Location of a New Indian Restaurant in Montreal, Quebec

Andrew Stutman

May 29, 2021

Table of Contents

I. Abstract

Tasked with helping stakeholders decide upon the location of a new Indian restaurant in Montreal, Quebec, I employed Montreal's postal code data, Foursquare API data, and geospatial data.

Analyzed data, involving clustering, indicated that Montreal has a variety of venues throughout Montreal, including restaurants. Analysis also indicated that vicinity of downtown Montreal contains the greatest density of venues among Montreal. Further, Foursquare API analysis revealed that Montreal has 10 Indian restaurants, which are all located in the vicinity of downtown Montreal.

This preliminary analysis should serve to inform stakeholders' further refined preferences and involved communications as to the location of the new Indian restaurant.

II. Introduction

This report is part of Coursera's Applied Data Science Capstone, which is the final course in the IBM Data Science Professional Certificate. The involved course assignment requires leveraging Foursquare location data to conceive of and solve a problem.

II. a. Problem

Stakeholders are seeking to open a new Indian restaurant in Montreal, Quebec, and they are challenged with deciding upon the specific location of the restaurant.

II. b. Data

As the data scientist employed by the stakeholders, I used a dataset of Montreal's current postal codes, Foursquare API data of Montreal venues, and geospatial data to resolve clusters of Montreal venues and locations of existing Indian restaurants in Montreal. This data analysis provides insight as to potential locations of the stakeholders' new Indian restaurant.

III. Methodology

I accessed Montreal's postal codes from a widipedia.org webpage (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_H), containing a table of Montreal's postal codes and boroughs. I web scraped the table to create a dataframe of Montreal's postal codes and borough.

I then used geocoder to add longitude and latitude data to the dataframe.

Finding that two postal codes are not applicable for this analysis (i.e., H0H is reserved for Santa Claus, H0P is not assigned to borough), I cleaned the dataframe and reset the index.

Next, I used folium to map Montreal's postal codes as displayed below.
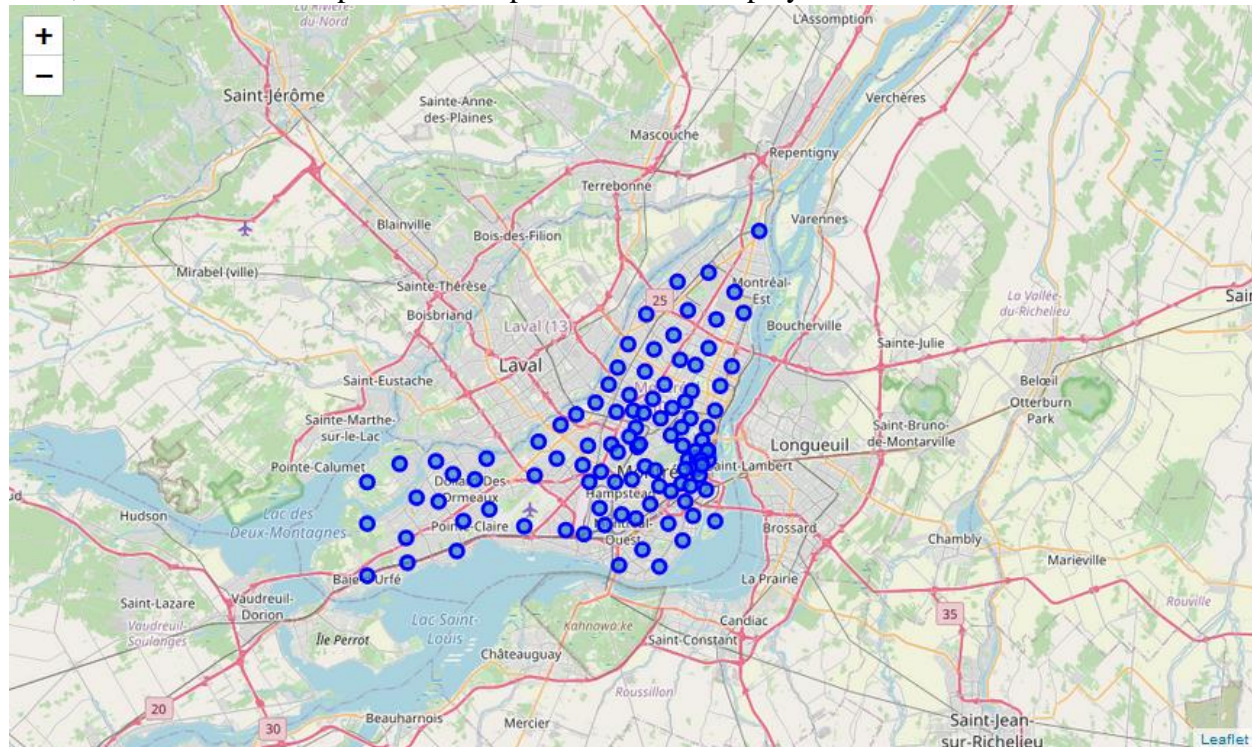

I then created a function using Foursquare API to create a dataframe of Montreal's venues.

Analysis of Montreal's venues, as provided by Foursquare, indicates that there are 263 unique venue categories, including 'Indian Restaurant.'

I then created a further dataframe that grouped Montreal's venues by postal code to facilitate clustering analysis. To this dataframe, I added columns of the 10 most common venue categories by postal code.

After clustering the dataframe into five clusters, and adding the associated cluster to each postal code record, I cleaned the dataframe remove null/NaN values.

I again used folium to map the dataframe of Montreal's postal codes clustered by Foursquare venues data, resulting in the following visualization:
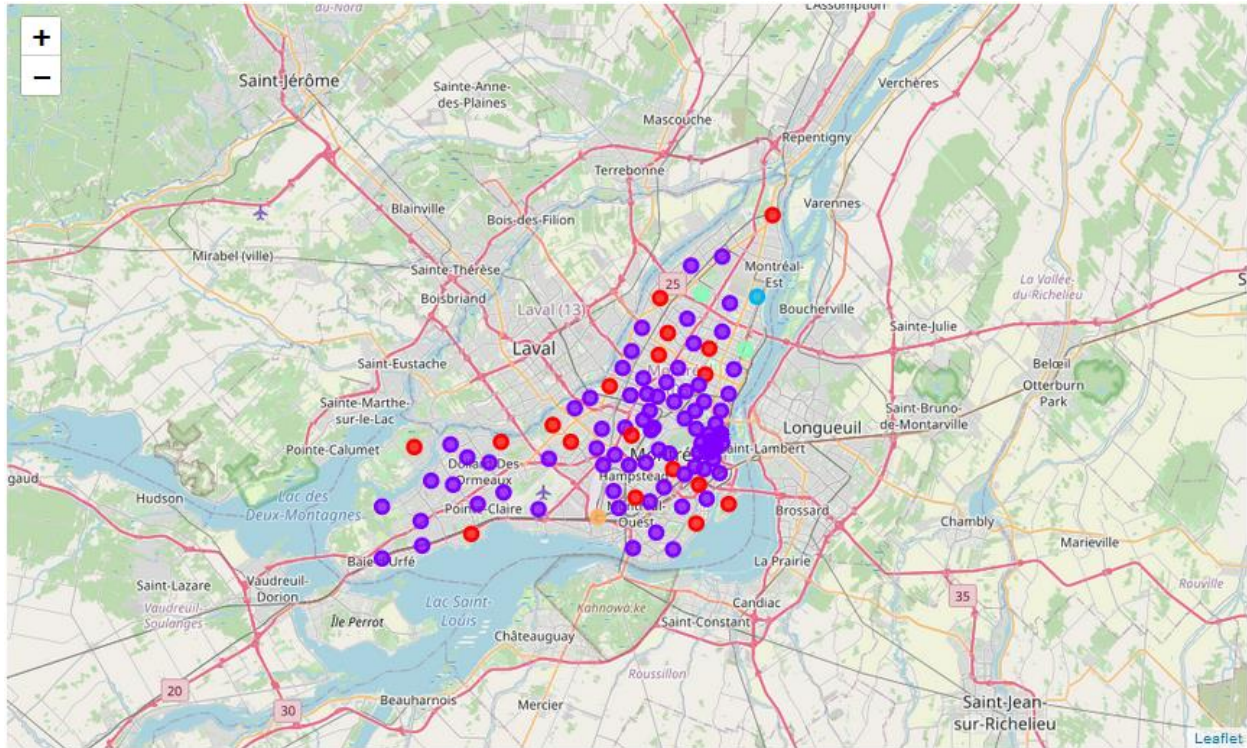
Analysis of the five clusters' common venues (with the associated map marker color in parentheses) are:

1 (red marker): recreation, shopping, restaurants

2 (purple marker): restaurants, hotels, shopping/markets

3 (blue marker): Mercier North; shopping/markets, yoga, restaurants

4 (green marker): Anjou West and Mercier Southeast; recreation, shopping/markets, restaurants

5 (orange marker): Lachine East; storage, yoga, restaurants, markets

The above clusters map and analysis indicate that there are various venues throughout Montreal, including restaurants.

Next, I created a dataframe of Montreal's existing Indian restaurants from a dataframe of Montreal's venues as created from Foursquare data, which produced 10 records.

I then added Montreal's 10 Indian restaurants to the above clusters map (indicated by black and green circles with blue and white hovering markers), producing the following map:
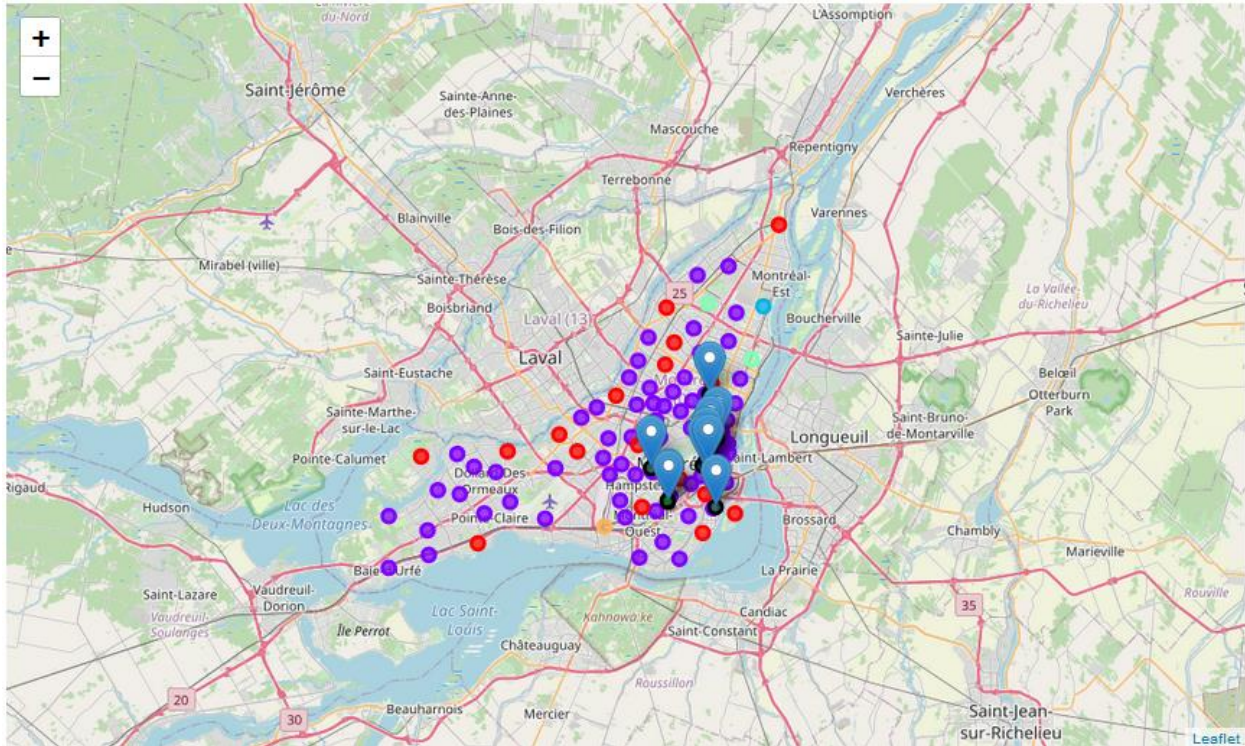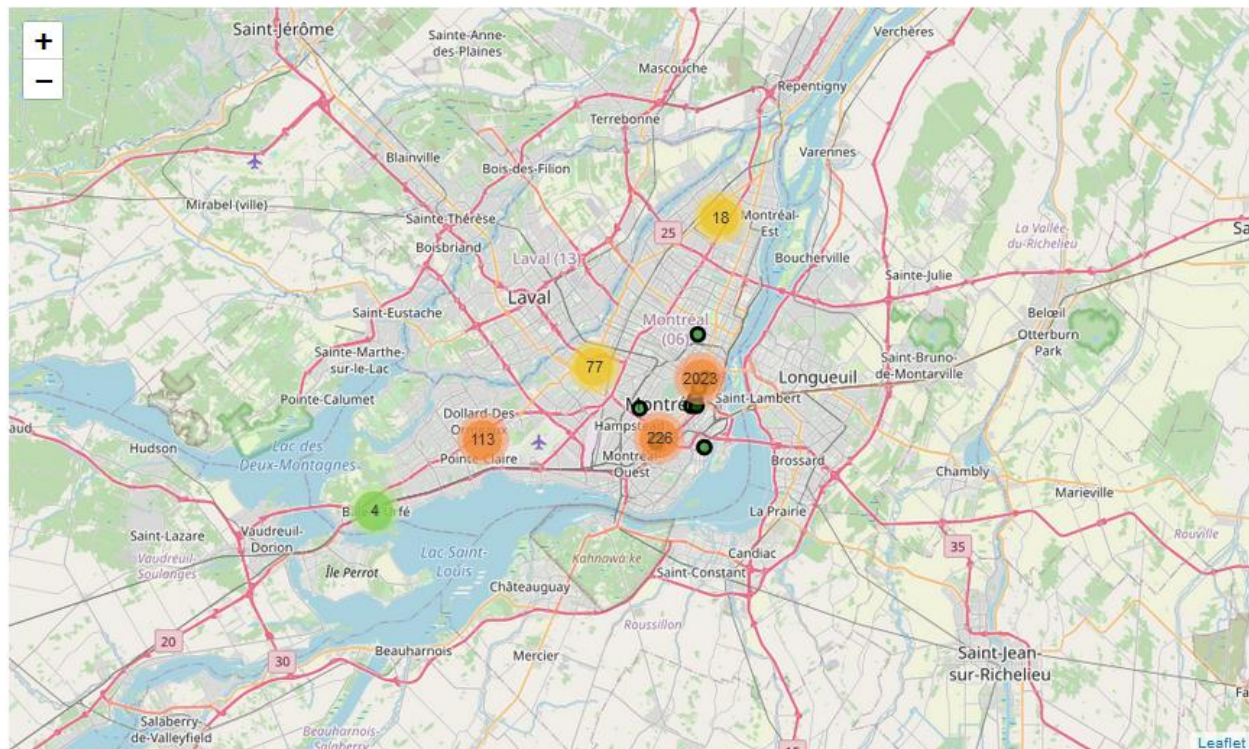
The visual displays that Montreal's 10 Indian restaurants are located in the vicinity of downtown Montreal (in the central/eastern side of Montreal).

To visualize the geographic density of Montreal's venues as provided by Foursquare, I again used folium to create Marker Clusters of Montreal's venues, which produced the following map:
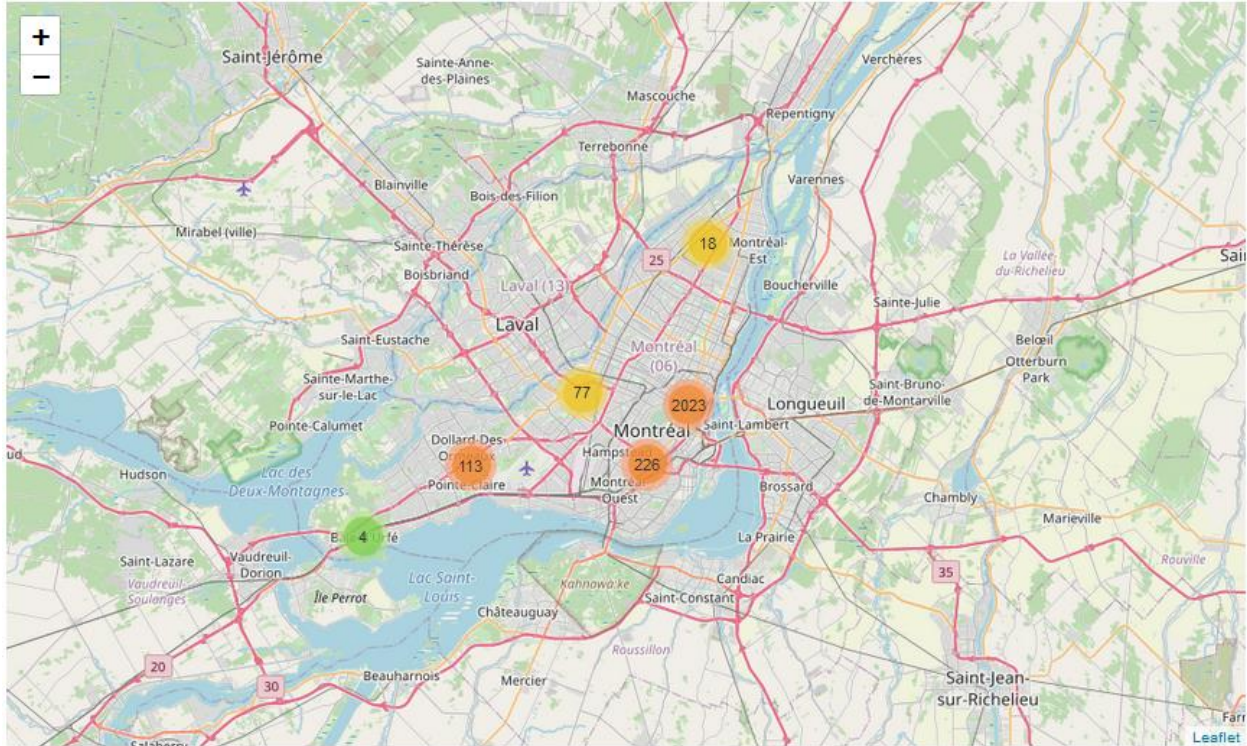
The above visual indicates that vicinity of downtown Montreal has the greatest density of venues in Montreal.

I then added Montreal's 10 Indian restaurants (black and green markers) to the map indicating the geographic density of Montreal's venues, which produced the following map:

The above map indicates that Montreal's 10 existing Indian restaurants are located in the vicinity of downtown Montreal, which is the area in which Montreal has the greatest geographic density of venues.

IV. Results

As found in the methodology section above, I use clustering of the processed Foursquare API venues data to determine 5 clusters of Montreal venues that have some similarity. All 5 clusters indicated varieties of venues throughout Montreal, including restaurants.

I also generated a Marker Cluster maps, which reveal the geographic density of venues. The vicinity of downtown Montreal features the greatest density of venues.

According to Foursquare, Montreal has 10 Indian restaurants, which are all located in the vicinity of downtown Montreal.

V. Discussion

Given that Montreal's 10 Indian restaurants are all relatively close to each other on the eastern-central side of Montreal (including Maison India and Start of India located virtually across the street, Rue Sherbrook Ouest, from each other), and that cluster analysis reveals venues throughout the entirety of Montreal (i.e., commerce exists across Montreal, including restaurants), stakeholders may have opportunities to locate their new Indian restaurant away

from the 10 existing Indian restaurants (corresponding to areas of venue density that are lower than that of eastern-central/downtown Montreal).

Placement of the new Indian restaurant away from the 10 existing closely-located Indian restaurants may be a strategy that the stakeholders should consider to potentially capture more patronage as compared to placing the new restaurant among the existing Indian restaurants.

VI. Conclusion

The purpose of this project was to assist stakeholders with deciding upon the location of a new Indian restaurant in Montreal, Quebec. By identifying common venues among Montreal's postal codes and clusters of venues from Foursquare data, geographic density of venues, and identifying the location of Montreal's existing Indian restaurants, stakeholders now have preliminary knowledge for further communications with their employed data scientist to narrow down the restaurant's potential location.

The stakeholders should continue to work with the data scientist to communicate location preferences (e.g., in areas of greater or lower venue/restaurant density, parking availability for customers, crime data, prices of available facilities/real estate to house the restaurant and associated costs/renovation, etc.) for further insight.

VII. References

https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_H

VIII. Acknowledgement

I would like to acknowledge Coursera.org, IBM, IBM staff associated with Coursera's Applied Data Science Capstone course and IBM's Data Science Professional Certificate, and fellow peers enrolled in Coursera's Applied Data Science Capstone course for this opportunity and support. Thank you.