

---

# Suggesting Restaurant Location in Vancouver

SILVIA ESPERANZA ÁLVAREZ ZÚÑIGA

Data Science Capstone Project – The Battle of Neighborhoods

## Abstract

*The purpose of this project is to provide and advice related to a physical location. The case analyzes the neighborhoods in Vancouver, British Columbia in Canada, and tries to get the optimum location for opening a new Mexican cuisine restaurant. The process developed contemplates the use of Foursquare to get the venues in Vancouver and the k-means machine learning algorithm for clustering the neighborhoods, based on the venues, and find out which would be the best neighborhood to make the recommendation.*

## I. INTRODUCTION

Vancouver is one of the most important cities in Canada. With the highest population density in Canada, Vancouver is also one of the most ethnically and linguistically diverse cities in the country [1]. Having all the characteristics of a cosmopolitan city, including a great economy, Vancouver is an ideal place to open a new restaurant, in this case an authentic Mexican cuisine restaurant. Designated by UNESCO as a cultural treasure, traditional Mexican cuisine is one of the most popular cuisines in the world.

In order to be successful when starting a business, a very important aspect to consider is the physical location to settle in. Specifically, this report is oriented to make a recommendation, to a group of investors interested on supporting a business of a new authentic Mexican cuisine restaurant, on a convenient region (or regions) to establish this restaurant in Vancouver.

Foursquare services are used to discover the main region o regions of interest in Vancouver for the investors to take a decision and recommendations will be highlighted according to the obtained results.

## II. DATA

The data necessary to accomplish the project are the following:

1. Neighborhoods in Vancouver
  - a. Postal code
  - b. Neighborhood

This data was obtained by scrapping the following Wikipedia page using Beautiful Soup:  
[https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_V](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_V) [2]

2. Latitude and longitude for each Neighborhood:

- a. Postal code
- b. Neighborhood
- c. Latitude
- d. Longitude

Latitude and longitude were obtained using Geopy (Python) library and using postal code as the argument.

We were working with a table like the one presented on Table 1 for making the analysis.

*Table 1. Postal codes and neighborhoods with latitude and longitude*

Postal Code	Neighborhood	Latitude	Longitude
V5K	Hastings-Sunrise	49.28196685	-123.0400039
V5L	Grandview-Woodland	49.27883935	-123.0668426
V5M	Hastings-Sunrise, Renfrew-Collingwood	49.25805275	-123.0401552
V5N	Grandview-Woodland, Kensington-Cedar Cottage	49.25345095	-123.066309
V5P	Kensington-Cedar Cottage, Victoria-Fraserview	49.22236995	-123.0683154
V5R	Renfrew-Collingwood	49.2400053	-123.0412036
V5S	Killarney	49.21567675	-123.041565
V5T	Mount Pleasant	49.261967	-123.0911945
V5V	Kensington-Cedar Cottage, Riley Park-Little Mountain	49.24763425	-123.0906046
V5W	Riley Park-Little Mountain, Kensington-Cedar Cottage, Oakridge, Sunset	49.2329609	-123.0912845
V5X	Oakridge, Marpole, Sunset	49.21517155	-123.0970026
V5Y	Mount Pleasant, Riley Park-Little Mountain	49.2512164	-123.1100756
V5Z	Fairview, South Cambie	49.25087085	-123.1213886
V6A	Strathcona, Chinatown, Downtown Eastside	49.2707128	-123.1007137
V6B	Downtown, Gastown, Harbour Centre, International Village, Victory Square, Yaletown	49.2747627	-123.1180934
V6C	Waterfront, Coal Harbour, Canada Place	49.28803365	-123.1147796
V6E	West End, Davie Village	49.2848351	-123.1262791
V6G	West End, Stanley Park	49.29972285	-123.1377908
V6H	Fairview, Granville Island, Shaughnessy	49.25774675	-123.1329205
V6J	Shaughnessy, Kitsilano, Quilchena	49.26256705	-123.1457577
V6K	Kitsilano, Greektown	49.26937685	-123.1650655
V6L	Arbutus Ridge, Dunbar-Southlands	49.2505265	-123.1658481
V6M	Shaughnessy, Oakridge, Kerrisdale, Arbutus Ridge	49.2378181	-123.1452847
V6N	Kerrisdale, Dunbar-Southlands, Musqueam	49.2379179	-123.1626348
V6P	Kerrisdale, Oakridge, Marpole	49.2260859	-123.1166419
V6R	Kitsilano, West Point Grey, Jericho	49.26740765	-123.1963317
V6T	UBC	49.2585106	-123.2266618
V6Z	Downtown	49.2780432	-123.1277779
V7X	Bentall Centre	49.2875369	-123.1203888
V7Y	Pacific Centre	49.28241235	-123.1192165

For simplicity, the neighborhoods with the same postal code are put together and there is a latitude and a longitude for each line.

With these previous dataset it would be possible to get Foursquare location data and make the analysis.

Foursquare [4] permitted to get the main venues related to the listed postal codes, analyze the data and finally decide which regions would be convenient for our new authentic Mexican cuisine restaurant.

### III. METHODOLOGY

A Canadian postal code is a six-character string that forms part of a postal address in Canada [3]. Postal codes beginning with V are located within the Canadian province of British Columbia [2], therefore Vancouver postal codes begin with V, as well as those for other cities in British Columbia. Our source in Wikipedia [2] contains all postal codes beginning with V. For this analysis we considered only the first three characters, hereafter named as **postal code**.

In some cities, like Vancouver, there is more than one neighborhood with the same postal code. And in some other cities there is no even one neighborhood specified. As it was mentioned before the initial data was obtained by scrapping the Wikipedia page using Beautiful Soup and we obtained a data frame containing Postal Code, City and Neighborhood.

*Table 2. Postal codes starting with V*

	PostalCode	City	Neighborhood
0	V1A	Kimberley	Not assigned
1	V2A	Penticton	Not assigned
2	V3A	Langley Township	Not assigned
3	V4A	Surrey	Not assigned
4	V5A	Burnaby	Government Road
5	V5A	Burnaby	Lake City
6	V5A	Burnaby	SFU
7	V5A	Burnaby	Burnaby Mountain
8	V6A	Vancouver	Strathcona
9	V6A	Vancouver	Chinatown

The data collected as mentioned above was cleaned up and processed before using it. To clean this data we filtered the rows following these rules: a) We only would take into account neighborhoods in Vancouver city. b) We would group neighborhoods, separated by a comma, with the same postal code per row. We obtained a data frame containing the Postal Code and the neighborhoods associated to that postal code (29 rows):

*Table 3. Postal codes and Neighborhoods in Vancouver*

	PostalCode	Neighborhood
0	V5K	Hastings-Sunrise
1	V5L	Grandview-Woodland
2	V5M	Hastings-Sunrise, Renfrew-Collingwood
3	V5N	Grandview-Woodland, Kensington-Cedar Cottage
4	V5P	Kensington-Cedar Cottage, Victoria-Fraserview

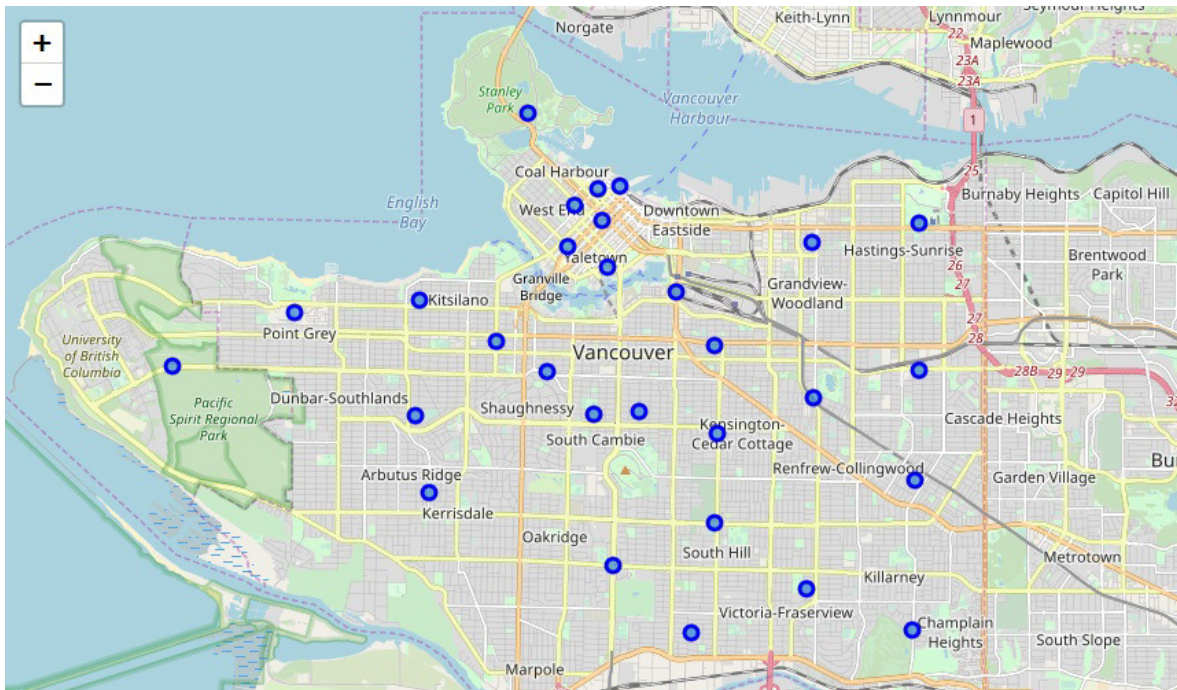
Using Nominatim from Geopy in Python, we found the location coordinates using the postal code as argument. Our generated dataset was contained in a data frame consisting of Postal Codes, Neighborhoods (or Neighborhood), Latitude and Longitude, for each row:

*Table 4. Postal codes, neighborhoods, latitudes, longitudes*

	PostalCode	Neighborhood	Latitude	Longitude
0	V5K	Hastings-Sunrise	49.281967	-123.040004
1	V5L	Grandview-Woodland	49.278839	-123.066843
2	V5M	Hastings-Sunrise, Renfrew-Collingwood	49.258053	-123.040155
3	V5N	Grandview-Woodland, Kensington-Cedar Cottage	49.253451	-123.066309
4	V5P	Kensington-Cedar Cottage, Victoria-Fraserview	49.222370	-123.068315

In the map shown below we can see the neighborhoods or groups of neighborhoods (spots associated to each postal code) that were considered for this analysis. For the purpose of this project we will refer to each spot or row in the data frame (Table 4) as **neighborhood**, and as **Vancouver region** to this zone.

*Map 1. Neighborhoods in Vancouver region.*

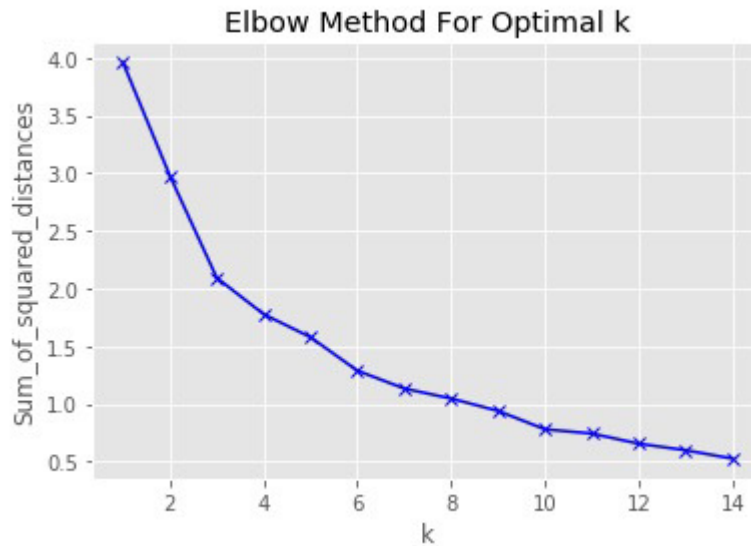


We used Fourquare to get the venues located in each neighborhood, within a 500 m radius and limited to 100 venues per neighborhood. We obtained 971 venues in all associated to 187 unique categories.

To decide which neighborhood would be appropriated for a new Mexican restaurant we implemented k-means machine learning algorithm. With this algorithm, clusters with groups of neighborhoods with similar characteristics were generated considering the venues located on each neighborhood. This would allow us to detect the different cluster's profiles, assign them a name, and choose our best neighborhood.

In order to determine the best k value for our k-means algorithm we used the elbow method that indicated us  $k = 6$  as the optimum value, as we can see on the graphic below:

*Graphic 1 . Elbow method to find k*



## IV. RESULTS

As a result of applying k-means algorithm and after analyzing its venues, we obtained the following clusters (see Appendix A for the full description of the clusters):

**Cluster 1:** Coffee and Asian Food Venues (11 neighborhoods)

**Cluster 2:** Park Venues (1 neighborhood)

**Cluster 3:** Trail Venues (1 neighborhood)

**Cluster 4:** Domestic Venues (1 neighborhood)

**Cluster 5:** Restaurant and Bakery Venues (1 neighborhood)

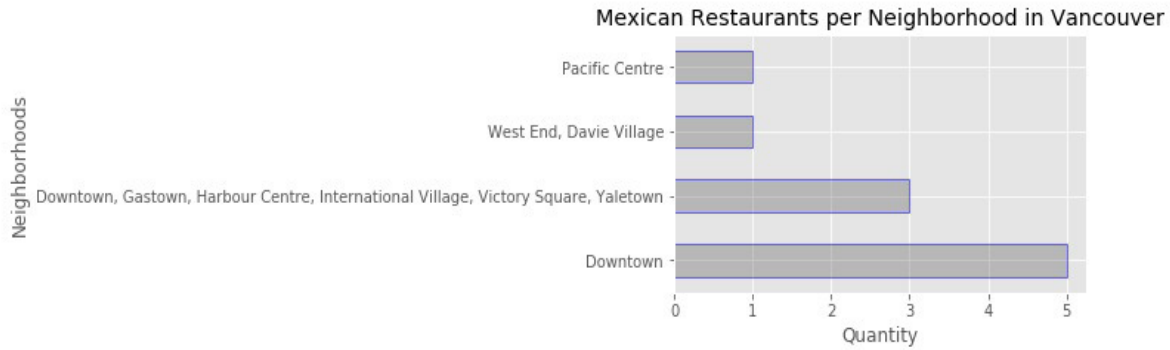
**Cluster 6:** Hotel and All World Cuisine Restaurants Venues (15 neighborhoods)

It can be seen that the cluster's names were assigned by observing the prevailing type of venues on each cluster. These observations allowed us to decide which cluster would be the best to look for our optimum neighborhood.

## V. DISCUSSION

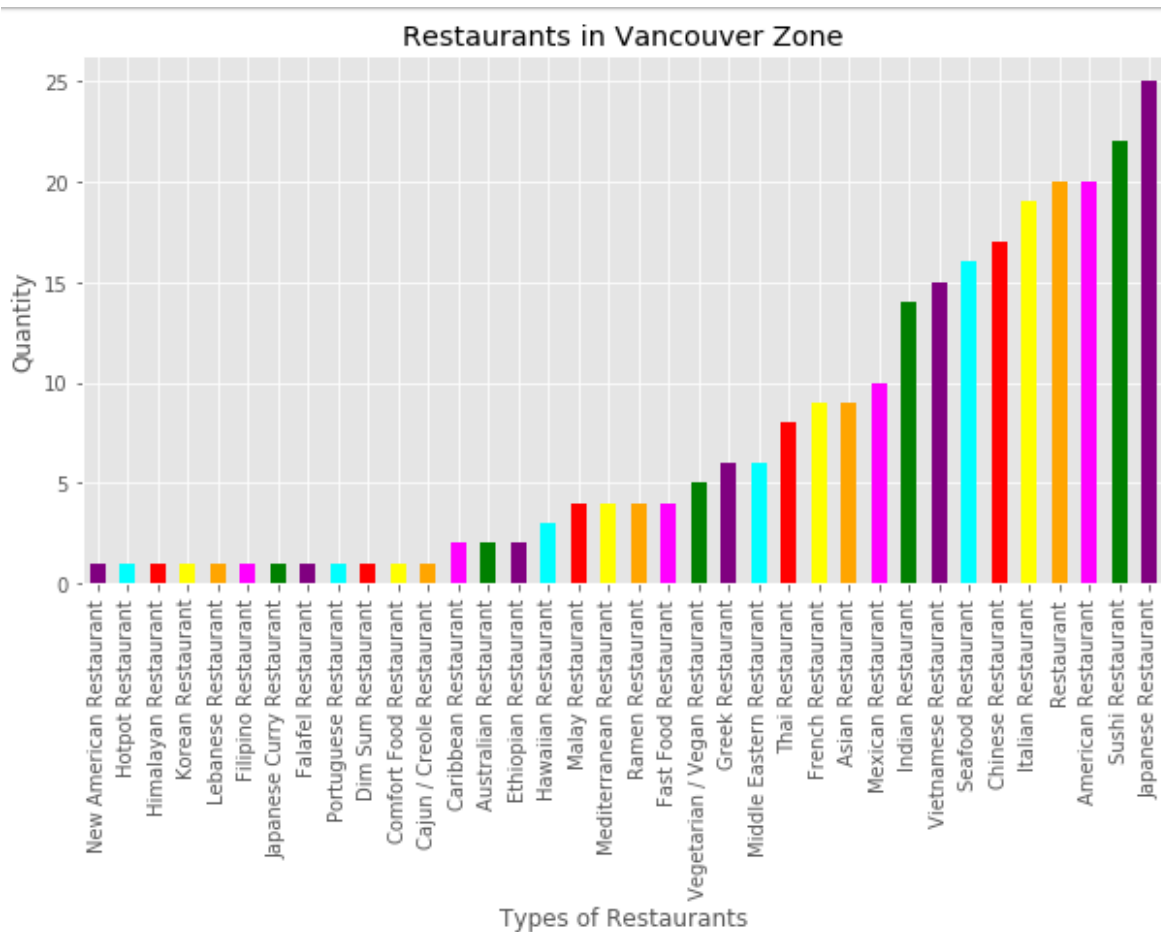
The following bar chart shows the amount of Mexican restaurants established in Vancouver region and the neighborhood where these are located:

**Graphic 2. Mexican Restaurants per Neighborhood in Vancouver region**



Additionally, the Graphic 3 indicates how many restaurants (of any kind) are working in Vancouver region.

**Graphic 3. Restaurants in Vancouver region**





As we can see there are 10 Mexican restaurants contrasting with 25 Japanese restaurants. With this we could think on a non-saturated market for our kind of restaurant.

After the clusters were generated we created a new data frame containing, additionally to the previous data, the top 10 most common venues. In the cases where there were less than 10 top common venues (for a neighborhood), this was indicated with a hyphen ‘-’ on the corresponding columns.

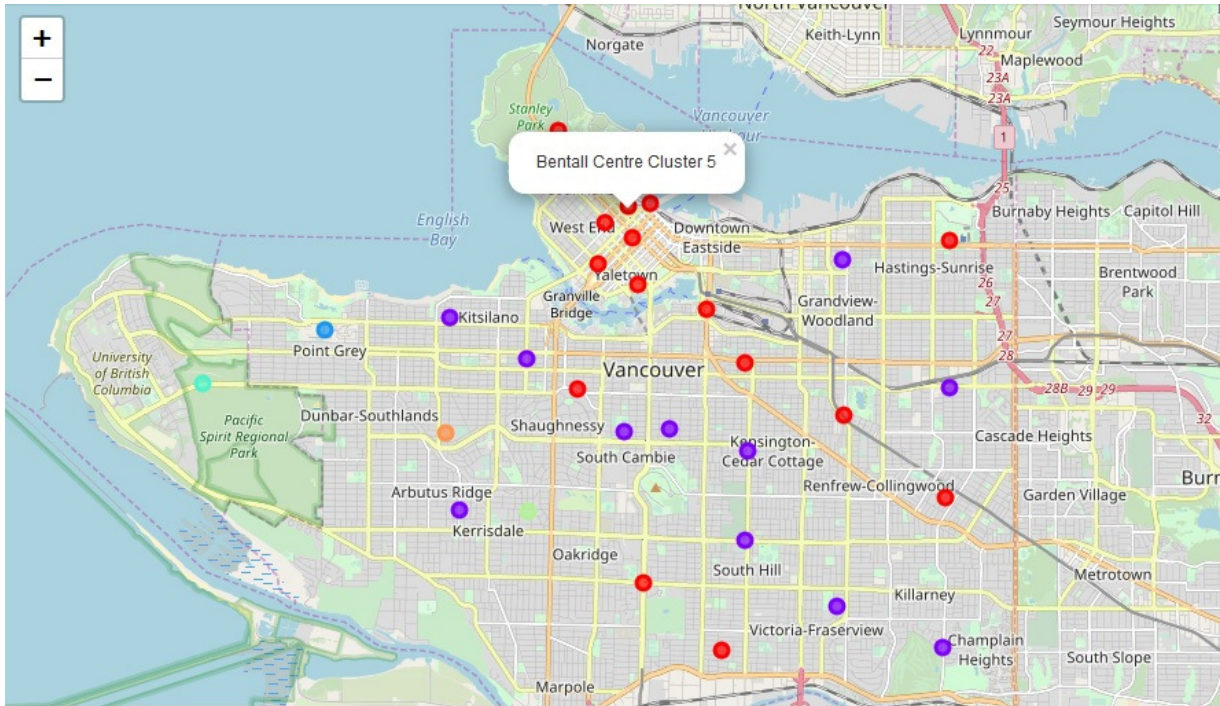
*Table 5. Cluster 6*

	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	Hastings-Sunrise	Event Space	Theme Park	Hockey Arena	Stadium	Theme Park Ride / Attraction	Sandwich Place	Market	Farm	Fair	Beer Garden
3	Grandview-Woodland, Kensington-Cedar Cottage	Café	Dog Run	Gym / Fitness Center	Child Care Service	Church	Lake	Bookstore	Concert Hall	Beach	Gym
5	Renfrew-Collingwood	Hotel	Asian Restaurant	Fish & Chips Shop	Bus Stop	Bar	Park	-	-	-	-
7	Mount Pleasant	Ethiopian Restaurant	Sushi Restaurant	Grocery Store	Pizza Place	Liquor Store	Market	Sandwich Place	Sports Bar	Japanese Restaurant	Pub
10	Oakridge, Marpole, Sunset	Park	History Museum	Indian Restaurant	Electronics Store	Bus Stop	-	-	-	-	-
13	Strathcona, Chinatown, Downtown Eastside	Brewery	Bakery	Ice Cream Shop	Gastropub	Coffee Shop	Beer Bar	Liquor Store	Café	Salad Place	Food Truck
14	Downtown, Gastown, Harbour Centre, International...	Hotel	Café	Restaurant	Italian Restaurant	Coffee Shop	Yoga Studio	Seafood Restaurant	Bakery	Vegetarian / Vegan Restaurant	Japanese Restaurant
15	Waterfront, Coal Harbour, Canada Place	Coffee Shop	Boat or Ferry	Café	Hotel	American Restaurant	Food Truck	Japanese Restaurant	Bar	Pizza Place	Seafood Restaurant
16	West End, Davie Village	Hotel	French Restaurant	Japanese Restaurant	Restaurant	Steakhouse	Food Truck	Italian Restaurant	Dessert Shop	Clothing Store	Café
17	West End, Stanley Park	Aquarium	Park	Theme Park Ride / Attraction	Zoo Exhibit	Bus Station	Garden	Exhibit	Event Space	Music Venue	Other Great Outdoors
18	Fairview, Granville Island, Shaughnessy	Breakfast Spot	Coffee Shop	Furniture / Home Store	Bakery	Bus Stop	Gourmet Shop	Lingerie Store	Cajun / Creole Restaurant	Café	Burger Joint
24	Kerrisdale, Oakridge, Marpole	Park	Bubble Tea Shop	Construction & Landscaping	Track	Golf Course	Bus Stop	-	-	-	-
27	Downtown	Hotel	Mexican Restaurant	Coffee Shop	Sushi Restaurant	Japanese Restaurant	Bakery	Café	Hostel	Diner	Gay Bar
28	Bentall Centre	Hotel	American Restaurant	Coffee Shop	Dessert Shop	Café	Restaurant	Steakhouse	Seafood Restaurant	Cosmetics Shop	Miscellaneous Shop
29	Pacific Centre	Hotel	Food Truck	Steakhouse	Seafood Restaurant	Coffee Shop	Clothing Store	Electronics Store	Café	Restaurant	Lounge

Observing the main characteristics of each cluster, initially we could consider Cluster 6 as the best cluster for our case, taking into account that its principal venues are hotels (we could think on tourism) and restaurants from the world (not only one or two type of cuisines prevailing). If we see again *Graphic 2. Mexican Restaurants per Neighborhood in Vancouver region*, and data in *Table 4. Cluster 6*, we can find that all Mexican restaurants in Vancouver region are located precisely in neighborhoods from Cluster 6. We also can observe that all this 4 neighborhoods (indexes 14, 16, 27 and 29) have “Hotel” as 1st Common Venue.

In accordance with the above we could consider like optimum a neighborhood from Cluster 6 and, if possible, a neighborhood with “Hotel” as 1st Common Venue. Therefore, our advice to the investors is to consider Bentall Centre as the optimum neighborhood to locate the new authentic Mexican cuisine restaurant

**Map . Clusters in Vancouver region.**



## VI. CONCLUSION

We used k-means machine learning algorithm to construct a classifier and applied it to a case of the real world. Finding the best zone to establish a business is a primary issue when starting a project like the one described in this document. This methodology can be improved by the use of more accurate and complete data, but still we consider were able to present a fairly good advice based on the developed process.

## VII. REFERENCES

- [1] Wikipedia: <https://en.wikipedia.org/wiki/Vancouver>
- [2] Wikipedia: [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_V](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_V)
- [3] [https://en.wikipedia.org/wiki/Postal\\_codes\\_in\\_Canada](https://en.wikipedia.org/wiki/Postal_codes_in_Canada)
- [4] <https://es.foursquare.com/>



# Appendix

### A. DESCRIPTION OF THE OBTAINED CLUSTERS

## A.1 Cluster 1

Cluster 1 is mainly conformed, as we can see, by Coffee Shop venues and Asian Restaurant venues. For this reason we named it “Coffee and Asian Restaurant Venues.

### Cluster 1: Coffee and Asian Food Venues (11 neighborhoods)

	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
1	Grandview-Woodland	Coffee Shop	Asian Restaurant	Pizza Place	Theater	Brewery	Italian Restaurant	Sushi Restaurant	Japanese Restaurant	Convenience Store	Breakfast Spot
2	Hastings-Sunrise, Renfrew-Collingwood	Coffee Shop	Grocery Store	Office	Japanese Restaurant	Sporting Goods Shop	Insurance Office	Sandwich Place	Bus Station	Pizza Place	Furniture / Home Store
4	Kensington-Cedar Cottage, Victoria-Fraserview	Pizza Place	Park	Indian Restaurant	Motorcycle Shop	Pet Store	Pharmacy	Convenience Store	Middle Eastern Restaurant	Sandwich Place	Café
6	Killarney	Shopping Mall	Juice Bar	Sushi Restaurant	Coffee Shop	Fast Food Restaurant	Salon / Barbershop	Sandwich Place	Chinese Restaurant	Bank	Bakery
8	Kensington-Cedar Cottage, Riley Park-Little Mo...	Coffee Shop	Bus Stop	Chinese Restaurant	Vietnamese Restaurant	Gourmet Shop	Greek Restaurant	Café	Sandwich Place	Filipino Restaurant	Supermarket
9	Riley Park-Little Mountain, Kensington-Cedar C...	Chinese Restaurant	Pizza Place	Bus Stop	Bakery	Diner	Fried Chicken Joint	Field	Fast Food Restaurant	Coffee Shop	Tea Room
11	Mount Pleasant, Riley Park-Little Mountain	Chinese Restaurant	Dessert Shop	Bakery	Sushi Restaurant	Grocery Store	Coffee Shop	-	-	-	-
12	Fairview, South Cambie	Sushi Restaurant	Park	Bus Stop	Chinese Restaurant	Greek Restaurant	Bubble Tea Shop	Coffee Shop	Sporting Goods Shop	Liquor Store	Gourmet Shop
19	Shaughnessy, Kitsilano, Quilchena	Coffee Shop	Thai Restaurant	American Restaurant	Pizza Place	Yoga Studio	Japanese Restaurant	Sushi Restaurant	Diner	Furniture / Home Store	Sporting Goods Shop
20	Kitsilano, Greektown	Coffee Shop	Park	Café	Toy / Game Store	Sushi Restaurant	Beach	Restaurant	Pub	Bus Stop	Food Truck
23	Kerrisdale, Dunbar-Southlands, Musqueam	Coffee Shop	Park	Convenience Store	Mediterranean Restaurant	Sporting Goods Shop	Fast Food Restaurant	Sushi Restaurant	Italian Restaurant	Pharmacy	Asian Restaurant

## A.2 Cluster 2

Cluster 2 only has one neighborhood (or group of neighborhoods) and one venue in that neighborhood. So the name assigned is the name of the venue.

### Cluster 2: Park Venues (1 neighborhood)

[illegible]

### A.3 Cluster 3

For Cluster 3 is the same case that the previous one. So the name assigned is the name of the venue.

**Cluster 3: Trail Venues (1 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
26	UBC	Trail	-	-	-	-	-	-	-	-	-

### A.4 Cluster 4

Cluster 4 only has one neighborhood (or group of neighborhoods) too, but some venues. We decided to name it “Domestic Venues”.

**Cluster 4: Domestic Venues (1 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
22	Shaughnessy, Oakridge, Kerrisdale, Arbutus Ridge	Home Service	Bus Stop	Moving Target	Electronics Store	-	-	-	-	-	-

### A.5 Cluster 5

Like the previous cluster, Cluster 5 only has one neighborhood (or group of neighborhoods) and some venues. The name for this venue is “Restaurant and Bakery Venues”.

**Cluster 5: Restaurant and Bakery Venues (1 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
21	Arbutus Ridge, Dunbar-Southlands	Caribbean Restaurant	Bakery	Italian Restaurant	-	-	-	-	-	-	-

## A.6 Cluster 6

The most representative venues on Cluster 6 are Hotel and all world cuisine Restaurant Venues. For this reason we named it “Hotel and All World Cuisine Restaurants Venues”.

**Cluster 6: Hotel and All World Cuisine Restaurants Venues (15 neighborhoods)**

	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	Hastings-Sunrise	Event Space	Theme Park	Hockey Arena	Stadium	Theme Park Ride / Attraction	Sandwich Place	Market	Farm	Fair	Beer Garden
3	Grandview-Woodland, Kensington-Cedar Cottage	Café	Dog Run	Gym / Fitness Center	Child Care Service	Church	Lake	Bookstore	Concert Hall	Beach	Gym
5	Renfrew-Collingwood	Hotel	Asian Restaurant	Fish & Chips Shop	Bus Stop	Bar	Park	-	-	-	-
7	Mount Pleasant	Ethiopian Restaurant	Sushi Restaurant	Grocery Store	Pizza Place	Liquor Store	Market	Sandwich Place	Sports Bar	Japanese Restaurant	Pub
10	Oakridge, Marpole, Sunset	Park	History Museum	Indian Restaurant	Electronics Store	Bus Stop	-	-	-	-	-
13	Strathcona, Chinatown, Downtown Eastside	Brewery	Bakery	Ice Cream Shop	Gastropub	Coffee Shop	Beer Bar	Liquor Store	Café	Salad Place	Food Truck
14	Downtown, Gastown, Harbour Centre, International...	Hotel	Café	Restaurant	Italian Restaurant	Coffee Shop	Yoga Studio	Seafood Restaurant	Bakery	Vegetarian / Vegan Restaurant	Japanese Restaurant
15	Waterfront, Coal Harbour, Canada Place	Coffee Shop	Boat or Ferry	Café	Hotel	American Restaurant	Food Truck	Japanese Restaurant	Bar	Pizza Place	Seafood Restaurant
16	West End, Davie Village	Hotel	French Restaurant	Japanese Restaurant	Restaurant	Steakhouse	Food Truck	Italian Restaurant	Dessert Shop	Clothing Store	Café
17	West End, Stanley Park	Aquarium	Park	Theme Park Ride / Attraction	Zoo Exhibit	Bus Station	Garden	Exhibit	Event Space	Music Venue	Other Great Outdoors
18	Fairview, Granville Island, Shaughnessy	Breakfast Spot	Coffee Shop	Furniture / Home Store	Bakery	Bus Stop	Gourmet Shop	Lingerie Store	Cajun / Creole Restaurant	Café	Burger Joint
24	Kerrisdale, Oakridge, Marpole	Park	Bubble Tea Shop	Construction & Landscaping	Track	Golf Course	Bus Stop	-	-	-	-
27	Downtown	Hotel	Mexican Restaurant	Coffee Shop	Sushi Restaurant	Japanese Restaurant	Bakery	Café	Hostel	Diner	Gay Bar
28	Bentall Centre	Hotel	American Restaurant	Coffee Shop	Dessert Shop	Café	Restaurant	Steakhouse	Seafood Restaurant	Cosmetics Shop	Miscellaneous Shop
29	Pacific Centre	Hotel	Food Truck	Steakhouse	Seafood Restaurant	Coffee Shop	Clothing Store	Electronics Store	Café	Restaurant	Lounge