

# **The Battle of the Neighborhoods – Open a Tea Room in Vancouver**

**Capstone Project – IBM Data Science Professional Certificate – Coursera**

**Alina Sansevich**

**June 2021**

## **Table of Contents**

- 1. Introduction
  - 1.1 Background
  - 1.2 Business Problem
- 2. Data
  - 2.1 Data sources
- 3. Methodology
  - 3.1 Web scraping and procuring geospatial data
  - 3.2 Exploring Vancouver's Neighborhoods
  - 3.3 Cluster Neighborhoods
    - 3.3.1 Elbow Method to determine the optimal value of k
- 4. Analysis
  - 4.1 Examine clusters
- 5. Results and Discussion
- 6. Conclusion
- 7. References

# 1. Introduction

## 1.1 Background

Vancouver is a major city in western Canada, located in the region of British Columbia. It is one of the most ethnically and linguistically diverse cities in Canada. Almost half of its residents are not native English speakers and belong to visible minority groups, mostly Asians. With its location on the Pacific Rim and at the western terminus of Canada's transcontinental highway and rail routes, Vancouver is one of the nation's largest industrial centres, and it's scenic location makes it a major tourist destination.

In many asian countries, a teahouse is traditionally a place which offers tea to its customers, alongside small plates of food. People gather at teahouses to chat, socialize and enjoy tea, and young people often meet at teahouses for dates. Tea rooms are also popular in Commonwealth countries, particularly Canada, with its harsh winters when afternoon tea is popular. The menu will generally have pastries, butter tarts or other small desserts.

All these factors combined make Vancouver a perfect city to open a Tea Room.

## 1.2 Business Problem

This project aims to find an optimal location for a "Tea Room & Pastries" shop. Specifically, **this report is targeted to stakeholders interested in opening a Tea Room in Vancouver, Canada.** The goal is to detect locations that are not already crowded with tea rooms and similar venues, taking into account the city's neighborhood demographic diversity and the proximity to tourist attractions to evaluate each location's potential attractiveness to customers.

# 2. Data

## 2.1 Data sources

The postal codes of British Columbia [1] were web scraped to find Vancouver's neighborhoods names and postal codes. To get the center coordinates of each postal code the Geopy library [2] was used. The data about venues (name, type and location) for each neighborhood was obtained using Foursquare API [3], and Vancouver's Wikipedia page [4], the city of Vancouver's website [5] and Tourism Vancouver's website [6] were consulted to gather information about the city and it's demographics.

### 3. Methodology

The following factors guided this analysis:

- The number of existing similar venues in the neighborhood (other tea rooms, breakfast spots, etc.).
- The proximity to tourist attractions.
- Each neighborhoods' demographics and its residential or commercial character.

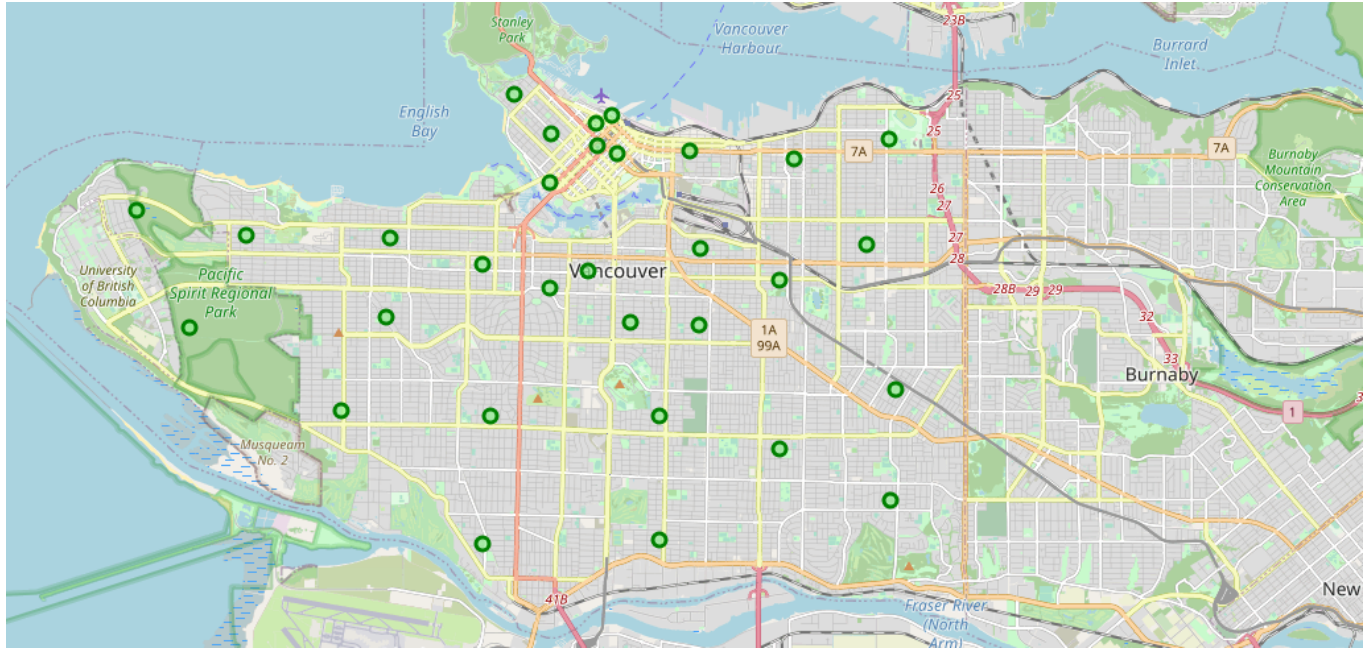
*Note: Vancouver has 31 postal codes assigned to different zones of the city. These postal codes describe better the city's areas (from a commercial and/or residential point of view) than the neighborhoods names, which have a more historical origin and are sometimes very small geographically. In my analysis I will use 'Neighborhoods', 'city areas' or 'postal code-defined areas' when referring to the different zones being analysed. When using the neighborhood's traditional names I will explicitly state so.*

#### 3.1 Web scraping and procuring geospatial data

Using BeautifulSoup and Geopy libraries the postal codes and names of Vancouver's neighborhoods and their location data were collected and combined into one table.

	PostalCode	City	Neighborhood	Latitude	Longitude
0	V5K	Vancouver	North Hastings-Sunrise	49.283277	-123.043293
1	V5L	Vancouver	North Grandview-Woodland	49.279853	-123.068277
2	V5M	Vancouver	South Hastings-Sunrise / North Renfrew-Colling...	49.265358	-123.049119
3	V5N	Vancouver	South Grandview-Woodland / NE Kensington-Cedar...	49.259254	-123.072063
4	V5P	Vancouver	SE Kensington-Cedar Cottage / Victoria-Fraserview	49.230668	-123.071843

The resulting dataframe had 31 rows, one for each postal code, meaning it contained all of Vancouver's neighborhoods. A map showing the city's postal code-defined areas was created using the Folium library to visualize the geospatial data obtained for the whole city (Map 1).



Map 1. Vancouver's neighborhoods.

### 3.2 Exploring Vancouver's Neighborhoods

The Foursquare API was then used to explore the neighborhoods. The limit was set at 100 venues and the radius at 1000 meters for each postal code from their given latitude and longitude coordinates. Data about all venue categories was requested. The results were saved in .csv format and read to create a new dataframe:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	North Hastings-Sunrise	49.283277	-123.043293	The Fair at the PNE	49.282971	-123.042109	Fair
1	North Hastings-Sunrise	49.283277	-123.043293	Livestock Barns	49.284037	-123.039278	Farm
2	North Hastings-Sunrise	49.283277	-123.043293	Pacific Coliseum	49.285823	-123.042727	Hockey Arena
3	North Hastings-Sunrise	49.283277	-123.043293	Tamam Fine Palestinian Cuisine	49.281070	-123.051438	Middle Eastern Restaurant
4	North Hastings-Sunrise	49.283277	-123.043293	Bao Chau	49.281232	-123.048971	Vietnamese Restaurant

The Foursquare API returned 2186 venues. The dataset had data for all 31 postal code-defined areas being considered. The maximum limit of 100 venues was returned for 14 postal code-defined areas, and there were another 10 postal code-defined areas for which the dataset contained over 40 venues, indicating that the information gathered was sufficient to complete the desired analysis. The Matplotlib and Seaborn libraries were applied to plot these results and better describe all findings (Figure 1). To visualize the locations of all venues returned by the Foursquare API a map was created (Map 2).

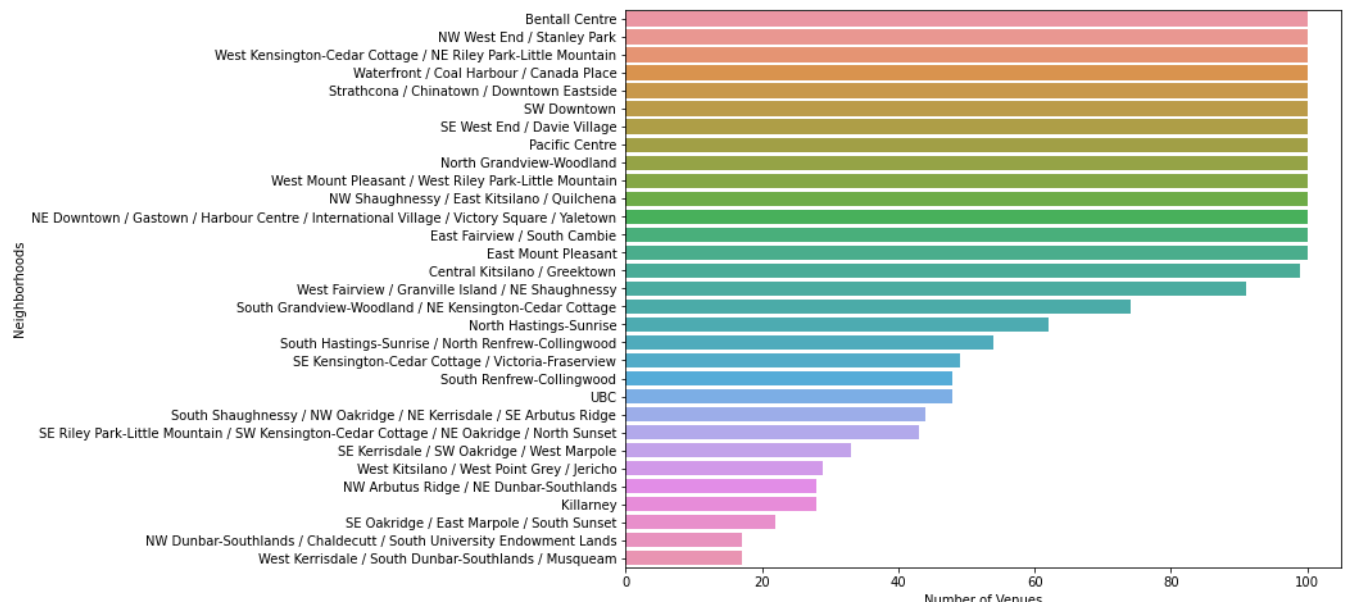


Figure 1. Number of venues per neighborhood returned by the Foursquare API.

These venues were grouped in 236 unique categories. In the following dataframe the 15 most frequent venue categories are shown:

Venue Category	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude
Coffee Shop	145	145	145	145	145	145
Café	80	80	80	80	80	80
Japanese Restaurant	76	76	76	76	76	76
Hotel	65	65	65	65	65	65
Park	64	64	64	64	64	64
Bakery	63	63	63	63	63	63
Sushi Restaurant	56	56	56	56	56	56
Chinese Restaurant	52	52	52	52	52	52
Vietnamese Restaurant	51	51	51	51	51	51
Restaurant	50	50	50	50	50	50
Pizza Place	48	48	48	48	48	48
Dessert Shop	42	42	42	42	42	42
Sandwich Place	40	40	40	40	40	40
Grocery Store	35	35	35	35	35	35
Breakfast Spot	34	34	34	34	34	34

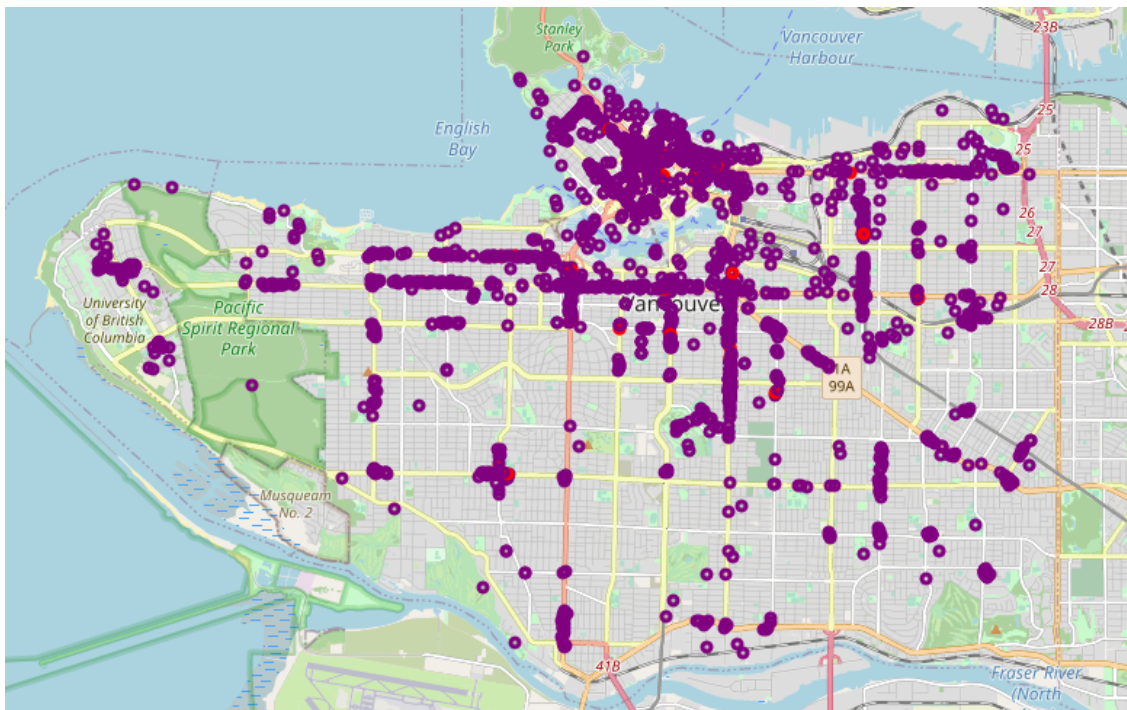
There were 14 venues in the 'Tea Room' category, 34 venues in the 'Breakfast Spot' category and 1 venue in the 'Creperie' category. These categories were selected to perform the

clustering analysis, since these types of venues could all be potential competitors for a new Tea Room. The decision was made to not include 'Coffee Shop' and related categories in this analysis because these type of venues were not considered as *direct* competitors for the following reasons:

When someone chooses to go to a Tea Room, that person is choosing:

1. Tea, not coffee.
2. The experience of drinking tea in the relaxed, friendly, usually cozy environment that is a Tea Room.
3. The pastries, sandwiches and/or desserts that usually are served along with the tea.

Instead, Coffee Shops are frequented by a more diverse public. Many people like to buy their coffee "on the go", others may stay in the coffee shop to study, work remotely, have an interview, or just hang out with friends, but the overall experience (and the main product: coffee) is different. So these types of venues were not considered as potential competitors when evaluating an area to open a Tea Room.



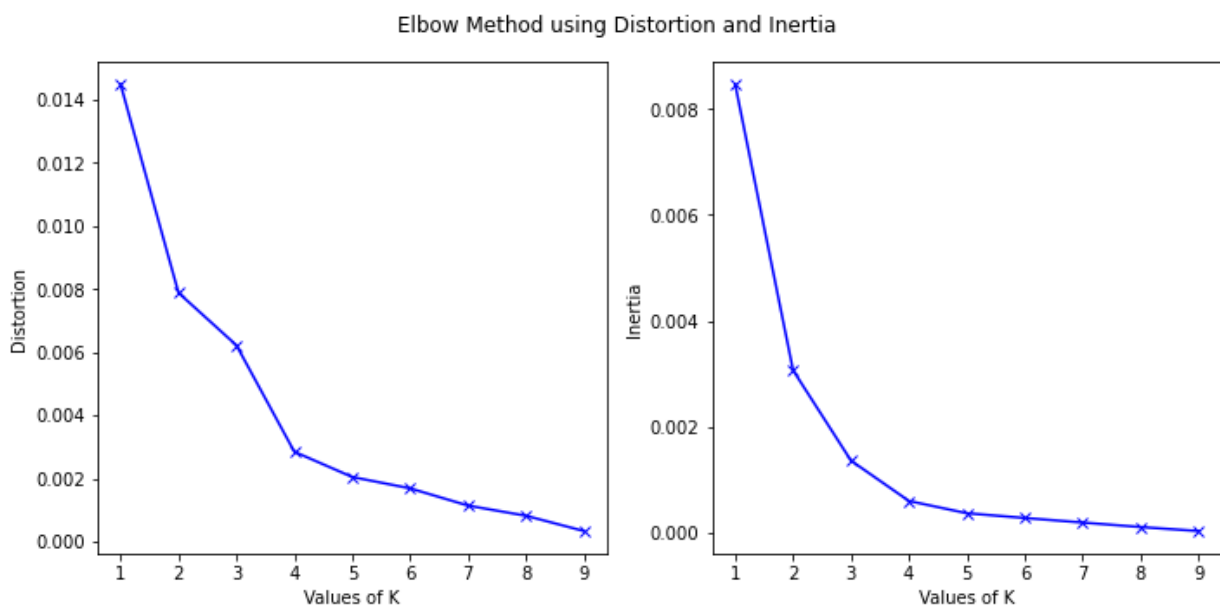
Map 2. Vancouver's venues. All the venues in this dataset are displayed here with a purple marker, except for the ones that belong to the 'Tea Room', 'Breakfast Spot' or 'Creperie' categories, that are shown with a red marker.

### 3.3 Cluster Neighborhoods

A clustering analysis was performed using the scikit-learn K-Means algorithm to find areas where Tea Rooms and similar venues are already located. For this, the categorical data about the venues was turned into numerical data using **One Hot encoding**, where for each neighborhood individual venues are turned into the frequency of how many of those venues are located in each neighborhood. Only the 'Tea Room', 'Breakfast Spot' or 'Creperie' venue categories data was used.

#### 3.3.1 Elbow Method to determine the optimal value of k

In order to start working with the k-Means algorithm it was necessary to determine the optimal number of clusters into which the data may be clustered. To do this we iterated the values of k from 1 to 9 and calculate the value of the distortions (*Distortion: the average of the squared distances from the cluster centers of the respective clusters*) and inertia (*Inertia: the sum of squared distances of samples to their closest cluster center*) for each value of k in the given range. The optimal number of clusters would be the value of k at the "elbow" (the point after which the distortion/inertia starts decreasing in a linear fashion).



**Figure 2. Elbow Method to determine the optimal value of k. The left plot shows the result using Distortion and the right plot shows the result using Inertia.**

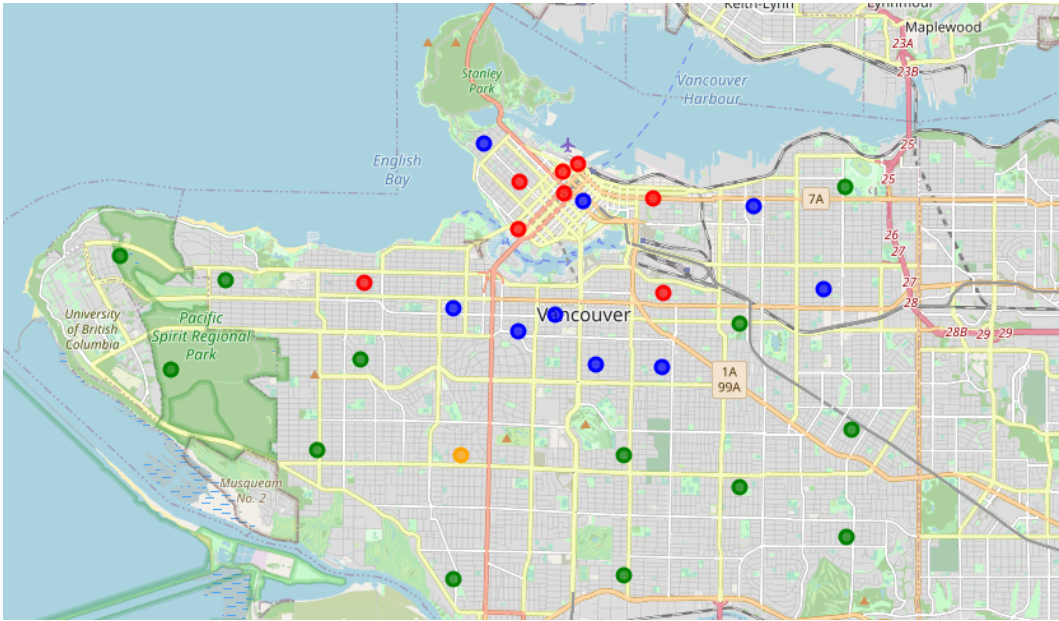
As shown in the plots, for this dataset the optimal value of k was 4. The k-Means algorithm was used with this value to create the clusters and continue with the analysis. The following dataframe includes the cluster labels obtained as well as the frequencies of the tea-related venues for each neighborhood:



	PostalCode	City	Neighborhood	Latitude	Longitude	Cluster Labels	Breakfast Spot	Tea Room	Creperie
0	V5K	Vancouver	North Hastings-Sunrise	49.283277	-123.043293	2	0.000000	0.00	0.0
1	V5L	Vancouver	North Grandview-Woodland	49.279853	-123.068277	1	0.030000	0.00	0.0
2	V5M	Vancouver	South Hastings-Sunrise / North Renfrew-Colling...	49.265358	-123.049119	1	0.037037	0.00	0.0
3	V5N	Vancouver	South Grandview-Woodland / NE Kensington-Cedar...	49.259254	-123.072063	2	0.000000	0.00	0.0
4	V5P	Vancouver	SE Kensington-Cedar Cottage / Victoria-Fraserview	49.230668	-123.071843	2	0.000000	0.00	0.0
...	...	...	...	...	...	...	...	...	...
26	V6S	Vancouver	NW Dunbar-Southlands / Chaldecutt / South Univ...	49.251210	-123.225418	2	0.000000	0.00	0.0
27	V6T	Vancouver	UBC	49.271268	-123.239269	2	0.000000	0.00	0.0
28	V6Z	Vancouver	SW Downtown	49.275885	-123.131725	0	0.000000	0.01	0.0
29	V7X	Vancouver	Bentall Centre	49.285967	-123.119531	0	0.010000	0.01	0.0
30	V7Y	Vancouver	Pacific Centre	49.282182	-123.119247	0	0.010000	0.01	0.0

#### 4. Analysis

To visualize the clusters, a map was created assigning different colors to each cluster's markers (Map 3).



Map 3. K-Means Clustering of Vancouver's Neighborhoods. Cluster 0 is shown with red markers, Cluster 1 is shown with blue markers, Cluster 2 is shown with green markers and Cluster 3 is shown with orange markers.



## 4.1 Examine clusters

At first glance, I found that:

- The southern and western areas of Vancouver all belong to one cluster (Cluster 2 - green).
- The northern areas are segmented in 2 clusters (Cluster 0 - red, and Cluster 1 - blue).
- Finally, the last cluster has only one area in it (Cluster 3 - orange).

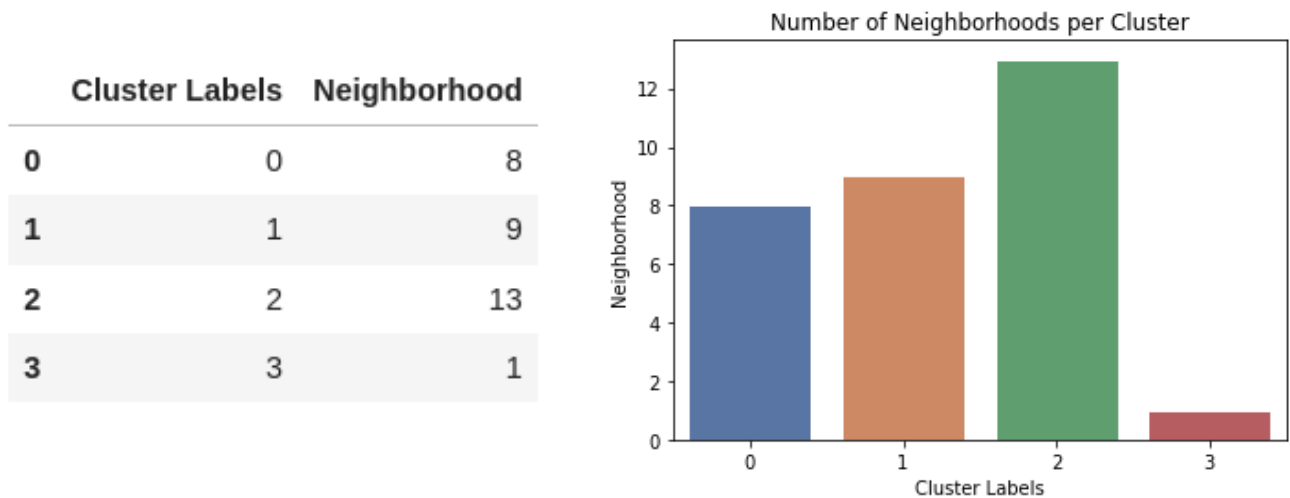


Figure 3. Number of neighborhoods grouped in each cluster.

I found that Cluster 3 had only one neighborhood (postal code area), there were 13 neighborhoods (postal codes areas) that had no "Tea&Pastries" shops (Cluster 2), Cluster 0 had 8 neighborhoods (postal codes areas) and Cluster 1 had 9 neighborhoods (postal codes areas).



Figure 4. Frequency of Tea & Pastries venues for each cluster.

From Figure 4, I could see that:

- There were no "Tea & Pastries" shops in neighborhoods grouped in Cluster 2 (shown green in Map 3).
- Cluster 1 was the cluster with more Breakfast Spots, followed by Cluster 0 (shown in Map 3 with colors blue and red respectively).
- There is only one Creperie (in Cluster 1), so this venue category wouldn't represent a potentially big competitor.
- Cluster 3 has the highest frequency of Tea Rooms and no Breakfast Spots.

In order to find good locations for a new Tea Room, the number of 'Tea & Pastries' venues present in each cluster was analysed by creating a pivot table:

Cluster Labels	Venue Category Neighborhood	Breakfast Spot	Creperie	Tea Room
0	Bentall Centre	1.0	0.0	1.0
	Central Kitsilano / Greentown	1.0	0.0	1.0
	East Mount Pleasant	1.0	0.0	1.0
	Pacific Centre	1.0	0.0	1.0
	SE West End / Davie Village	1.0	0.0	0.0
	SW Downtown	0.0	0.0	1.0
	Strathcona / Chinatown / Downtown Eastside	1.0	0.0	1.0
	Waterfront / Coal Harbour / Canada Place	1.0	0.0	1.0
	East Fairview / South Cambie	3.0	0.0	0.0
1	NE Downtown / Gastown / Harbour Centre / International Village / Victory Square / Yaletown	3.0	0.0	1.0
	NW Shaughnessy / East Kitsilano / Quilchena	4.0	0.0	1.0
	NW West End / Stanley Park	3.0	0.0	1.0
	North Grandview-Woodland	3.0	0.0	0.0
	South Hastings-Sunrise / North Renfrew-Collingwood	2.0	0.0	0.0
	West Fairview / Granville Island / NE Shaughnessy	3.0	0.0	0.0
	West Kensington-Cedar Cottage / NE Riley Park-Little Mountain	3.0	0.0	1.0
	West Mount Pleasant / West Riley Park-Little Mountain	3.0	1.0	1.0
	South Shaughnessy / NW Oakridge / NE Kerrisdale / SE Arbutus Ridge	0.0	0.0	2.0
3				

The pivot table shows how many venues of each category of interest can be found in each area. To analyze the results for each cluster in more detail I created bar plots with the number of venues per postal code area.

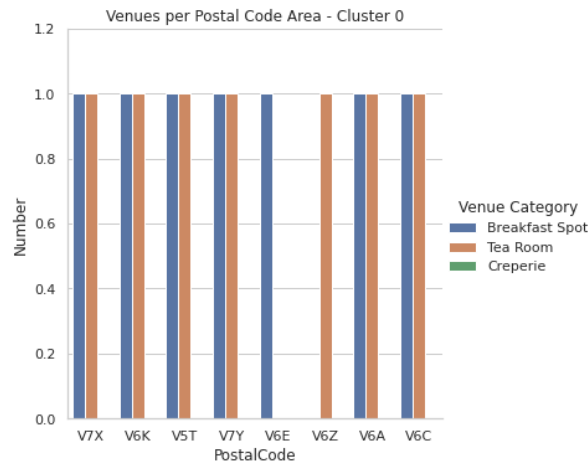


Figure 5. Number of venues per postal code area for Cluster 0.

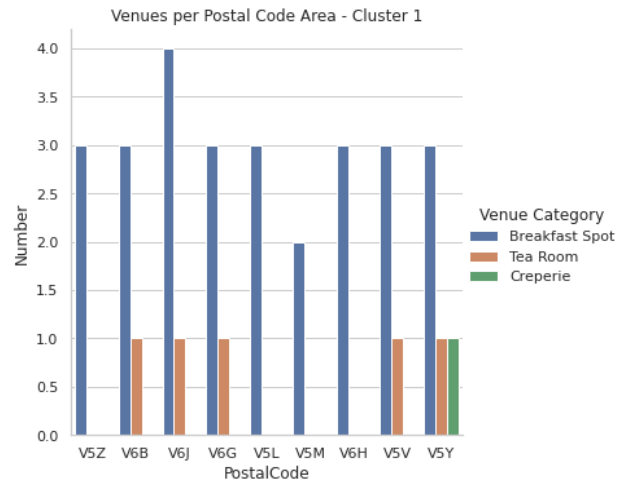


Figure 6. Number of venues per postal code area for Cluster 1.

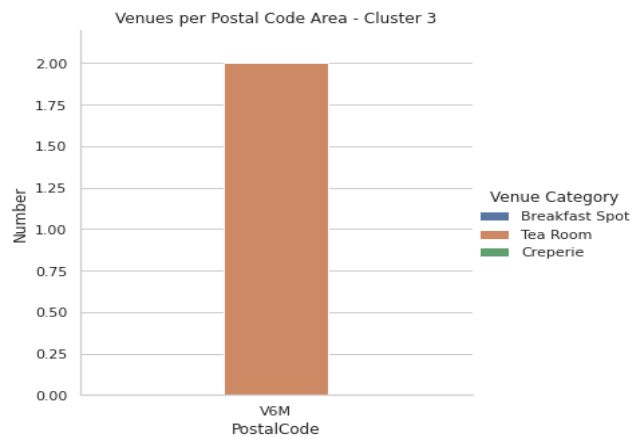
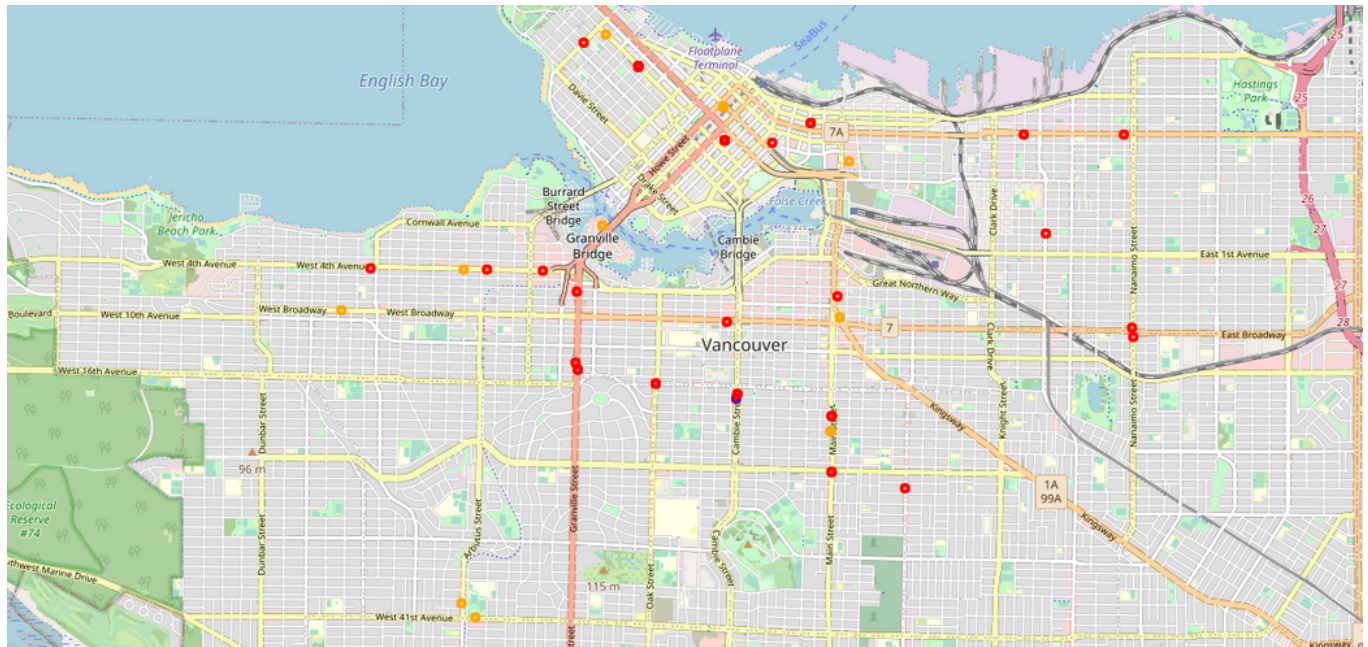


Figure 7. Number of venues per postal code area for Cluster 3.

Cluster 0 had almost one 'Breakfast Spot' and one 'Tea Room' per postal code. In Cluster 1 I found many more 'Breakfast Spots' than 'Tea Rooms', and there were 4 postal codes that had no 'Tea Rooms'. Finally, in Cluster 3 there were only two Tea Rooms, and no 'Breakfast Spots' nor 'Creperies'. A map was created to visualize the venues belonging to the categories of interest:



**Map 4.** 'Tea & Pastries' venues of Vancouver. 'Breakfast Spots' are shown in red, 'Tea Rooms' are shown in orange and 'Creperie' is shown in purple.

As shown in Map 4, I found that most Tea Rooms in Vancouver were spread out between neighborhoods. The only exception to this were the two Tea Rooms found on Cluster 3, which were located just around the corner from each other. To evaluate each area in more detail and select good locations for a new Tea Room, I created a dataframe with the 10 most common venues in each neighborhood belonging to the 3 clusters that already had tea related venues:

	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	Bentall Centre	Hotel	Dessert Shop	American Restaurant	Café	Restaurant	Coffee Shop	Cosmetics Shop	Japanese Restaurant	Steakhouse	Plaza
1	Central Kitsilano / Greektown	Coffee Shop	Café	Sushi Restaurant	Pub	Japanese Restaurant	Grocery Store	Vegetarian / Vegan Restaurant	Park	Bakery	Convenience Store
2	East Fairview / South Cambie	Coffee Shop	Japanese Restaurant	Café	Park	Pharmacy	Sushi Restaurant	Chinese Restaurant	Mexican Restaurant	Indian Restaurant	Breakfast Spot
3	East Mount Pleasant	Brewery	Coffee Shop	Bakery	Vietnamese Restaurant	Pizza Place	Mexican Restaurant	Ice Cream Shop	Taco Place	Sushi Restaurant	Music Venue
4	Killarney	Grocery Store	Shopping Mall	Bus Stop	Pharmacy	Coffee Shop	Italian Restaurant	Restaurant	Burger Joint	Cosmetics Shop	Liquor Store
5	NE Downtown / Gastown / Harbour Centre / Inter...	Hotel	Coffee Shop	Taco Place	Restaurant	Sandwich Place	Concert Hall	Dessert Shop	Café	Italian Restaurant	Breakfast Spot
6	NW Arbutus Ridge / NE Dunbar-Southlands	Italian Restaurant	Japanese Restaurant	Sushi Restaurant	Bank	Coffee Shop	Dessert Shop	Pharmacy	Pizza Place	Park	Sandwich Place
7	NW Dunbar-Southlands / Chaldecutt / South Univ...	Shopping Mall	Bakery	Noodle House	Coffee Shop	Frozen Yogurt Shop	Sushi Restaurant	Farmers Market	Park	Supermarket	Bank
8	NW Shaughnessy / East Kitsilano / Quilchena	Coffee Shop	Bakery	Furniture / Home Store	Restaurant	Japanese Restaurant	Breakfast Spot	Yoga Studio	Vegetarian / Vegan Restaurant	French Restaurant	Grocery Store
9	NW West End / Stanley Park	Coffee Shop	Dessert Shop	Japanese Restaurant	Café	Sushi Restaurant	Korean Restaurant	Sculpture Garden	Grocery Store	Ramen Restaurant	Vietnamese Restaurant

As a last step in my analysis, I checked out what other venues were present in the areas assigned to postal codes V6E (from Cluster 0), V5Z, V5L, V5M and V6H (from Cluster 1) that have no Tea Rooms.

### Top 10 venues in postal code V6E (Cluster 0):

neighborhoods_venues_sorted[neighborhoods_venues_sorted['PostalCode'] == 'V6E']													
	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	PostalCode	City
17	SE West End / Davie Village	Hotel	Japanese Restaurant	Dessert Shop	Bakery	Sushi Restaurant	Ramen Restaurant	Restaurant	Café	Sandwich Place	Park	V6E	Vancouver

### Top 10 venues in postal codes V5Z, V5L, V5M and V6H (Cluster 1):

neighborhoods_venues_sorted[neighborhoods_venues_sorted['PostalCode'] == 'V5Z']													
	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	PostalCode	City
2	East Fairview / South Cambie	Coffee Shop	Japanese Restaurant	Café	Park	Pharmacy	Sushi Restaurant	Chinese Restaurant	Mexican Restaurant	Indian Restaurant	Breakfast Spot	V5Z	Vancouver

neighborhoods_venues_sorted[neighborhoods_venues_sorted['PostalCode'] == 'V5L']													
	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	PostalCode	City
10	North Grandview-Woodland	Coffee Shop	Brewery	Café	Pizza Place	Sushi Restaurant	Grocery Store	Japanese Restaurant	Italian Restaurant	Indian Restaurant	Breakfast Spot	V5L	Vancouver

neighborhoods_venues_sorted[neighborhoods_venues_sorted['PostalCode'] == 'V5M']													
	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	PostalCode	City
20	South Hastings-Sunrise / North Renfrew-Collingwood	Pizza Place	Vietnamese Restaurant	Furniture / Home Store	Breakfast Spot	Chinese Restaurant	Bus Stop	Sandwich Place	Park	Dim Sum Restaurant	Pet Store	V5M	Vancouver

neighborhoods_venues_sorted[neighborhoods_venues_sorted['PostalCode'] == 'V6H']													
	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	PostalCode	City
26	West Fairview / Granville Island / NE Shaughnessy	Coffee Shop	Furniture / Home Store	Japanese Restaurant	Bank	Sandwich Place	Café	Breakfast Spot	Pizza Place	Pharmacy	Park	V6H	Vancouver

### Finally, the top 10 venues in postal code V6M (Cluster 3):

neighborhoods_venues_sorted[neighborhoods_venues_sorted['PostalCode'] == 'V6M']													
	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	PostalCode	City
22	South Shaughnessy / NW Oakridge / NE Kerrisdal	Chinese Restaurant	Coffee Shop	Sushi Restaurant	Japanese Restaurant	Bus Stop	Asian Restaurant	Vietnamese Restaurant	Pharmacy	Tea Room	Dessert Shop	V6M	Vancouver

## 5. Results and Discussion

My analysis shows that although there are many Tea Rooms in Vancouver, they are spread out over the city's neighborhoods. As shown in Map 4, there are areas with no Tea Rooms fairly close to the city center.

The k-Means algorithm found four clusters that clearly segment the city according to each areas' inherent characteristics. Vancouver is a city that has grown from an early settlement in the peninsula, in what is now the "Gastown" neighborhood. The northern areas of the city closest to the peninsula, and the peninsula itself, are home to Vancouver's commercial and financial districts, and most of the city's tourist attractions. The k-Means algorithm has found 2 clusters in this area, Cluster 0 and Cluster 1. Here I found all "Breakfast Spots" and most of the city's Tea Rooms (see Map 4).

The areas included in Cluster 2 are mostly residential. The "western half" of this cluster includes some of the most affluent residential neighborhoods, characterized by larger than average lot sizes, with stately homes on tree-lined streets. The "southern half" of Cluster 2 covers some of the most ethnically diverse neighborhoods in Vancouver (see Map 3). These are also mostly residential areas that used to be predominantly working class but rising house prices have made these areas increasingly popular with young professionals and families who can no longer afford houses in areas such as the west side. These areas have become an attractive location for young professionals and artists, and have received an influx of immigrants from China and other South-East Asian countries. In several neighborhoods included in Cluster 2 only a portion of residents speak English as a mother tongue, compared to roughly half of residents in Vancouver as a whole. In general, we could say that Cluster 2 groups areas of Vancouver that are generally not a popular destination for tourists. The neighborhoods included in this cluster are probably not the best locations to open a new Tea Room and for this reason I have not explored them any further.

Cluster 3 includes only one postal code defined area. The neighborhoods in this area are almost-entirely residential, with high average annual household income and the highest average house price in Vancouver. It is also the site of many historical homes, with elegant tree-lined streets and large properties built before or during World War II. The neighborhoods in this postal code are mostly British in character and demographics, and have received an influx of affluent immigrants from Hong Kong. As Cluster 2, this area also is generally not a popular destination for tourists. It does have Tea Rooms, possibly frequented mostly by its residents.

As shown in Figures 5, 6 and 7, there are postal code defined-areas that have no Tea Rooms, so I decided to analyse those areas in more detail to find the best possible location for a new Tea Room:

*Cluster 0 – V6E – SE West End / Davie Village:*

In these neighborhoods, the most common venues (after Hotels) are mainly restaurants. This area has only one 'Breakfast Spot' that is not in the top 10. This could be an area to explore by stakeholders. As an argument against this location, Davie Village may not be the most likely area where people would be interested in visiting a Tea Room. This is "Vancouver's gayborhood", the LGBT-friendly neighborhood that in the evening, comes to life as the real epicentre of Vancouver's gay nightlife scene. The target audience for a Tea Room may just not be there.

*Cluster 1 – V5Z – East Fairview / South Cambie:*

Here we find that coffee shops are the 1st and 3rd most common venues, and "Breakfast Spot" can be found only on the 10th place, with 3 venues in this area (see Figure 7). This may be a good area for stakeholders to explore. It is close to one of the city's largest parks and is known for a large cluster of medical facilities.

*Cluster 1 – V5M – South Hastings–Sunrise / North Renfrew–Collingwood:*

Here we find that 'Breakfast Spot' is the 4th most common venue (so there are possibly more competitors for a new Tea Room), with 2 venues in the area (see Figure 7). The 3rd place is "Furniture/Home Store", the 6th place is "Bus Stop" and the 10th place is "Pet Store". All this means that this area is not so visited by tourists. Since it is located to the east side of Vancouver further away from Downtown and all the tourist attractions, it may not be so interesting for stakeholders.

*Cluster 1 – V6H – West Fairview / Granville Island / NE Shaughnessy:*

'Breakfast Spot' is the 7th most common venue, with 3 venues in the area (see Figure 7). This area is located at the very center of Vancouver's downtown, close to many tourist attractions, so it may be a very interesting area to explore, since there are no Tea Rooms in it.

*Cluster 1 – V5L – North Grandview–Woodland:*

In this case we find that 'Breakfast Spot' is again in the 10th place, with 3 venues in the area (see Figure 7). On the 6th place we find the 'Grocery Store' category, which indicates that this area has more residents living in it. Since it is a commercial and residential area relatively close to Downtown, this could be a very interesting area for stakeholders. "Commercial Drive" is located in this neighborhood. It is a street packed with small shops that has been a key part of the commercial landscape in Vancouver for nearly a century for its ethnic diversity



and the uniqueness of shops that populate it. It is probably the best candidate location to explore in order to open a new Tea Room.

*Cluster 3 - V6M - South Shaughnessy / NW Oakridge / NE Kerrisdale / SE Arbutus Ridge:*

Finally, in Cluster 3 we find that 'Tea Room' is the 9th most common venue category, with 2 venues in the area (see Figure 8). This is one of the most affluent areas of Vancouver, but it is mainly residential, so its commercial streets (where the two Tea Rooms are already located) may be too "crowded" to be of interest to stakeholders. Still, it could be good to explore it further before discarding it as an option for a new Tea Room.

The results presented in this study do not imply that these areas are actually the optimal locations for a new Tea Room. This is a preliminary analysis to find areas that are not crowded with existing venues that could be competitors for a new Tea Room. The recommended areas should therefore be considered only as a starting point for a more detailed analysis which could eventually result in finding the optimal location after other factors (like rental price, availability of business premises and city planning regulations to name a few) are taken into account.

Keeping all this in mind, after performing this analysis I would recommend the stakeholders to further explore the following locations:

**V5L - North Grandview-Woodland** → may be visited by residents and tourists, "Commercial Drive", proximity to Downtown.

**V5Z - East Fairview / South Cambie** → proximity to Downtown, not many competitors.

**V6H - West Fairview / Granville Island / NE Shaughnessy** → close to many areas of interest, not many competitors.

**V6M - South Shaughnessy / NW Oakridge / NE Kerrisdale / SE Arbutus Ridge** → not the first option but could have potential if further analysis were to show that there is enough market demand among the areas' residents.

## 6. Conclusion

The purpose of this project was to identify an optimal location for a "Tea Room & Pastries" shop in Vancouver, Canada. After obtaining the geolocation data of the city's neighborhoods and venue data from Foursquare API, a k-Means clustering analysis was performed to group the city's neighborhoods according to this goal. The analysis of the four clusters obtained revealed areas where no Tea Rooms are located. These areas were studied in more detail and

as a result, four areas of Vancouver can be presented to potential stakeholders as the most promising neighborhoods to open a Tea Room in Vancouver.

## **7. References**

- [1] The postal codes of British Columbia  
([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] Geopy (<https://pypi.org/project/geopy/>)
- [3] Foursquare (<https://foursquare.com/>)
- [4] Vancouver's Wikipedia page (<https://en.wikipedia.org/wiki/Vancouver>)
- [5] The city of Vancouver's website (<https://vancouver.ca/>)
- [6] Tourism Vancouver's website (<https://www.tourismvancouver.com/>)