IBM Data Science Capstone Project

(IBM Data Science Professional Certificate, Coursera)

Rent Prices and Social Environment Analysis of Vancouver Neighbourhoods

Prepared by: Beibyt Azymbek

February 2020

Table of Contents

1. Introduction

1.1 Background

Vancouver is the largest city in the province of British Columbia, and the eighth largest municipality in Canada. The Greater Vancouver metropolitan area (which includes neighbouring cities such as Burnaby, Richmond, and Surrey) is the third largest in Canada. The City is famous for its beautiful surrounding nature and vibrant lifestyle it provides for the habitants [1]. With its scenic views, mild climate, and friendly people, Vancouver is known around the world as both a popular tourist attraction and one of the best places to live.

Vancouver has now also become one of the fastest growing cities in the world by attracting large populations of internal and international migrants. According to the World Population Review, population of Greater Vancouver metropolitan area has risen from 1.96 million in 1990 to 2.58 million in 2020 [2].

Due to its geography, being surrounded by the Pacific Ocean from West and Rocky Mountains on East, Vancouver has limited space for its expansion and being able to accommodate its growing population, which reasonably leads to high housing and living expenses. With the size of only 114 square kilometers, City of Vancouver is one of most expensive to live cities in Canada and in the world, especially in terms of accommodation and real estate. According to the 15th Annual Demographia International Housing Affordability Survey, Vancouver has been listed as 2nd least affordable city in the world as of 2019 [3].

1.2 Problem

Above stated issue has now become a major concern for not only current habitants of the area but prospective migrants as well, who seek affordable rental living when considering relocation. Usually a lot of time and effort is spent on deciding what neighbourhood in the City to settle in. Work conducted in this paper is a data-driven analytical approach of comparing different neighbourhoods of Vancouver in terms of their rental prices and social environments. As part of the analysis, this work aims to identify what factors, specifically what types of venues and their presence in a neighbourhood affect rental prices.

1.3 Interest

The project will be interesting for people who plan renting accommodation in Vancouver or generally consider moving to the City. By providing analytics and comparison on the neighbourhoods, results of this work would assist in finding optimal neighbourhood in accordance with the needs and budget of those who plan to live there. I believe that results of this research will help individuals to find an accommodation in a suitable neighbourhood with all essential infrastructure and venues within their budget and be able to make their lives more comfortable.

2. Data

2.1 Data Sources and Description

The following data sources and methods on obtaining them were used within this work:

- Vancouver Rental Market Statistics Summary provided by Zones from Canada Mortgage and Housing Corporation [4]. This is a database on housing and mortgage information managed by a state corporation, which provides information in interactive as well as in textual format (csv). Within this work, I have downloaded csv file for Vancouver housing market data specifically for 1-bedroom residences, which is one of the most common residence sizes rented in highly priced cities such as Vancouver. There is obviously data available for larger (2+ bedroom) residences which can easily be obtained and similar analysis could be conducted on. This source provided both list of neighbourhoods of Vancouver as well as average rental price for 1-bedroom residences within those neighbourhoods.
- Geocoder Python package was used to obtain geographical coordinates of the neighbourhoods. This information would allow us to pass coordinates as parameters when creating neighbourhood maps and collecting information on venues within those neighbourhoods.
- Foursquare Application Programming Interface (API) was used to gather information on venues, categories of venues, their location, quantity and quality (based on user reviews) within each of the neighbourhoods of Vancouver [5]. This information was used to obtain and classify social environmental background of the neighbourhoods and identify potential factors which might directly impact on the rent prices in that neighbourhood.

2.2 Data Cleansing and Preparation

Following data cleansing and preparation activities have been performed before further processing:

- Data on rental market (csv file) had too many unnecessary details which had to be cleansed before proceeding with analysis. For example, vacancy and availability rates, information on number of available units was not relevant within our subject of work and had to be removed from for further processing.
- Data at hand was simplified to cover only those related to 1-bedroom residences as usually they are enough to provide indicators on the associated neighbourhood prices as any other larger residence types, i.e. if renting 1-bedroom apartment in neighbourhood A would cost higher than renting same 1-bedroom apartment in neighbourhood B, high chances are the same will be true for 2 and 3-bedroom apartments as well.
- Some column names or names of the Vancouver neighbourhoods had unnecessary characters, such as commas or brackets, which would cause false results for classification algorithms.
- Some data types of captured information, such as average rent price had to be converter from object (string) into integer.

3. Methodology

Our methodology will analyze Vancouver neighbourhoods on their average rental prices and social environments these neighbourhoods provide for their habitants. Specifically, we will go through following steps:

1. Identify list of Vancouver neighbourhoods and obtain average rent prices for these neighbourhoods
2. Get neighbourhood coordinates and map neighbourhoods to have visual understanding of the region
3. Cluster neighbourhoods based on average rent prices and conduct preliminary analysis
4. Gather information on venues for each of the neighbourhoods
5. Analyze venues in the neighbourhoods by identifying the most common venue categories in each of the neighbourhoods to understand how they differentiate between each other
6. Cluster neighbourhoods based on new venues dataset and visualize clusters on the map
7. Analyze if social environment (venues) of the neighbourhoods impact on rent prices, and if yes, then which venue categories specifically have most effect

3.1 Identify list of Vancouver neighbourhoods and obtain average rent prices for these neighbourhoods

Gathered data from a csv file after being pre-processed and formatted to suit our needs, as described in 2.2, is now stored in a pandas dataframe. Following is the snapshot of the pre-processed dataframe wih only needed and formatted data fields.

	Neighbourhood	Average_Rent
0	University Endowment Lands	2097
1	West Vancouver	2024
2	North Vancouver DM	1881
3	Westside/Kerrisdale	1773
4	Downtown	1738
5	English Bay	1701
6	Kitsilano/Point Grey	1662
7	West End/Stanley Park	1615
8	South Granville/Oak	1545
9	Southeast Vancouver	1539
10	Vancouver	1480

3.2 Get neighbourhood coordinates and map neighbourhoods to have visual understanding of the region

Now, as we have gathered the list of neighbourhoods and their average rent prices, we can proceed by obtaining geographical coordinates for the these neighbourhoods. We will use Geocoder library to obtain this information. By creating a function to gather coordinates of the

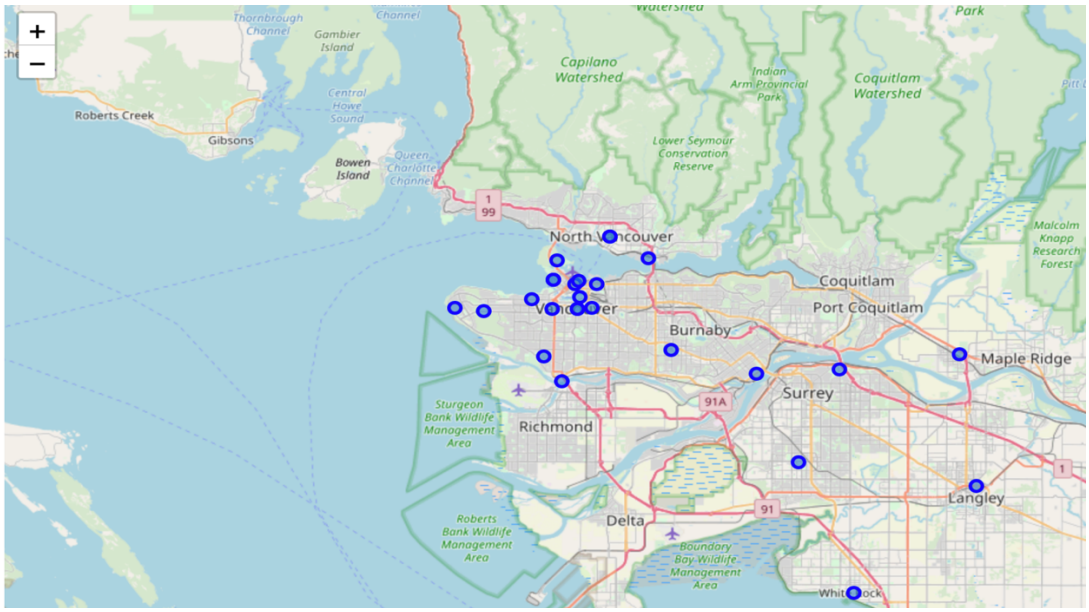
neighbourhoods and then running it on each of the neighbourhoods we obtained a list of the coordinates, as shown below.

```
[[49.25857000000008, -123.22014999999999],
 [49.26114904019019, -123.25304149399976],
 [49.32108000000005, -123.07520999999997],
 [49.22097000000008, -123.15072999999995],
 [49.28108000000003, -123.11589999999995],
 [49.132837756669176, -122.85958479339024],
 [49.26833000000005, -123.16541999999998],
 [49.301060000000064, -123.13572999999997],
 [49.260830000000055, -123.14099999999996],
 [49.270110000000045, -123.11009999999999],
 [49.260380000000055, -123.11335999999994],
 [49.32108000000005, -123.07520999999997],
 [49.30307749685801, -123.03149972143655],
 [49.11325200000002, -122.6573235],
```

Then we include obtained coordinates into our previous dataframe containing list of the neighbourhoods and associated average rent prices. Below is the resulting dataframe snapshot.

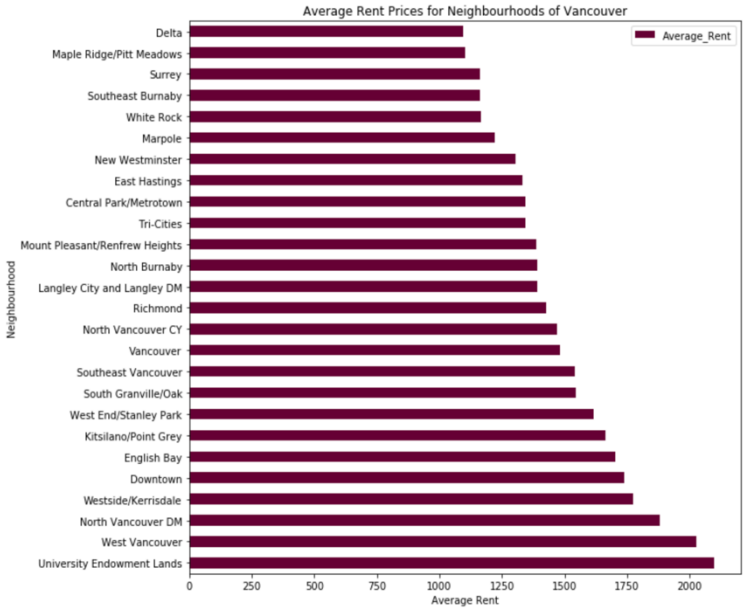
	Neighbourhood	Average_Rent	Latitude	Longitude
0	University Endowment Lands	2097	49.258570	-123.220150
1	West Vancouver	2024	49.261149	-123.253041
2	North Vancouver DM	1881	49.321080	-123.075210
3	Westside/Kerrisdale	1773	49.220970	-123.150730
4	Downtown	1738	49.281080	-123.115900
5	English Bay	1701	49.132838	-122.859585
6	Kitsilano/Point Grey	1662	49.268330	-123.165420
7	West End/Stanley Park	1615	49.301060	-123.135730
8	South Granville/Oak	1545	49.260830	-123.141000
9	Southeast Vancouver	1539	49.270110	-123.110100
10	Vancouver	1480	49.260380	-123.113360
11	North Vancouver CY	1471	49.321080	-123.075210
12	Richmond	1427	49.303077	-123.031500
13	Langley City and Langley DM	1392	49.113252	-122.657324
14	North Burnaby	1392	49.284611	-123.139880
15	Mount Pleasant/Renfrew Heights	1385	49.261780	-123.096080
16	Tri-Cities	1344	49.260380	-123.113360
17	Central Park/Metrotown	1343	49.226660	-123.005270
18	East Hastings	1332	49.281163	-123.090760
19	New Westminster	1305	49.207060	-122.908410
20	Marpole	1221	49.200670	-123.130530
21	White Rock	1164	49.023920	-122.796790
22	Southeast Burnaby	1161	49.284611	-123.139880
23	Surrey	1161	49.210498	-122.813874
24	Maple Ridge/Pitt Meadows	1102	49.222531	-122.676635
25	Delta	1095	49.284150	-123.112080

As we now have gathered the coordinates for the neighbourhoods, let's show them on the map of Vancouver to visualize their locations. Using Folium library we were able to create following map of Vancouver.



3.3 Cluster neighbourhoods based on average rent prices and conduct preliminary analysis

Before proceeding with clustering of neighbourhoods based on solely average rent prices, let's first look at how neighbourhoods differ from each other in terms of the rent prices. Let's create a simple bar chart to visualize the average rent prices in Vancouver neighbourhoods.

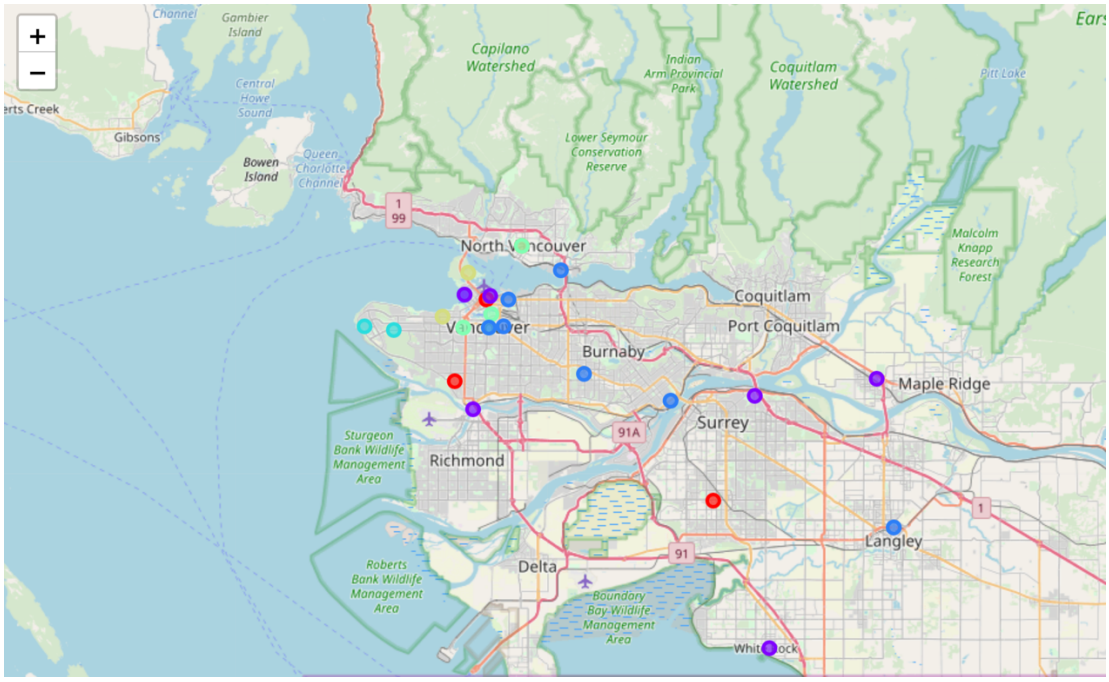


Next, let's now cluster our neighbourhoods based on average rent prices and see what results we will get. For clustering I will use K-means clustering algorithm, which is very helpful approach for performing unsupervised clustering based on similarities of the input features. In our case the feature is the average rent price of each of the neighbourhoods. We will run K-means algorithm for 7 clusters, which has been identified as most optimal number of clusters for our given dataset using elbow parameter.

Then we will include obtained set of cluster labels into our main dataframe. Following snapshot depicts the header of the merged dataframe.

	Neighbourhood	Average_Rent	Cluster Labels	Latitude	Longitude
0	University Endowment Lands	2097	3	49.258570	-123.220150
1	West Vancouver	2024	3	49.261149	-123.253041
2	North Vancouver DM	1881	6	49.321080	-123.075210
3	Westside/Kerrisdale	1773	0	49.220970	-123.150730
4	Downtown	1738	0	49.281080	-123.115900

Now, let's show our clustered neighbourhoods on the Vancouver map and then analyze if we can derive any visual insights from these clustering. We again use Folium library to show clustered neighbourhoods on the map of Vancouver.



The map above shows some insight on why some clusters of neighbourhoods have higher average rent prices than others. For example, on the map we can see that neighbourhood clusters 3 and 0, representing highest average rent prices, are located on the west end of Vancouver City as well as within high-end boroughs such as Downtown Vancouver, English Bay and South Vancouver.

3.4 Gather information on venues for each of the neighbourhoods

Let's now proceed with gathering information on venues in each of these neighbourhoods and then perform some analysis and clustering to see what we can derive. We will use Foursquare API to obtain the venues information for our neighbourhoods.

First, we have created a function to obtain max 100 venues within a 2000 meters radius of each neighbourhood center. Then we ran this function on each of the neighbourhoods and store gathered data into a dataframe. Below snapshot depicts the first 5 rows of the resulting table.

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	University Endowment Lands	49.258570	-123.220150	Ranger Station	49.258738	-123.222438	Trail
1	University Endowment Lands	49.258570	-123.220150	Cleveland Trail, Pacific Spirit Park	49.255623	-123.218226	Trail
2	West Vancouver	49.261149	-123.253041	Beaty Biodiversity Museum	49.263347	-123.251000	Science Museum
3	West Vancouver	49.261149	-123.253041	Mercante	49.263765	-123.255148	Italian Restaurant
4	West Vancouver	49.261149	-123.253041	The Point Grill	49.261740	-123.255626	Sandwich Place

3.5 Analyze venues in the neighbourhoods by identifying the most common venue categories in each of the neighbourhoods to understand how they differentiate between each other.

Now let’s proceed with analyzing collected information on venues. For example, let's check how many venues were returned for each neighbourhood.

	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighbourhood						
Central Park/Metrotown	45	45	45	45	45	45
Delta	100	100	100	100	100	100
Downtown	100	100	100	100	100	100
East Hastings	26	26	26	26	26	26
English Bay	1	1	1	1	1	1
Kitsilano/Point Grey	34	34	34	34	34	34
Langley City and Langley DM	12	12	12	12	12	12
Maple Ridge/Pitt Meadows	18	18	18	18	18	18
Marpole	16	16	16	16	16	16
Mount Pleasant/Renfrew Heights	72	72	72	72	72	72

Next, we will analyze each neighbourhood for their venue categories and frequency of occurences. Then we will group rows by neighbourhood and by taking the mean of the frequency of occurrence of each category. At the end we will create a new dataframe to display top 10 venue categories for each of our neighbourhoods. Following snapshot shows first 3 rows of the dataframe.

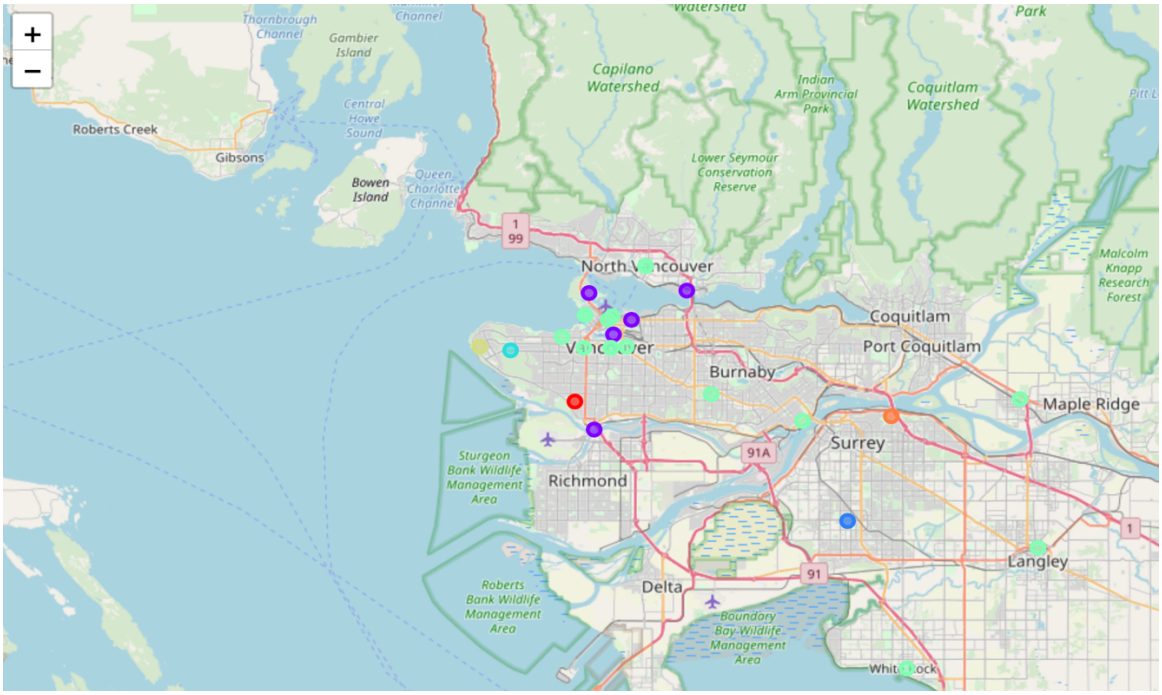
	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Central Park/Metrotown	Bakery	Clothing Store	Coffee Shop	American Restaurant	Hotel	Bookstore	Cosmetics Shop	Gym / Fitness Center	Electronics Store	Movie Theater
1	Delta	Coffee Shop	Café	Hotel	Italian Restaurant	Restaurant	Sandwich Place	Middle Eastern Restaurant	Lounge	Pub	Japanese Restaurant
2	Downtown	Hotel	Coffee Shop	Café	Restaurant	Steakhouse	Concert Hall	Electronics Store	Italian Restaurant	Ramen Restaurant	Bookstore

3.6 Cluster neighbourhoods based on new venues dataset and visualize clusters on the map.

Now, let's cluster the neighbourhoods. We will again run K-means algorithm to cluster the neighbourhoods into 7 clusters. We will then include produced cluster labels into our existing above dataframe, which is now displayed as follows.

	Neighbourhood	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	University Endowment Lands	49.258570	-123.220150	3	Trail	Zoo Exhibit	Ethiopian Restaurant	French Restaurant	Food Truck	Food Court	Food	Financial or Legal Service	Fast Food Restaurant	Falafel Restaurant
1	West Vancouver	49.261149	-123.253041	5	Coffee Shop	Park	Italian Restaurant	Plaza	Fast Food Restaurant	Sandwich Place	Museum	Science Museum	Ethiopian Restaurant	Food
2	North Vancouver DM	49.321080	-123.075210	4	Coffee Shop	Sushi Restaurant	Grocery Store	Middle Eastern Restaurant	Italian Restaurant	Gym / Fitness Center	Sandwich Place	Mediterranean Restaurant	Café	Portuguese Restaurant
3	Westside/Kerrisdale	49.220970	-123.150730	0	Supermarket	Spanish Restaurant	Café	Park	Event Space	Food Truck	Food Court	Food	Financial or Legal Service	Fast Food Restaurant
4	Downtown	49.281080	-123.115900	4	Hotel	Coffee Shop	Café	Restaurant	Steakhouse	Concert Hall	Electronics Store	Italian Restaurant	Ramen Restaurant	Bookstore

Finally, let's visualize our new neighbourhood clusters based on venue categories using Folium library.



3.7 Analyze if social environment (venues) of the neighbourhoods impact on rent prices, and if yes, then which venue categories specifically have most effect

First, let's examine each produced cluster and determine the discriminating venue categories that distinguish each cluster.

Cluster 1. We can see that cluster 1 consists of only one neighbourhood, for which the most common venue type is a supermarket. We can also notice that other next most common venue types are diverse and do not specifically tell anything about the cluster.

	Neighbourhood	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
3	Westside/Kerrisdale	49.22097	-123.15073	0	Supermarket	Spanish Restaurant	Café	Park	Event Space	Food Truck	Food Court	Food	Financial or Legal Service	Fast Food Restaurant

Cluster 2. For Cluster 2 there are also no differentiating venue categories, however, looking at the map above and locating the Cluster 2 neighbourhoods, we can tell that most of them are located in touristic areas with variety of venues and provide a very vibrant lifestyle.

	Neighbourhood	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
7	West End/Stanley Park	49.301060	-123.13573	1	Aquarium	Park	Theme Park Ride / Attraction	Zoo Exhibit	Event Space	Pub	Playground	Outdoor Sculpture	Other Great Outdoors	Music Venue
9	Southeast Vancouver	49.270110	-123.11010	1	Coffee Shop	Pet Store	Art Gallery	Sushi Restaurant	Boat or Ferry	Gastropub	Brewery	Italian Restaurant	Café	Liquor Store
12	Richmond	49.303077	-123.03150	1	Bus Station	Fast Food Restaurant	Coffee Shop	Sporting Goods Shop	Gas Station	Gastropub	Music Store	Pastry Shop	French Restaurant	Discount Store
18	East Hastings	49.281163	-123.09076	1	Coffee Shop	Diner	Sandwich Place	Brewery	Park	Restaurant	Deli / Bodega	Pub	Food Truck	Spa
20	Marpole	49.200670	-123.13053	1	Restaurant	Bus Stop	Hotel Pool	Grocery Store	Rental Car Location	Scenic Lookout	Night Market	Shoe Store	Café	Burger Joint

Cluster 3. This is another single-neighbourhood cluster which has food and restaurant venue categories as primary differentiating features.

	Neighbourhood	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
5	English Bay	49.132838	-122.859585	2	Sandwich Place	Zoo Exhibit	Ethiopian Restaurant	Food Truck	Food Court	Food	Financial or Legal Service	Fast Food Restaurant	Falafel Restaurant	Exhibit

Cluster 4. A very small cluster which can be differentiated by a very specific venue category - Trail. We can derive that this neighbourhood has a lot of options for those who love spending time in nature.

	Neighbourhood	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	University Endowment Lands	49.25857	-123.22015	3	Trail	Zoo Exhibit	Ethiopian Restaurant	French Restaurant	Food Truck	Food Court	Food	Financial or Legal Service	Fast Food Restaurant	Falafel Restaurant

Cluster 5. A large cluster with a strong differentiating factor - most common venue categories are coffee shops, bakeries and cafes. Looking at the above map, we also can derive that neighbourhoods of this cluster are located in the heart of Vancouver and provide lots of attractions for its habitants.

	Neighbourhood	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
2	North Vancouver DM	49.321080	-123.075210	4	Coffee Shop	Sushi Restaurant	Grocery Store	Middle Eastern Restaurant	Italian Restaurant	Gym / Fitness Center	Sandwich Place	Mediterranean Restaurant	Café	Portuguese Restaurant
4	Downtown	49.281080	-123.115900	4	Hotel	Coffee Shop	Café	Restaurant	Steakhouse	Concert Hall	Electronics Store	Italian Restaurant	Ramen Restaurant	Bookstore
6	Kitsilano/Point Grey	49.268330	-123.165420	4	Coffee Shop	Park	Pizza Place	Café	Sushi Restaurant	Tanning Salon	Diner	Pub	Chinese Restaurant	Restaurant
8	South Granville/Oak	49.260830	-123.141000	4	Coffee Shop	Furniture / Home Store	Japanese Restaurant	Breakfast Spot	American Restaurant	Thai Restaurant	Sushi Restaurant	Electronics Store	Pharmacy	Restaurant
10	Vancouver	49.260380	-123.113360	4	Coffee Shop	Chinese Restaurant	Japanese Restaurant	Pizza Place	Fast Food Restaurant	Ramen Restaurant	Sushi Restaurant	Dessert Shop	Indian Restaurant	Liquor Store
11	North Vancouver CY	49.321080	-123.075210	4	Coffee Shop	Sushi Restaurant	Grocery Store	Middle Eastern Restaurant	Italian Restaurant	Gym / Fitness Center	Sandwich Place	Mediterranean Restaurant	Café	Portuguese Restaurant
13	Langley City and Langley DM	49.113252	-122.657324	4	Breakfast Spot	Sporting Goods Shop	Portuguese Restaurant	Office	Department Store	Furniture / Home Store	Rental Car Location	Pet Store	Martial Arts Dojo	Gym / Fitness Center
14	North Burnaby	49.284611	-123.139880	4	Coffee Shop	Sculpture Garden	Café	Ice Cream Shop	Italian Restaurant	Burger Joint	Spanish Restaurant	Dessert Shop	Sandwich Place	Fast Food Restaurant
15	Mount Pleasant/Renfrew Heights	49.261780	-123.096080	4	Coffee Shop	Diner	Pizza Place	Sushi Restaurant	Sandwich Place	Breakfast Spot	Bar	Thrift / Vintage Store	Lounge	Indian Restaurant

Cluster 6. Another small cluster consisting of one neighbourhood with diverse common venues categories that can be found within it.

	Neighbourhood	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
1	West Vancouver	49.261149	-123.253041	5	Coffee Shop	Park	Italian Restaurant	Plaza	Fast Food Restaurant	Sandwich Place	Museum	Science Museum	Ethiopian Restaurant	Food

Cluster 7. Last cluster with also one neighbourhood in it. Interestingly, most common venue, which is a music venue, is a very specific and rare venue category in comparison to other clusters.

	Neighbourhood	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
23	Surrey	49.210498	-122.813874	6	Music Venue	Construction & Landscaping	Trail	Zoo Exhibit	Event Space	Food Truck	Food Court	Food	Financial or Legal Service	Fast Food Restaurant

Let's now merge above dataframe and information on average rent prices of the neighbourhoods to see if we can derive any relations between most frequent venue categories and rent prices. Following snapshot displays first half of the final dataframe.

	Neighbourhood	Average_Rent	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	University Endowment Lands	2097	49.258570	-123.220150	3	Trail	Zoo Exhibit	Ethiopian Restaurant	French Restaurant	Food Truck	Food Court	Food	Financial or Legal Service	Fast Food Restaurant	Falafel Restaurant
1	West Vancouver	2024	49.261149	-123.253041	5	Coffee Shop	Park	Italian Restaurant	Plaza	Fast Food Restaurant	Sandwich Place	Museum	Science Museum	Ethiopian Restaurant	Food
2	North Vancouver DM	1881	49.321080	-123.075210	4	Coffee Shop	Sushi Restaurant	Grocery Store	Middle Eastern Restaurant	Italian Restaurant	Gym / Fitness Center	Sandwich Place	Mediterranean Restaurant	Café	Portuguese Restaurant
3	Westside/Kerrisdale	1773	49.220970	-123.150730	0	Supermarket	Spanish Restaurant	Café	Park	Event Space	Food Truck	Food Court	Food	Financial or Legal Service	Fast Food Restaurant
4	Downtown	1738	49.281080	-123.115900	4	Hotel	Coffee Shop	Café	Restaurant	Steakhouse	Concert Hall	Electronics Store	Italian Restaurant	Ramen Restaurant	Bookstore
5	English Bay	1701	49.132838	-122.859585	2	Sandwich Place	Zoo Exhibit	Ethiopian Restaurant	Food Truck	Food Court	Food	Financial or Legal Service	Fast Food Restaurant	Falafel Restaurant	Exhibit
6	Kitsilano/Point Grey	1662	49.268330	-123.165420	4	Coffee Shop	Park	Pizza Place	Café	Sushi Restaurant	Tanning Salon	Diner	Pub	Chinese Restaurant	Restaurant
7	West End/Stanley Park	1615	49.301060	-123.135730	1	Aquarium	Park	Theme Park Ride / Attraction	Zoo Exhibit	Event Space	Pub	Playground	Outdoor Sculpture	Other Great Outdoors	Music Venue
8	South Granville/Oak	1545	49.260830	-123.141000	4	Coffee Shop	Furniture / Home Store	Japanese Restaurant	Breakfast Spot	American Restaurant	Thai Restaurant	Sushi Restaurant	Electronics Store	Pharmacy	Restaurant
9	Southeast Vancouver	1539	49.270110	-123.110100	1	Coffee Shop	Pet Store	Art Gallery	Sushi Restaurant	Boat or Ferry	Gastropub	Brewery	Italian Restaurant	Café	Liquor Store
10	Vancouver	1480	49.260380	-123.113360	4	Coffee Shop	Chinese Restaurant	Japanese Restaurant	Pizza Place	Fast Food Restaurant	Ramen Restaurant	Sushi Restaurant	Dessert Shop	Indian Restaurant	Liquor Store

4. Results

Looking at the above table, it's hard to derive any direct relationships between the frequency of specific venue category in a neighbourhood and the average rent price in that neighbourhood. However, there are some hints showing that there can be relationships between these factors, for example:

- neighbourhoods with parks, trails and other natural venues tend to have higher average rent prices.
- neighbourhoods which have cultural venues such as museums, concert halls, plazas and event spaces tend to rank among high average rent priced neighbourhoods

Nevertheless, we can state that more frequent presence of any specific type of venues does not play a key role in defining rent prices in any of the neighbourhoods. Actually, looking at the maps we can state that geographical location plays a more if not most critical role in defining the rental prices for accommodation. We should also consider presence of key infrastructure, such as public transportation, schools and universities, which provide a comfortable living experience in a given neighbourhood.

Another factors that highly affect average rent prices can include social and economical aspects of the neighbourhoods, such as crime rates, household income and others.

5. Discussion

As mentioned before, Vancouver is a large metropolis with a high population density and restricted geographical area. Above this, there are many other factors that might impact on the rental prices in the City’s neighbourhoods. As there is such a complexity, very different approaches can be used in clustering and classification studies. Moreover, it is obvious that not every classification method can yield the same high quality results.

K-means algorithm used as part of this clustering study has given relatively insightful information on how neighbourhoods are clustered in this City. However, the results of the algorithm did not show direct relationship of how presence of venues in the specific neighbourhood affects the average rent prices on accommodation in that neighbourhood.

Although results of the clustering algorithm in a tabular format did not explicitly state any associations, visualization of the clusters on a map did provide some understanding on how geographical location and proximity to specific natural landmarks and center of the City play bigger role on defining the average rent prices.

As in this work we have only considered a single aspect, frequency of venue categories, other above-mentioned factors might be used in future research to obtain more encompassing analysis and lead to results of higher accuracy.

6. Conclusion

In conclusion, conducted analytical work can be summarized in followings:

- there is no a defined and direct relationship between frequent presence of any of the venue categories in a neighbourhood and average rent prices in that neighbourhood
- neighbourhoods with parks, trails and other natural venues tend to have higher average rent prices
- neighbourhoods which have cultural venues such as museums, concert halls, plazas and event spaces tend to rank among high average rent priced neighbourhoods
- further analysis of the matter by taking other factors, such as geographical location, infrastructure, social aspects, population density and household income of the residents might deliver higher-quality and more encompassing results.

7. References

[1] City of Vancouver. Information about Vancouver
https://vancouver.ca/about-vancouver.aspx

[2] Vancouver population
http://worldpopulationreview.com/world-cities/vancouver-population/

[3] 15th Annual Demographia International Housing Affordability Survey
http://demographia.com/media_rls_2019.pdf

[4] CMHC Rental Market Report Data Tables
https://www.cmhc-schl.gc.ca/en/data-and-research/data-tables

[5] Foursquare API
https://developer.foursquare.com/

[6] Geocoder Python Library
https://geocoder.readthedocs.io/index.html