store	Furniture / Home Store
3	Cote-des-Neiges--Notre-Dame-de-Grace	744729.0	18.11%	45.484893	-73.631757	1	Coffee Shop	Pharmacy	Deli / Bodega	Fast Food Restaurant	Italian Restaurant	Liquor Store	Sandwich Place
18	L'Ile-Bizard--Sainte-Genevieve	518800.0	NaN	45.493960	-73.903146	2	Golf Course	Women's Store	Gaming Cafe	Discount Store	Donut Shop	Drugstore	Farmers Market
12	LaSalle	447021.0	19.21%	45.428989	-73.613800	1	Department Store	Grocery Store	Pharmacy	Supermarket	Fast Food Restaurant	Gym / Fitness Center	Discount Store

Figure 11: Dataframe containing the list of neighborhoods with average price of a house, 5 years price growth, coordinates, cluster labels, and 10 most common venues.

The data from Fig.11 was used to generate a choropleth map of Montreal showing neighborhoods by house prices with cluster labels.

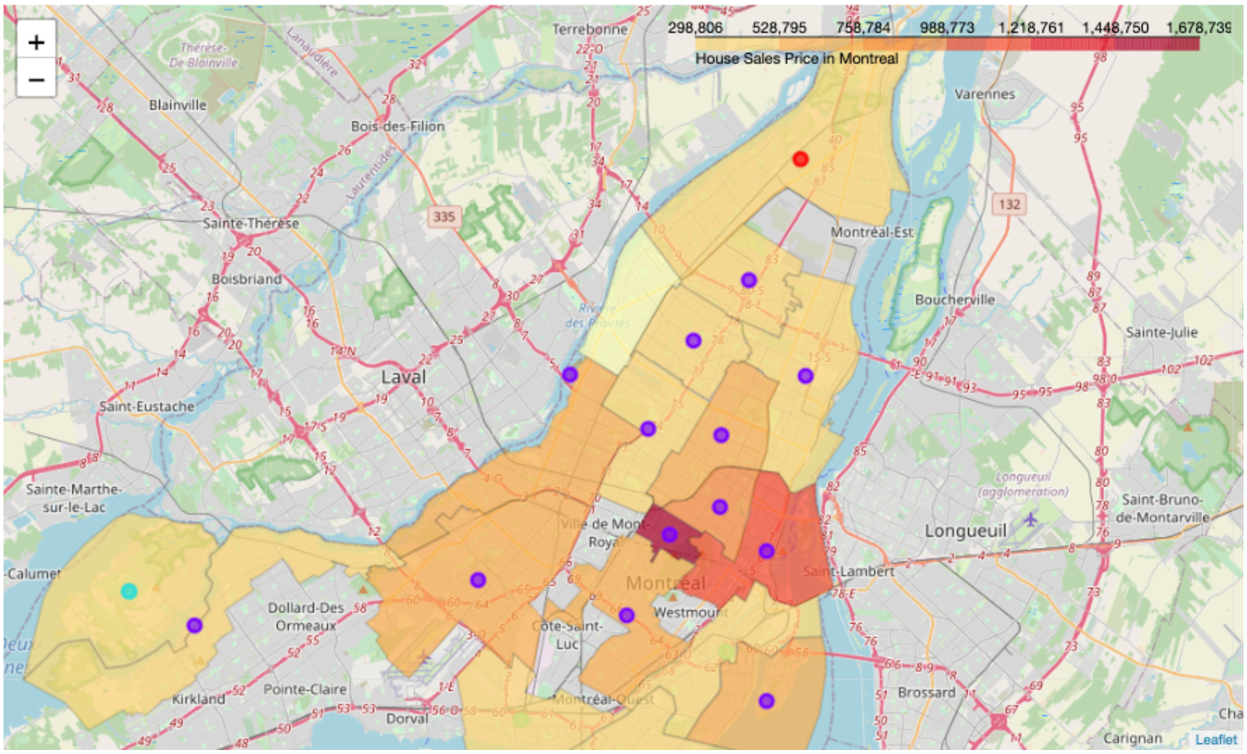

Figure 12: Choropleth map of Montreal showing the house prices in neighborhoods of Montreal. The colors are according the scale on the top right of the image. Circles indicate the clusters.

Examining the clusters

Based on the map in Fig. 12, I labeled the clusters according to their geographical position within the city and/or the most common venue category, as follows:

1) **Eastern end cluster**

The Eastern end cluster located at the Eastern tip of the island of Montreal contains a single neighborhood of Pointe-aux-Trembles-Rivieres-des-Prairies which is relatively larger in area but smaller by population. This neighborhood seems to be well connected to the city with commuter rail but has fewer choices for venues and restaurants. The average house prices are some of the lowest on the island, but the 5-year price growth is also the lowest. Thus, while it may be a good place to buy a house for someone who prefers the peace and quiet of a suburb, the return on investment may not be optimal.

	Neighborhood	Average Price	5 Year Price Growth	Latitude	Longitude	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	9th Mo Comm Ven
5	Pointe-aux-Trembles-Rivieres-des-Prairies	312335.0	8.47%	45.658857	-73.537199	0	Tram Station	Train Station	Women's Store	Gaming Cafe	Diner	Discount Store	Donut Shop	Drugstore	Farm Mar

Figure 13: The Eastern end cluster containing only the neighborhood of Pointe-aux-Trembles-Rivieres-des-Prairies

2) The city proper cluster

The city proper cluster the largest cluster, containing 14 neighborhoods in total, with each of them known to have distinct cultural flavors. However, they all seem to share many physical characteristics, such as a broad choice of venues, such as restaurants, cafés, parks, grocery and convenience stores, and pizza places. The house prices, however, show a broad range. Some of the neighborhoods, such as Ville Marie (downtown) and Outremont are some of the most expensive in the city. However, we can also see that there are a few neighborhoods with affordable house prices and a strong 5 year price growth.

Based on the analysis of venues and prices, one of the best neighborhoods to buy a house appears to be Rosemont--La-Petite-Patrie, with relatively affordable housing, a wide selection of restaurants and venues, and the strongest 5 year price growth in Montreal, making for a much larger potential return on investment.

Buying a house in a few other neighborhoods in this also can be a good investment, most notably, Verdun--Ile-des-Soeurs, LaSalle, and Villeray-Saint-Michel-Parc-Extension. These three neighborhoods have relatively lower house prices and a stronger growth.

Overall, this cluster contains some prime spots for buying a house for people new to the market in Montreal.

	Neighborhood	Average Price	5 Year Price Growth	Latitude	Longitude	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
0	Ahuntsic-Cartierville	565560.0	19.68%	45.576388	-73.662712	1	Hockey Arena	Breakfast Spot	Italian Restaurant	Liquor Store	Middle Eastern Restaurant	Park	Café
9	Anjou	396957.0	4.66%	45.612503	-73.565355	1	BBQ Joint	Convenience Store	Construction & Landscaping	Motorsports Shop	Sporting Goods Shop	Women's Store	Furniture / Home Store
3	Cote-des-Neiges--Notre-Dame-de-Grace	744729.0	18.11%	45.484893	-73.631757	1	Coffee Shop	Pharmacy	Deli / Bodega	Fast Food Restaurant	Italian Restaurant	Liquor Store	Sandwich Place
12	LaSalle	447021.0	19.21%	45.428989	-73.613800	1	Department Store	Grocery Store	Pharmacy	Supermarket	Fast Food Restaurant	Gym / Fitness Center	Discount Store
13	Mercier-Hochelaga-Maisonneuve	359739.0	14.10%	45.575953	-73.534009	1	Gym	Baseball Field	Coffee Shop	Pool	Women's Store	Furniture / Home Store	Discount Store
4	Outremont	1665210.0	NaN	45.515644	-73.608670	1	Restaurant	Park	Bakery	Pizza Place	Metro Station	Italian Restaurant	Ice Cream Shop
1	Pierrefonds--Roxboro	376115.0	10.05%	45.481054	-73.867003	1	Pizza Place	Convenience Store	Performing Arts Venue	Grocery Store	Fast Food Restaurant	Diner	Women's Store
8	Plateau-Mont-Royal	838636.0	19.59%	45.526094	-73.580749	1	Café	Bookstore	Bakery	Coffee Shop	Cocktail Bar	Creperie	Pizza Place
6	Rosemont--La-Petite-Patrie	585431.0	43.33%	45.553353	-73.580415	1	Sushi Restaurant	Sandwich Place	Grocery Store	Flower Shop	Furniture / Home Store	Diner	Discount Store
14	Saint-Laurent	606828.0	18.40%	45.498161	-73.712855	1	Women's Store	Bank	Breakfast Spot	Coffee Shop	Thai Restaurant	Furniture / Home Store	Grocery Store
15	St-Leonard	453933.0	NaN	45.589375	-73.595435	1	Hockey Arena	Discount Store	Drugstore	Restaurant	Italian Restaurant	Convenience Store	Dessert Shop
19	Verdun--Ile-des-Soeurs	536734.0	35.74%	45.451877	-73.555246	1	Gym	Soccer Field	Golf Course	Spa	Women's Store	Frozen Yogurt Shop	Diner
7	Ville-Marie	1018513.0	33.00%	45.509380	-73.555461	1	Café	French Restaurant	Italian Restaurant	Restaurant	Speakeasy	Pizza Place	Plaza
16	Villeray-Saint-Michel-Parc-Extension	390595.0	19.88%	45.555761	-73.620297	1	Fast Food Restaurant	Hardware Store	Pizza Place	Liquor Store	Restaurant	Café	Theater

Figure 14: The city proper cluster containing 14 neighborhoods, with a wide selection of venues, house prices and growth potential.

3) Île Bizard island cluster

The Île Bizard island cluster contains just one neighborhood of l'Île-Bizard--Sainte-Genevieve. A large section of this cluster is occupied by the Bois-de-l'Île-Bizard Nature Park and many golf courses, making it one of the largest green spaces in Montreal. The average house prices are higher than many of the suburbs and the 5 year price growth data was unavaible. This cluster may be a good place to buy a house for those who want to get away from the hustle of the city and closer to nature.

	Neighborhood	Average Price	5 Year Price Growth	Latitude	Longitude	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	9th M Comr Ve
18	L'Ile-Bizard--Sainte-Genevieve	518800.0	NaN	45.49396	-73.903146	2	Golf Course	Women's Store	Gaming Cafe	Discount Store	Donut Shop	Drugstore	Farmers Market	Fast Food Restaurant	Flo S

Figure 15: The Île Bizard island cluster containing only the neighborhood of l'Île-Bizard-Sainte-Genevieve

4) Southwestern cluster

The southwestern cluster contains two of the neighborhoods along the Lachine Canal. Both of them offer a wide selection of restaurants, shops, and venues, along with opportunities for outdoor activies in parks. Both neighborhoods have relatively affordable house prices and some of the strongest 5-year price growth figures in Montreal. These factors make buying a house in either of these neighborhoods a good investment with high returns.

	Neighborhood	Average Price	5 Year Price Growth	Latitude	Longitude	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
10	Lachine	442126.0	26.22%	45.444809	-73.689583	3	Park	Restaurant	Convenience Store	Women's Store	Dessert Shop	Discount Store	Donut Shop	Drugstore
17	Sud-Ouest	525435.0	28.71%	45.470778	-73.577176	3	Park	Construction & Landscaping	Food Truck	Women's Store	Gaming Cafe	Donut Shop	Drugstore	Farmers Market

Figure 16: The Southwestern cluster contains the neighborhoods of Lachine and Sud-Ouest and offer some of the best return on investment among all the neighborhoods in Montreal.

Discussion

My aim was to identify clusters of neighborhoods in the city of Montreal with similar characteristics using machine learning algorithms and compare the average house prices between neighborhoods to recommend places where a customer can buy a house. My analysis using k-means clustering was able to generate 4 clusters of neighborhoods. I generated map of Montreal with these clusters overlaid on top of a map showing neighborhoods by house prices. This map provides a visual reference for my analysis and can also be used by someone else to derive their own conclusions from it.

While this analysis can be a starting point to make a decision about buying a house, further analysis is required to make accurate predictions about the suitability of a neighborhood for investment. Some limitations of this analysis are as follows:

1) The average house price only considers the price of 2-bedroom house for a family of 4, while the area of the house and the amenities are not included in the dataset.
2) The k-means algorithm only resulted in 4 clusters, which was the optimal number of clusters obtained from the silhouette method. While 3 of the clusters were smaller and distinct, the largest cluster, the city proper cluster contained 70% of the neighborhoods. This result complicated the analysis for the cluster because the diversity of neighborhoods and a wide range of house prices.
3) This analysis did not include differences between the neighborhoods based on livability metrics such as crime, education, and infrastructure. These factors have a large influence on decisions regarding buying a house.

Future analysis on this topic should involve taking these factors into account and performing the analysis with a larger number of clusters.

Recommendations

Coupled with the average house prices and the 5-year price growth, I was able to provide a few recommendations for where to buy a house to maximize the return on investment, while being in a neighborhood with many options for dining and activities. Based on these factors in my analysis, the best neighborhoods to buy a house in Montreal seem to be in the boroughs of **Rosemont--La-Petite-Patrie, Sud-Ouest, Verdun--Ile-des-Soeurs, LaSalle, and Villeray-Saint-Michel-Parc-Extension.**

Conclusion

In this study, I generated clusters of neighborhoods in Montreal with the help of Foursquare location data of the venues in the city. This analysis was based on the distribution of venues contained within those neighborhoods. After this, I studied these clusters in conjunction with the data about average house prices and the 5-year price growth in a neighborhood. Based on this analysis, I was able to provide recommendations for a person wanting to buy a house in Montreal in a neighborhood that has a wide selection of venues and restaurants, affordable house prices, and a stronger growth potential. The main purpose of my analysis was thus to maximize the return on investment in the decision-making process for the purchase of a house. Along with this analysis, other factors that should be considered in this decision-making process are crime rates, infrastructure, educational facilities, among others.

In conclusion, while Montreal's real estate market is affordable, more so with the current slump in house prices caused by the Covid-19 pandemic, analysis like this one may help new buyers to make an informed decision.