IBM Data Science Professional Certificate

Capstone Project

The Battle of Neighbourhoods

A New Restaurant in Vancouver, BC

Final Report

Presented by:
Nathalia Contreras

The Battle of Neighbourhoods

A New Restaurant in Vancouver, BC

1. ## Introduction/Business Problem

Vancouver is a metropolitan city located in British Columbia in Canada. It is considered a dynamic and multicultural city; mainly recognised for its diverse green areas and its great variety of activities for everyone. In Vancouver, people can find a plethora of restaurants that offer various cuisine that goes from chinese to italian or mediterranean.

French cuisine is one of the most popular existing cuisines. Some people consider it to have a complex preparation. Nonetheless, others consider it to be unique and influential due to its ingredients, techniques, and presentation. When thinking of french food, the word 'sophistication' and 'pricey' immediately come to mind, however that is up to the occasion and the menu. In summary, french gastronomy is about the mixture of flavours and textures that one can try in every bite.

Now, combining the two ideas mentioned above, there is a businessperson (stakeholder), who is looking forward to opening a french cuisine restaurant in the city of Vancouver, British Columbia, Canada.

The current project aims to find the most suitable location for a french cuisine restaurant in Vancouver.

2. ## Data

Vancouver consists of 22 local areas, which are also known as 'Neighbourhoods'.

In the official website of the city of Vancouver, there is a section called "The Open Data Catalogue", which provides access to many datasets in different formats.

From the Open Data Catalogue - "Neighbourhoods" option - the following link is presented:
- https://data.vancouver.ca/datacatalogue/localAreaBoundary.htm

This dataset was downloaded in CSV format.It has the official names of each local area in Vancouver.

The geographic coordinates (latitude and longitude) of each Local Area (Neighbourhood) are going to be obtained by using: from *geopy.geocoders import Nominatim*.

Therefore, one dataset including the names and IDs of each local area, latitude, and longitude, is going to be created.

The Foursquare API is going to be used in order to retrieve information about venues (Name, Latitude, Longitude, and Category) for each Local Area. Then, information about French Restaurants will be shown, so as to see how those venues are distributed throughout Vancouver.

3. Methodology

a. Data Cleaning:

Data containing the Local Areas Boundaries in Vancouver was downloaded in CSV format from the following link: https://data.vancouver.ca/datacatalogue/localAreaBoundary.htm
The CSV file only contains the name of each local area and an ID for each.

The file was uploaded and converted to a dataframe. The first 5 rows are:

	MAPID	NAME
0	SUN	Sunset
1	MP	Mount Pleasant
2	RP	Riley Park
3	CBD	Downtown
4	KITS	Kitsilano

The shape of this first dataframe is: 22 rows and 2 columns.

To get the geographical coordinates more precisely, it is important to add the city of each local area, in order to avoid confusion with other areas with similar names. In that sense, the string ", Vancouver, BC" will be added after each local area. The dataset now will be:

	MAPID	LocalArea
0	SUN	Sunset, Vancouver, BC
1	MP	Mount Pleasant, Vancouver, BC
2	RP	Riley Park, Vancouver, BC
3	CBD	Downtown, Vancouver, BC
4	KITS	Kitsilano, Vancouver, BC

b. Add coordinates to each local area

To obtain the geographic coordinates (latitude and longitude) of each Local Area, the following library is going to be used: *from geopy.geocoders import Nominatim*.
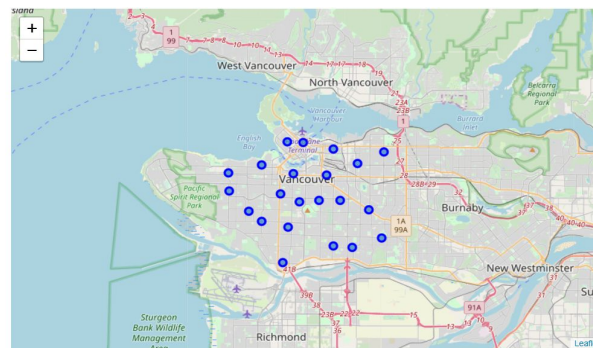
After obtaining the coordinates, these are combined with the first dataframe, creating a new one, which is the following:

	MAPID	LocalArea	Latitude	Longitude
0	SUN	Sunset, Vancouver, BC	49.219593	-123.090239
1	MP	Mount Pleasant, Vancouver, BC	49.263330	-123.096588
2	RP	Riley Park, Vancouver, BC	49.247438	-123.102966
3	CBD	Downtown, Vancouver, BC	49.283393	-123.117456
4	KITS	Kitsilano, Vancouver, BC	49.269410	-123.155267

This last dataframe is the main one because it contains all the data needed for further steps.

c. Visualise local areas in Vancouver

A map of Vancouver with all its local areas is shown in order to know the geographic location of each one:

d. Use Foursquare API

Foursquare API is needed in order to get the top 100 venues within 500 metres for each Local Area in Vancouver. For each venue, data about its name, latitude, longitude and category are going to be extracted.

	LocalArea	LocalArea Latitude	LocalArea Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Sunset, Vancouver, BC	49.219593	-123.090239	Kalai's Dosa Hut	49.218998	-123.091160	South Indian Restaurant
1	Sunset, Vancouver, BC	49.219593	-123.090239	New Novelty Restaurant and Sweets	49.223925	-123.090885	Dessert Shop
2	Sunset, Vancouver, BC	49.219593	-123.090239	Vision 2000	49.224026	-123.090798	Cosmetics Shop
3	Sunset, Vancouver, BC	49.219593	-123.090239	Sartaj Sweets & Restaurant	49.224046	-123.090841	Indian Restaurant
4	Mount Pleasant, Vancouver, BC	49.263330	-123.096588	Dude Chilling Park	49.263730	-123.096796	Outdoor Sculpture

This dataframe contains various venue categories; it has 153 unique categories. However, filtering is needed owing that the focus is only on 'Restaurants'.

Dataframe after filtering: *(It has 227 rows and 7 columns)*

	LocalArea	LocalArea Latitude	LocalArea Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Sunset, Vancouver, BC	49.219593	-123.090239	Kalai's Dosa Hut	49.218998	-123.091160	South Indian Restaurant
1	Sunset, Vancouver, BC	49.219593	-123.090239	Sartaj Sweets & Restaurant	49.224046	-123.090841	Indian Restaurant
2	Mount Pleasant, Vancouver, BC	49.263330	-123.096588	Sushiyama	49.262897	-123.097169	Sushi Restaurant
3	Mount Pleasant, Vancouver, BC	49.263330	-123.096588	Carp	49.262393	-123.096070	Sushi Restaurant
4	Mount Pleasant, Vancouver, BC	49.263330	-123.096588	Congee Noodle House 粥麵館 (Congee Noodle House)	49.263029	-123.102105	Chinese Restaurant

After grouping, counting, and sorting the data from above:

- The local areas of West End, and Grandview-Woodland have the largest number of restaurants, with 28 each.
- The local areas of Killarney, Shaughnessy, and Victoria-Fraserview have the smallest number of restaurants, with 1 each.

e. Analyse each local area

Applying *One Hot Encoding,* then grouping by Local Area to calculate the mean of the frequency of occurrence of each category. This dataframe has 21 rows and 7 columns. The reason there are now 21 rows is because 'Arbutus-Ridge, Vancouver, BC' has no restaurant venues, hence no venues in that local area were selected.

Now, calculate the top 5 common venues in each local area, and save the result in a dataframe:

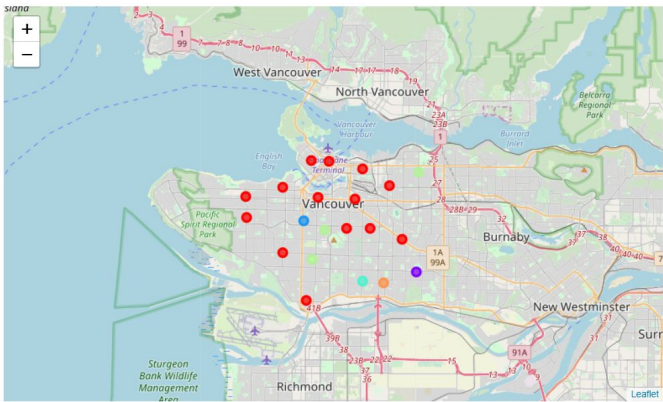|	LocalArea	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Downtown, Vancouver, BC	Vietnamese Restaurant	Seafood Restaurant	Restaurant	American Restaurant	Hawaiian Restaurant	Mexican Restaurant	Japanese Curry Restaurant	Australian Restaurant	French Restaurant	Chinese Restaurant
1	Dunbar-Southlands, Vancouver, BC	Sushi Restaurant	Italian Restaurant	Indian Restaurant	Fast Food Restaurant	Vietnamese Restaurant	Falafel Restaurant	Japanese Curry Restaurant	Hawaiian Restaurant	Greek Restaurant	French Restaurant
2	Fairview, Vancouver, BC	Asian Restaurant	Vietnamese Restaurant	Chinese Restaurant	Korean Restaurant	Indian Restaurant	Malay Restaurant	Fast Food Restaurant	Falafel Restaurant	Restaurant	Sushi Restaurant
3	Grandview-Woodland, Vancouver, BC	Sushi Restaurant	Italian Restaurant	Indian Restaurant	Japanese Restaurant	French Restaurant	Seafood Restaurant	Greek Restaurant	Vegetarian / Vegan Restaurant	Mexican Restaurant	Middle Eastern Restaurant
4	Hastings-Sunrise, Vancouver, BC	Vietnamese Restaurant	Sushi Restaurant	Falafel Restaurant	Japanese Curry Restaurant	Italian Restaurant	Indian Restaurant	Hawaiian Restaurant	Greek Restaurant	French Restaurant	Filipino Restaurant

f. Cluster local areas

The machine learning algorithm used for this project is k-means because the goal is to find similar local areas in Vancouver and by which aspects they are similar.

By running k-means, the local areas are going to be clustered into 6 clusters. The result will be saved in a dataframe which will also include the top 10 venues for each local area.

Dataframe: *(It has 21 rows and 15 columns)*

	MAPID	LocalArea	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	10th Most Common Venue
0	SUN	Sunset, Vancouver, BC	49.219593	-123.090239	3	Indian Restaurant	South Indian Restaurant	Vietnamese Restaurant	Falafel Restaurant	Japanese Curry Restaurant	Italian Restaurant	Hawaiian Restaurant	Greek Restaurant	French Restaurant	Filipino Restaurant
1	MP	Mount Pleasant, Vancouver, BC	49.263330	-123.096588	0	Sushi Restaurant	Vietnamese Restaurant	Indian Restaurant	Chinese Restaurant	Mexican Restaurant	New American Restaurant	Peruvian Restaurant	Ethiopian Restaurant	Restaurant	Asian Restaurant
2	RP	Riley Park, Vancouver, BC	49.247438	-123.102966	0	Restaurant	Thai Restaurant	Japanese Restaurant	Vegetarian / Vegan Restaurant	Sushi Restaurant	Chinese Restaurant	Vietnamese Restaurant	Middle Eastern Restaurant	Caribbean Restaurant	Seafood Restaurant
3	CBD	Downtown, Vancouver, BC	49.283393	-123.117456	0	Vietnamese Restaurant	Seafood Restaurant	Restaurant	American Restaurant	Hawaiian Restaurant	Mexican Restaurant	Japanese Curry Restaurant	Australian Restaurant	French Restaurant	Chinese Restaurant
4	KITS	Kitsilano, Vancouver, BC	49.269410	-123.155267	0	Sushi Restaurant	American Restaurant	French Restaurant	Japanese Restaurant	Thai Restaurant	Restaurant	Asian Restaurant	Indian Restaurant	Vegetarian / Vegan Restaurant	Seafood Restaurant

Now, for a better visualisation of each cluster, create a map:

g. Examine Clusters

Examine each cluster and determine the venue categories that distinguish each one.

Cluster 1:

	LocalArea	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	10th Most Common Venue
1	Mount Pleasant, Vancouver, BC	0	Sushi Restaurant	Vietnamese Restaurant	Indian Restaurant	Chinese Restaurant	Mexican Restaurant	New American Restaurant	Peruvian Restaurant	Ethiopian Restaurant	Restaurant	Asian Restaurant
2	Riley Park, Vancouver, BC	0	Restaurant	Thai Restaurant	Japanese Restaurant	Vegetarian / Vegan Restaurant	Sushi Restaurant	Chinese Restaurant	Vietnamese Restaurant	Middle Eastern Restaurant	Caribbean Restaurant	Seafood Restaurant
3	Downtown, Vancouver, BC	0	Vietnamese Restaurant	Seafood Restaurant	Restaurant	American Restaurant	Hawaiian Restaurant	Mexican Restaurant	Japanese Curry Restaurant	Australian Restaurant	French Restaurant	Chinese Restaurant
4	Kitsilano, Vancouver, BC	0	Sushi Restaurant	American Restaurant	French Restaurant	Japanese Restaurant	Thai Restaurant	Restaurant	Asian Restaurant	Indian Restaurant	Vegetarian / Vegan Restaurant	Seafood Restaurant
5	Dunbar-Southlands, Vancouver, BC	0	Sushi Restaurant	Italian Restaurant	Indian Restaurant	Fast Food Restaurant	Vietnamese Restaurant	Falafel Restaurant	Japanese Curry Restaurant	Hawaiian Restaurant	Greek Restaurant	French Restaurant
6	Kerrisdale, Vancouver, BC	0	Chinese Restaurant	Vietnamese Restaurant	Fast Food Restaurant	Sushi Restaurant	Thai Restaurant	Asian Restaurant	Italian Restaurant	Mediterranean Restaurant	Japanese Restaurant	Portuguese Restaurant
8	West Point Grey, Vancouver, BC	0	Japanese Restaurant	Vegetarian / Vegan Restaurant	Asian Restaurant	Sushi Restaurant	Fast Food Restaurant	Italian Restaurant	South American Restaurant	Falafel Restaurant	Japanese Curry Restaurant	Indian Restaurant
9	Marpole, Vancouver, BC	0	Sushi Restaurant	Chinese Restaurant	Japanese Restaurant	Vietnamese Restaurant	Thai Restaurant	Taiwanese Restaurant	Shanghai Restaurant	Dim Sum Restaurant	Falafel Restaurant	Fast Food Restaurant
12	Fairview, Vancouver, BC	0	Asian Restaurant	Vietnamese Restaurant	Chinese Restaurant	Korean Restaurant	Indian Restaurant	Malay Restaurant	Fast Food Restaurant	Falafel Restaurant	Restaurant	Sushi Restaurant
14	West End, Vancouver, BC	0	Japanese Restaurant	Greek Restaurant	Indian Restaurant	American Restaurant	Ramen Restaurant	Sushi Restaurant	Italian Restaurant	Restaurant	Chinese Restaurant	Falafel Restaurant
16	Renfrew-Collingwood, Vancouver, BC	0	Vietnamese Restaurant	Japanese Restaurant	Chinese Restaurant	Cantonese Restaurant	Shanghai Restaurant	Vegetarian / Vegan Restaurant	Asian Restaurant	Caribbean Restaurant	French Restaurant	Japanese Curry Restaurant
19	Kensington-Cedar Cottage, Vancouver, BC	0	Chinese Restaurant	Vietnamese Restaurant	Malay Restaurant	Greek Restaurant	Filipino Restaurant	Fast Food Restaurant	Japanese Curry Restaurant	Italian Restaurant	Indian Restaurant	Hawaiian Restaurant
20	Strathcona, Vancouver, BC	0	Vietnamese Restaurant	Restaurant	Japanese Restaurant	Asian Restaurant	Fast Food Restaurant	Japanese Curry Restaurant	Italian Restaurant	Indian Restaurant	Hawaiian Restaurant	Greek Restaurant
21	Grandview-Woodland, Vancouver, BC	0	Sushi Restaurant	Italian Restaurant	Indian Restaurant	Japanese Restaurant	French Restaurant	Seafood Restaurant	Greek Restaurant	Vegetarian / Vegan Restaurant	Mexican Restaurant	Middle Eastern Restaurant

Cluster 2:

	LocalArea	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	10th Most Common Venue
15	Killarney, Vancouver, BC	1	Italian Restaurant	Vietnamese Restaurant	Vegetarian / Vegan Restaurant	Japanese Restaurant	Japanese Curry Restaurant	Indian Restaurant	Hawaiian Restaurant	Greek Restaurant	French Restaurant	Filipino Restaurant

Cluster 3:

	LocalArea	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	10th Most Common Venue
11	Shaughnessy, Vancouver, BC	2	French Restaurant	Vietnamese Restaurant	Vegetarian / Vegan Restaurant	Japanese Restaurant	Japanese Curry Restaurant	Italian Restaurant	Indian Restaurant	Hawaiian Restaurant	Greek Restaurant	Filipino Restaurant

Cluster 4:

	LocalArea	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	10th Most Common Venue
0	Sunset, Vancouver, BC	3	Indian Restaurant	South Indian Restaurant	Vietnamese Restaurant	Falafel Restaurant	Japanese Curry Restaurant	Italian Restaurant	Hawaiian Restaurant	Greek Restaurant	French Restaurant	Filipino Restaurant

Cluster 5:

	LocalArea	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	10th Most Common Venue
10	Oakridge, Vancouver, BC	4	Vietnamese Restaurant	Sushi Restaurant	Fast Food Restaurant	Falafel Restaurant	Japanese Curry Restaurant	Italian Restaurant	Indian Restaurant	Hawaiian Restaurant	Greek Restaurant	French Restaurant
13	South Cambie, Vancouver, BC	4	Vietnamese Restaurant	Malay Restaurant	Sushi Restaurant	Falafel Restaurant	Japanese Curry Restaurant	Italian Restaurant	Indian Restaurant	Hawaiian Restaurant	Greek Restaurant	French Restaurant
17	Hastings-Sunrise, Vancouver, BC	4	Vietnamese Restaurant	Sushi Restaurant	Falafel Restaurant	Japanese Curry Restaurant	Italian Restaurant	Indian Restaurant	Hawaiian Restaurant	Greek Restaurant	French Restaurant	Filipino Restaurant

	LocalArea	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	10th Most Common Venue
18	Victoria-Fraserview, Vancouver, BC	5	Fast Food Restaurant	Vietnamese Restaurant	Vegetarian / Vegan Restaurant	Japanese Restaurant	Japanese Curry Restaurant	Italian Restaurant	Indian Restaurant	Hawaiian Restaurant	Greek Restaurant	French Restaurant

h. Find French Restaurants in Vancouver
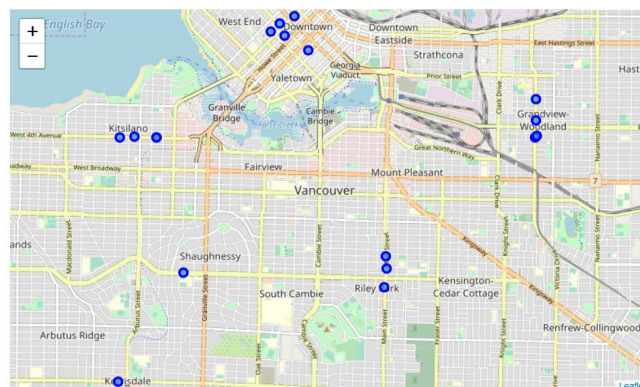
Explore French Restaurants in Vancouver using Foursquare API.

Dataframe: *(It has 18 rows and 7 columns)*

	LocalArea	LocalArea Latitude	LocalArea Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Riley Park, Vancouver, BC	49.247438	-123.102966	Coco et Olive	49.251760	-123.101089	French Restaurant
1	Riley Park, Vancouver, BC	49.247438	-123.102966	The French Table	49.250073	-123.100949	French Restaurant
2	Riley Park, Vancouver, BC	49.247438	-123.102966	Yugo Restaurant	49.247488	-123.101384	French Restaurant
3	Downtown, Vancouver, BC	49.283393	-123.117456	L'Hermitage	49.280139	-123.117480	French Restaurant
4	Downtown, Vancouver, BC	49.283393	-123.117456	Bacchus Lounge	49.282187	-123.122333	French Restaurant

Group, count, and sort values in dataframe:

	Venue
LocalArea	
Downtown, Vancouver, BC	4
Grandview-Woodland, Vancouver, BC	4
Kitsilano, Vancouver, BC	4
Riley Park, Vancouver, BC	3
Kerrisdale, Vancouver, BC	1
Shaughnessy, Vancouver, BC	1
West End, Vancouver, BC	1

Visualise French Restaurant on the map:

4. Results

The results about most popular Restaurant for each Local Area are the following:
- The local areas of West End, and Grandview-Woodland have the largest number of restaurants, with a total of 28 each.
- The local areas of Killarney, Shaughnessy, and Victoria-Fraserview have the smallest number of restaurants, with 1 each.

- Cluster 1: This cluster consists of 14 local areas. The most popular restaurant venues are: Sushi Restaurant and Vietnamese Restaurant
- Cluster 2: The most common restaurant venue is Italian Restaurant in the local area of Killarney.
- Cluster 3: French Restaurant is the most common venue in the local area of Shaughnessy.
- Cluster 4: Indian Restaurant is the most common venue in the local area of Sunset.
- Cluster 5: This cluster consists of 3 local areas, which are Oakridge, South Cambie, and Hastings-Sunrise; and the most common venue in all of them is Vietnamese Restaurant.
- Cluster 6: A Fast Food Restaurant is the most common venue in the local area of Victoria-Fraserview.

The result when exploring only French Restaurants are the following:
- Throughout Vancouver, there are 18 French Restaurant, which are distributed across 7 local areas: Downtown, Grandview-Woodland, Kitsilano, Riley Park, Kerrisdale, Shaughnessy, and West End.
- Downtown, Grandview-Woodland, and Kitsilano have 4 French Restaurants each, whereas Riley Park has 3, and Kerrisdale, Shaughnessy, and West End have 1 each.
- Observing the map of French Restaurant only, most French Restaurants are in residential areas, except for Downtown that not only is a residential area, but also a tourist area. Nonetheless, West End is also considered a tourist area, but only has 1 French Restaurant.
- From the clustering results, in Shaughnessy (cluster 3), the most common venue is a French Restaurant. What is more, there is only 1 French Restaurant in this local area. Hence, there is a high likelihood of a new French Restaurant to be popular in the local area of Shaughnessy.

5. Discussion

- Shaughnessy could be an ideal local area to open a French Restaurant due to the one that already exists is the most common venue in the area.
- Even though West End could be seem to be also an ideal location due to only have 1 French Restaurant, in this local area the most common restaurant is Japanese Restaurant, followed by Greek and Indian Restaurants. Plus, the one French Restaurant in this area is not in the top 10 of most common venues.
- Downtown has 4 French Restaurants. Although the most common restaurant is Vietnamese Restaurant, followed by Seafood Restaurant, the French Restaurant category is in the top 10 of most common venues. For that reason, this local area could also be a potential place to open a French Restaurant. Plus, due to Downtown being a residential and tourist area, the new French Restaurant could be frequented by locals and tourists, increasing the business' profitability.

- Grandview-Woodland with 4 French Restaurants could be also an ideal area to open a French Restaurant, due to this category being the fifth most common venue. This local area harbours international restaurants, and it is recognised for its rich multicultural history and features.

6. Conclusion

- The local areas to be considered to open a French Restaurant are: Shaughnessy, Downtown, and Grandview-Woodland
- Shaughnessy only has one restaurant, so the number of competitors is small.
- Downtown and Grandview-Woodland has 4 French Restaurants each, so the number of competitors is relatively small.
- Grandview-Woodland is more optimal to open a French Restaurant due to from an overall of 28 restaurant venues, 4 are French Restaurant, which is a category in the top 5 most common venue. Additionally, this local area is recognised for its multiculturalism.
- Further research might be needed according to the local area of interest to the client.