

The Battle of Neighborhoods: Exploring Montreal Neighborhoods to find the best location for Korean restaurant.

Olga McIngvale

May 9th, 2021

Introduction: Business Problem

Montreal is the second-most populous city in Canada and most populous city in the Canadian province of Quebec. Montreal has the second-largest economy in Canada based on GDP and the largest in Quebec. The city is today an important center of commerce, finance, industry, technology, culture, and world affairs. Montreal was referred to as "Canada's Cultural Capital" as well. The city's many cultural communities have given it a distinct local culture.

This project aims at exploring and segmenting the city of Montreal to give the recommendations to potential investors which were looking for the best place to open the Korean restaurant.

Data

- We will use the Wikipedia data "List of postal codes of Canada: H" for the Metropolitan Montreal (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_H) to get the information of the postal-codes and boroughs/neighborhoods names. The data from the Wikipedia page will be scrapped and wrangled, cleaned, and then read it into a pandas dataframe.
- To get the spatial data (latitude and longitude) we will be using the Open-source geocoding Nominatim from geopy. Latitude and longitude will be added to pandas dataframe.
- We will use Foursquare API to get the common venues in each neighborhood and then use this feature to group the neighborhoods into clusters to determine the best possible fit for the Korean restaurant.

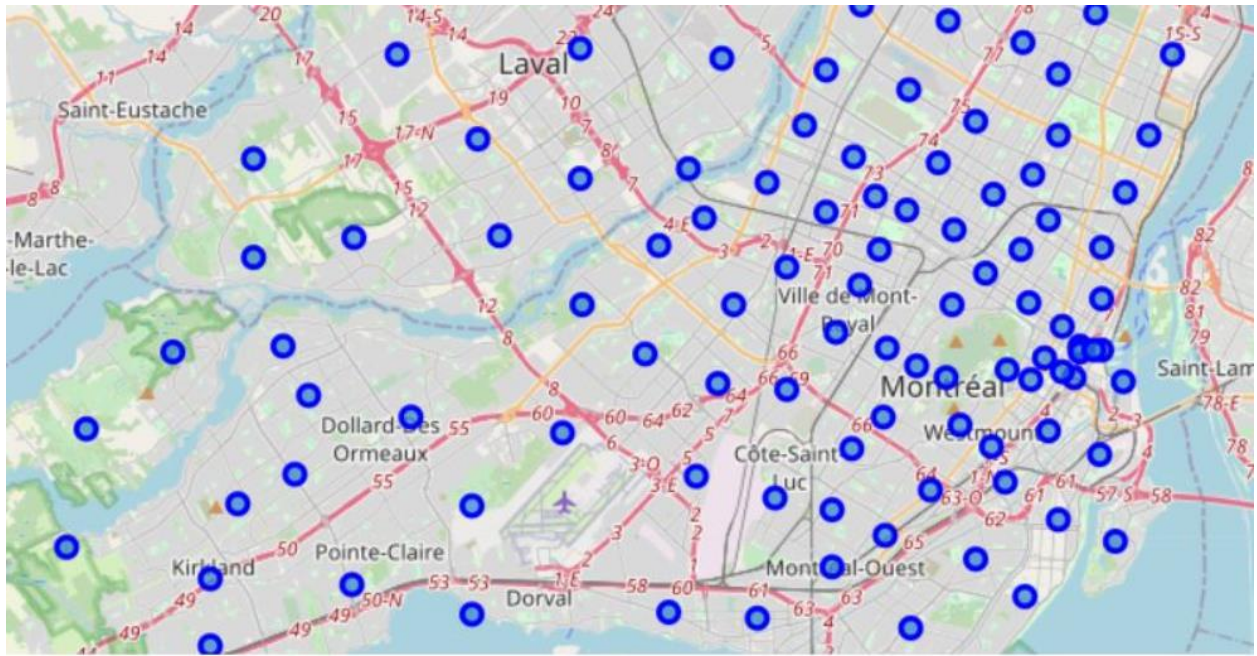
Data have been scrapped by using the BeautifulSoup package, then be cleaned and processed for further analyzing.

Once we have built a dataframe of the postal code of each neighborhood, to utilize the Foursquare location data, we got the latitude and the longitude coordinates of each neighborhood via the Geocoder Python package. Finally we have our clean dataset including the spatial data.

Methodology

In this project, we will search for venues using Foursquare API and then cluster our Montreal neighborhoods based on common venues via clustering algorithm Kmeans.

We used a great visualization library Folium, to create a map of Montreal using latitude and longitude values.



We will consider the center area of Montreal, downtown and around area with the higher density of people including tourists. So, the original dataframe has been sliced and created a new dataframe which narrowed considered neighborhoods quantity from 122 to 17, located in the city center.

Next, we started utilizing the Foursquare API to explore the selected Montreal neighborhoods and segment them.

By utilizing the Foursquare API 497 venues have been found, which have been grouped into 150 unique categories.

We are interested in all venues categories, not only those that are restaurants or cafes. Because it is good to know that even the location is not very crowded with food places (which we eventually will be looking for), but also what exactly is popular in that area, a background (which can attract more potential clients).

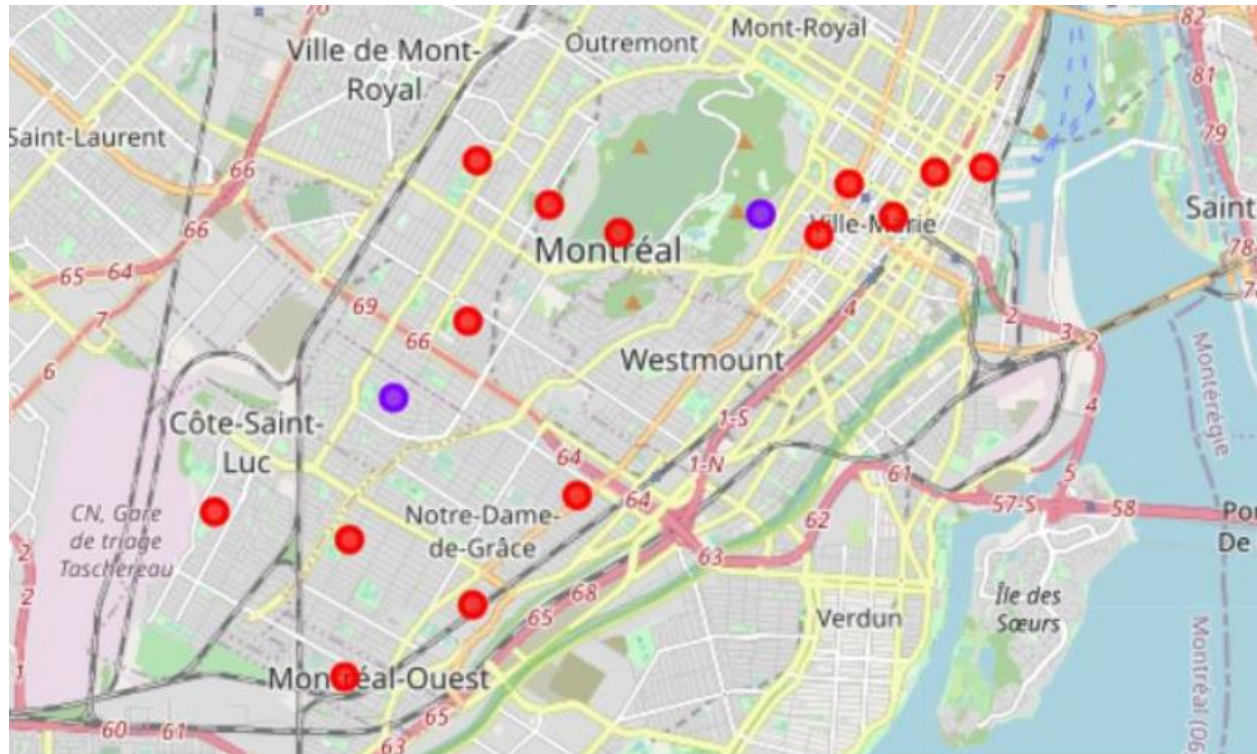
Analysis

To analyze each neighborhood, we grouped rows by neighborhood and by taking the mean of the frequency of occurrence of each category.

Then we checked each neighborhood along with the top 5 most common venues. It does not look detailed enough, so the new dataframe which displays the top 10 venues for each neighborhood has been created.

We have used the clustering algorithm (unsupervised process for the unlabeled dataset) K-means to cluster the Montreal neighborhoods into 3 clusters.

The new dataframe that includes the clusters as well as the top 10 venues for each neighborhood has been created.



Results and Discussion

Now, we can examine each cluster and determine the discriminating venue categories that distinguish each cluster.

According to the characteristics of each clustering, we could now easily and intuitively find our ideal location in the map. When we click on the spot in the map, the neighborhood name is shown on the pop-up label.

Since there are lots of restaurants in neighborhoods from cluster 1 (red circles) as the most common venues, we easily can see that locations from the cluster 2 (purple circles) are not crowded with restaurants at all. Since we have already limited locations by the city center, we will choose between Downtown Montreal Southwest and Hampstead / Côte Saint-Luc. Downtown Montreal Southwest could be the best one since it includes large green spaces and recreational facilities. That is a popular area and could be the best location to open the new Korean restaurant.

The cluster 2 presented below.

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
Downtown MontrealSouthwest	Historic Site	Park	Scenic Lookout	Mountain	Performing Arts Venue	New American Restaurant	Nightclub	Noodle House	Paper / Office Supplies Store	Pastry Shop
Hampstead / Côte Saint-Luc	Park	Big Box Store	Food Service	Yoga Studio	Performing Arts Venue	New American Restaurant	Nightclub	Noodle House	Paper / Office Supplies Store	Pastry Shop

This concludes our analysis. We have found the best location with low number of restaurants and no Korean restaurants nearby; the zone is close to city center and very popular.

Conclusion

Purpose of this project was to find the best location in Montreal, Canada to open a new Korean restaurant. Location should have low number of restaurants in general (particularly Korean or Asian restaurants).

By using Foursquare API we found the most common venues and then clustered our neighborhoods based on common venues.

The optimal location to open a new restaurant will be offered to stakeholders for the final decision. The other specific characteristics and additional factors can be taken into consideration, but it is not a part of this project.