

Where to find partners for food delivery service

Jihoo Park

June 11, 2020

1. Introduction

The goal of this project is to find a neighborhood in Montreal, where the restaurant business is most active. It is to provide food delivery service. In terms of restaurant business, therefore, it includes not only restaurants but also pubs, gourmet stores, coffee shops, delis and so on, which can be items for food delivery service.

2. Background

This project is for a startup company which launches a mobile application for food delivery service. The question they ask is, “What is the best neighborhood in Montreal to find partners?”

As a launching event, the company plans to open a temporary pop-up stand. They prefer the neighborhood which has many places of restaurant business. Their strategy at the stand is to advertise their service providing information, to visit the venues to encourage the owners as their partners, and to promote their service to the customers in the neighborhood as potential users of their application.

3. Data

The data to find the neighborhood will be obtained by exploring the city. Using Foursquare API, venues and their information will be provided. The information includes names, categories, coordinates, addresses, distances.

The exploration will be performed centering around Montreal downtown, where there is more possibility to have restaurant business spots.

After obtaining the data, it will be clustered by categories. This is to distinguish food related venues from others. Each cluster is expected to give information of each

venue including names, categories, addresses, and cluster numbers. Among them, categories and cluster names will be used to visualize on the map as a pop-up text.

By colors of each clustered group, it will be visible and clear to figure out the types of business on the map.

4. Methodology

(1) Main component

Using API by Foursquare:

To use an API which is provided by Foursquare, an URL was created including credentials, coordinates obtained and radius. With the URL, GET request was sent.

Obtaining data by JSON:

The data was obtained as a JSON file. Essential information was extracted from the file then normalized in a form of dataframe.

Clustered by categories:

The venues were clustered based on the values under the column, categories. The result of clustering and the existing dataframe were merged.

Visualized by Folium:

The clustered venues were visualized on the map using Folium. The map was generated centering around Montreal downtown.

(2) Data exploration

JSON file:

Since the data was obtained in a form of JSON, investigation of the file was necessary to understand it. The structure of data, including hierarchy among keys and values in dictionaries were checked. It was useful to figure out the location of proper data to extract.

Dataframe:

To transform the data from the JSON file to a dataframe, normalization was performed. Then the number of venues was checked. It is necessary to get enough data to obtain the meaningful result. The value of radius in the URL was modified several times to get 100 rows of data in the dataframe.

Map:

When generating the map of Montreal, the coordinates used came from the centric place of downtown: 1 Place Ville Marie. Because the address was in French, however, nothing was provided but an error. When trying to use the translated version, a wrong spot was displayed on the map.

As a second plan, the city name was used to get the coordinates. The spot was located in the center of the city, in downtown. Double checked the location using Google maps and Folium to keep the coordinates as the center of the map.

(3) Data analysis

- **Libraries**

Following libraries were imported for each purpose.

request: To handle the GET request

pandas: To analyze data in a form of dataframes

numpy: To handle data in a vectorized manner

folium: To visualize data on maps

- **URL and result**

Defined Foursquare credential and version.

Get the coordinates of Montreal using the searching key: Montreal, QC

Created an URL with the values and set radius and limit for the result.

Sent GET request then checked the result in the JSON file.

- **Data in dataframe**

Normalized the data from JSON to a dataframe. The columns were filtered.

Created a function that extracts the categories of the venues.

Cleaned column names by keeping only last terms.

- **Visualization**

Generated a map of Montreal with markers.

A red marker was added to the map to mark the center of Montreal.

Venues from the dataframe were added to the map as blue circle markers.

The places were clustered. The updated result was visualized on the map being displayed in the same colors according to the clusters.

(4) Machine learning

Clustering was performed for the venues obtained from the JSON file. K-means was used to cluster them by categories. Assigned the number of clusters as 7 then encoded One hot. After running K-means clustering, added a column for cluster labels. It was added into the existing dataframe.

For visualization of data, a color schema was set to clusters. It means that the color of the spots on the map are different depending on the categories that they belong to. The schema helps to classify the type of venues and to distinguish restaurant business spots from others. This should be helpful for the client to figure out the result at a glance.

5. Results

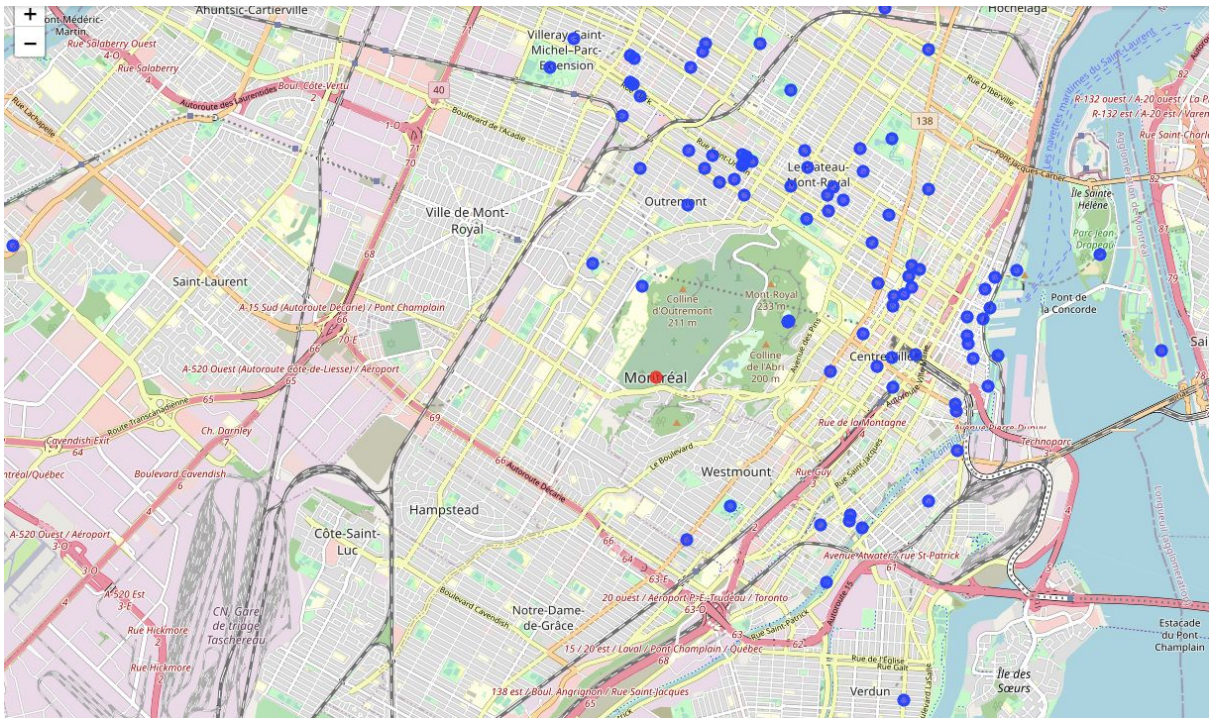
From the JSON file, 100 venues in Montreal and their information were acquired:

```
100 venues were returned.
```

The data was transformed into a dataframe. Through a modification of the table including filtering columns, extracting categories data, and cleaning column names keeping last terms, a visible dataframe was presented:

	name	categories	address	lat	lng	labeledLatLngs	distance	postal
0	Parc du Mont-Royal	Park	1260 Chemin Remembrance	45.504079	-73.587321	[{'label': 'display', 'lat': 45.50407921694641...	1953	H3
1	Westmount Park	Park	Parc Westmount	45.481574	-73.597476	[{'label': 'display', 'lat': 45.48157356036293...	2010	
2	Chalet du Mont-Royal	Historic Site	1196 voie Camillien-Houde	45.503904	-73.587460	[{'label': 'display', 'lat': 45.503904, 'lng':...	1935	H3
3	Want Apothecary	Clothing Store	4960 Sherbrooke Street, West	45.477508	-73.605049	[{'label': 'display', 'lat': 45.47750760210334...	2232	H3
4	Drawn & Quarterly	Bookstore	211 rue Bernard Ouest	45.524748	-73.604614	[{'label': 'display', 'lat': 45.52474756423502...	3097	H2

The venues were visualized on a map using Folium. The red marker presents the center of the city, and the blue ones display the venues:

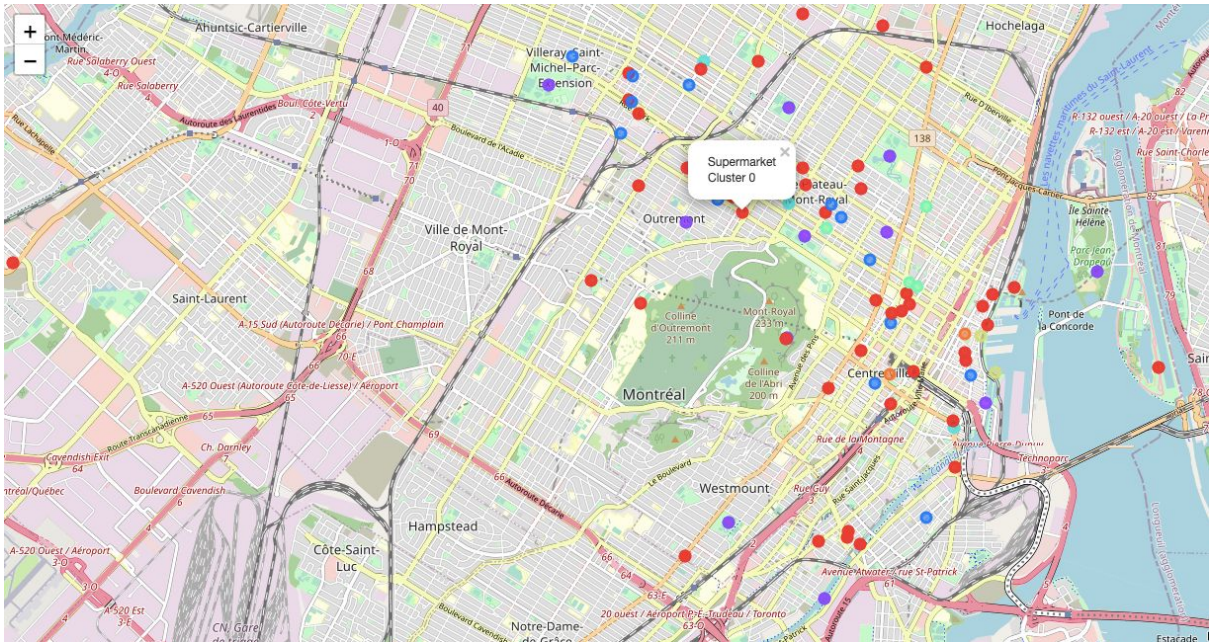


Most spots are scattered in downtown and Le Plateau - Mont-Royal.

As the next step, the spots were clustered using K-means. Now all venues belong to 7 clusters.

	Cluster Labels	categories_Asian Restaurant	categories_BBQ Joint	categories_Bagel Shop	categories_Bakery	categories_Coffee Shop
0	1	0	0	0	0	0
1	1	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0

And the clustered spots are visualized on the map as follows:



Spots are supposed to be presented in the same color by the cluster they belong to. When clicking on the spots, the categories and cluster names are popped up.

6. Discussion

During the investigation of visualization, an issue was found.

On the map, various types of venues were displayed in the same color. It was verified that these various kinds of business were included in the same cluster:

	categories	name	address
2	Historic Site	Chalet du Mont-Royal	1196 voie Camillien-Houde
3	Clothing Store	Want Apothecary	4960 Sherbrooke Street, West
4	Bookstore	Drawn & Quarterly	211 rue Bernard Ouest
7	Grocery Store	Milano Fruiterie	6862 boul. Saint-Laurent
9	Market	Marché Atwater	138 ave. Atwater
10	Farmers Market	Marché Jean-Talon	7070 avenue Henri-Julien
11	Historic Site	Vieux-Port de Montréal / Old Port of Montreal ...	Vieux-Port de Montréal
12	Performing Arts Venue	Place des Arts	175 rue Sainte-Catherine Ouest
14	Plaza	Place des Festivals	Quartier des Spectacles
15	Canal	Canal-de-Lachine	Lachine canal / Canal de Lachine
18	Deli / Bodega	Boucherie Lawrence	5237 St Laurent

This issue still occurred when increasing the number of clusters. Only the first cluster had the problem. The issue seemed less messy when the spots were classified into

45 clusters. However, 45 clusters for 100 spots does not seem to be meaningful, nor for visualization.

Also, there was another issue even when clustering them into 45 tables: unorganized category names. For example, “Mexican Restaurant” was in a different cluster than “Restaurant”. And “Coffee Shop” did not belong to the same cluster with “Café”.

Moreover, the client might want to distinguish the restaurant business from other types of venues on the map at a glance. Organization of the categories would be useful in this case.

The goal of organization was set to group them in 7 general categories. To perform this, a new column, “NewCat” was added into the dataframe. Then the values under the categories were assigned according to the conditions. For example, if a value included a word, “Restaurant” or “coffee”, they belonged to a new category, “Food & Drinks”.

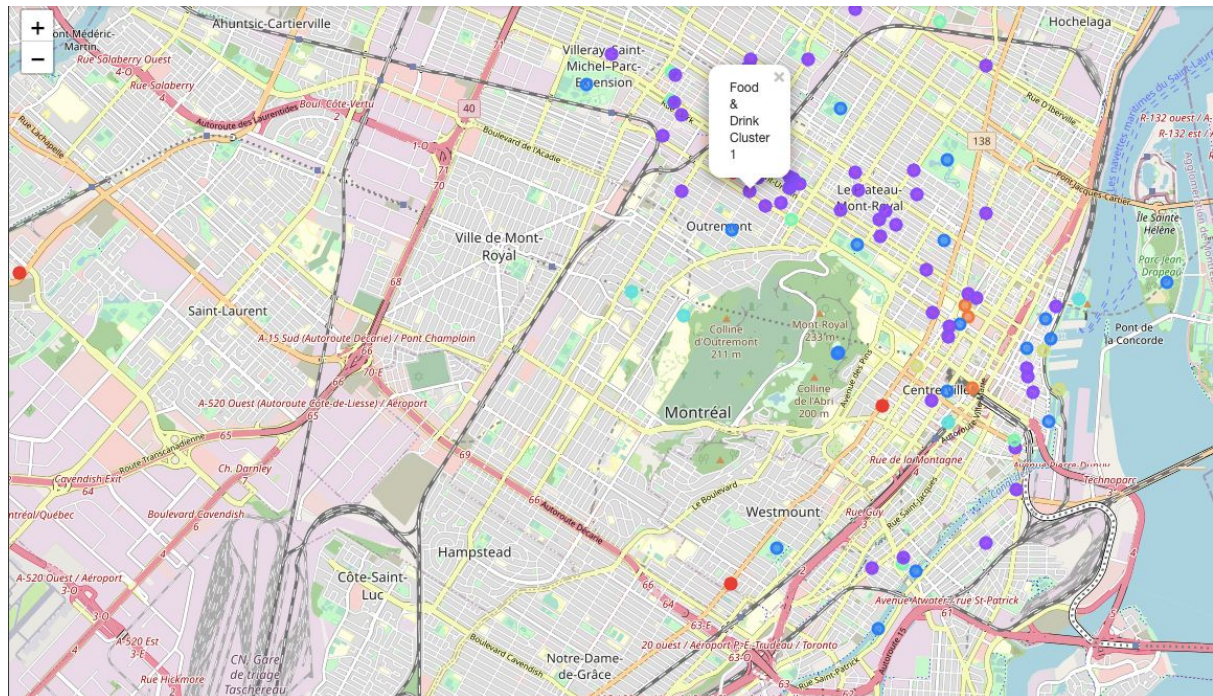
The new column was moved as the second one. The existing column, “categories” was dropped from the dataframe. And the new column, “NewCat” was renamed as “categories”.

In this way, the old category column had been replaced as the new one:

	name	categories	address	lat	lng
0	Parc du Mont-Royal	Public Place	1260 Chemin Remembrance	45.504079	-73.587321
1	Parc Westmount Park	Public Place	Parc Westmount	45.481574	-73.597476
2	Chalet du Mont-Royal	Public Place	1196 voie Camillien-Houde	45.503904	-73.587460
3	Want Apothecary	Store	4960 Sherbrooke Street, West	45.477508	-73.605049
4	Drawn & Quarterly	Store	211 rue Bernard Ouest	45.524748	-73.604614
5	Boulangerie Guillaume	Food & Drink	5132 boulevard St-Laurent	45.523405	-73.593662
6	Dispatch Coffee	Food & Drink	267 St-Zotique	45.528903	-73.616104

Clustering was performed based on the new result. Then new cluster labels were added to the existing dataframe.

The spots were visualized in new clusters as follows:



Most spots are presented as purple markers. Those are included in the Cluster 1, “Food & Drink”. The result of visualization shows that most types of places are restaurant business.

To double check if the values of categories are consistent in each cluster, they were investigated by clusters.

Cluster 0: Stores

categories		name	address
3	Store	Want Apothecary	4960 Sherbrooke Street, West
4	Store	Drawn & Quarterly	211 rue Bernard Ouest
78	Store	Casa del Habano	1434 rue Sherbrooke Ouest
99	Store	Pépinière Jasmin	6305 Boulevard Henri Bourassa

Cluster 1: Food & Drink

	categories	name	address
5	Food & Drink	Boulangerie Guillaume	5132 boulevard St-Laurent
6	Food & Drink	Dispatch Coffee	267 St-Zotique
13	Food & Drink	Cadet	1431, boul. Saint-Laurent
17	Food & Drink	Café Olimpico	124 rue Saint-Viateur Ouest
18	Food & Drink	Boucherie Lawrence	5237 St Laurent
20	Food & Drink	Damas	1201 Van Horne
21	Food & Drink	La Tamalera	226, avenue Fairmount Ouest
22	Food & Drink	La Bête à Pain	195 rue Young
24	Food & Drink	Larry's	9, avenue Fairmount Est

Cluster 2: Public Places

	categories	name	address
0	Public Place	Parc du Mont-Royal	1260 Chemin Remembrance
1	Public Place	Parc Westmount Park	Parc Westmount
2	Public Place	Chalet du Mont-Royal	1196 voie Camillien-Houde
8	Public Place	Parc La Fontaine	3933 rue du Parc-Lafontaine
11	Public Place	Vieux-Port de Montréal / Old Port of Montreal ...	Vieux-Port de Montréal
14	Public Place	Place des Festivals	Quartier des Spectacles
15	Public Place	Canal-de-Lachine	Lachine canal / Canal de Lachine
16	Public Place	Parc Sir-Wilfrid-Laurier	Parc Sir-Wilfrid-Laurier
25	Public Place	Basilique Notre-Dame (Basilique Notre-Dame de ...	110, rue Notre-Dame Ouest

Cluster 3: Sports

categories		name	address
27	Sports	CEPSUM	2100 boul. Édouard-Montpetit
57	Sports	Quai de l'Horloge / Clock Tower quay	Rue Quai de l'Horloge
58	Sports	Centre Bell	1909 Ave. des Canadiens-de-Montréal
77	Sports	Circuit Gilles-Villeneuve	1, circuit Gilles-Villeneuve
81	Sports	Midtown Le Sporting Club Sanctuaire	6105 avenue du Boisé
93	Sports	Académie Arts Martiaux Brossard	1505, boulevard Provencher

Cluster 4: Markets

categories		name	address
7	Market	Milano Fruiterie	6862 boul. Saint-Laurent
9	Market	Marché Atwater	138 ave. Atwater
10	Market	Marché Jean-Talon	7070 avenue Henri-Julien
34	Market	Adonis	225 rue Peel
65	Market	PA Nature	5029 Ave du Parc
89	Market	Odessa l'Entrepôt	100-4900 rue Molson

Cluster 5: Hotels & Spas

categories		name	address
7	Market	Milano Fruiterie	6862 boul. Saint-Laurent
9	Market	Marché Atwater	138 ave. Atwater
10	Market	Marché Jean-Talon	7070 avenue Henri-Julien
34	Market	Adonis	225 rue Peel
65	Market	PA Nature	5029 Ave du Parc
89	Market	Odessa l'Entrepôt	100-4900 rue Molson

Cluster 6: Entertainment

	categories	name	address
12	Entertainment	Place des Arts	175 rue Sainte-Catherine Ouest
19	Entertainment	96,9 CKOI	1100-800 rue De La Gauchetière Ouest
23	Entertainment	La Maison Symphonique de Montréal	1600 rue Saint-Urbain

It is checked that all clusters have consistent values for categories.

7. Conclusion

The client wants to open their pop-up stand in Montreal. They prefer the most active neighborhood for food related business to get more chances to encourage the owners to be their partners.

On the map, the venues which are included in “Food & Drink” category are presented in the neighborhood named, “Le Plateau - Mont-Royal”.

As a conclusion, the best area to open the pop-up stand is **Le Plateau - Mont-Royal**.