The Battle of Neighborhoods

Content

1. Introduction Section :

Discussion of the business problem and the audience who would be interested in this project.

1.1 Scenario and Background

The purpose is to search for a house for rent in the Canadian city of Vancouver. This house must meet certain requirements regarding the location of the same, environment, places of interest, proximity to means of transport. It is not a matter of making a comparison with another city in the world, but rather seeks to meet the minimum requirements of a target client. Although this problem could be solved simply by making use of the professional services of a real estate agent, all the techniques and knowledge acquired in the course of IBM will be applied.

1.2 Problem to be resolved:

As detailed in the previous section, you must search for a home in the city of Vancouver. The home must meet the following requirements:

- Must be in a neighborhood of the city with a family atmosphere.
- In the environment there must be parks to walk and play sports.
- It must be close to public transportation.
- There must be a sports offer in the vicinity.
- The area should be easily accessible by road.
- It must be located next to ski areas.
- The catering offer must be rich, coffee shops and restaurants must be in the immediate vicinity.

1.3 Interested Audience:

The audience interested in this project is a supposed target customer who must change his place of residence from a European capital, to the city of Vancouver. This is a manager of a multinational company with a high purchasing power and that has no economic limitations in terms of the rental price of the home in question since the cost of such rent will be borne by the company for which he works. The requirements on the location of the home seek to find an environment similar to the current place of residence of the client.

2. Data Section:

Description of the data and its sources that will be used to solve the problem

2.1 Data Required to resolve the problem

To make the choice of the home that the client is looking for in Vancouver, the following information is needed:

- Information about the neighborhoods of the city with their geographic location data.
- Information on the means of transport of the city with your location data.
- Data on the houses that are under rent in the city.
- Geographical data on the requirements of green leisure areas.
- Interesting information about the restoration offer of the city with its geographical coordinates of location.

2.2 Data sources and manipulation.

Data on the neighborhoods and zip codes of the city of Vancouver have been obtained here:

https://www.geonames.org/postalcode-search.html?q=vancouver&country=CA&adminCode1=BC

This data has been easily and quickly manageable directly by copying from the page and manipulating in an Excel file. Since it is a very small size table, it has been decided to carry out in this way so as not to steal time from the other parts of the project.

The data on points of interest such as transportation, leisure centers, parks, restaurants and others, have been obtained directly from Foursquare.

The geolocation of the neighborhoods carried out directly on the project, allowed to observe at a glance, which neighborhoods of the city were more interesting for the target client.

Since in this case our client does not present economic problems for the rental of the house, because it is financed by the company for which he works, the data on rental housing in the city of Vancouver have been ignored at first and all efforts have been focused on deciding, within the district chosen by the geoposión, which would be the best neighborhood that fulfilled the objectives sought.

This is the table we have obtained from the page
https://www.geonames.org/postalcode-
search.html?q=vancouver&country=CA&adminCode1=BC
and that has been handled on an excel sheet to create a .csv extension file

	Code	Borough	Neighbourhood
0	V5S	Killarney	Killarney
1	V5K	North Hastings	North Hastings
2	V5L	North Grandview	Woodlands
3	V5P	SE Kensington	Victoria
4	V5R	South Renfrew	Collingwood

We are going to assign coordinates to the data of the neighborhoods.
We are going to store the coordinates in a df.
Now we analyze the integrity of the table that we just created.

	Code	Borough	Neighbourhood	Latitude	Longitude
0	V5S	Killarney	Killarney	49.215345	-123.041225
1	V5K	North Hastings	North Hastings	49.281665	-123.039980
2	V5L	North Grandview	Woodlands	49.280700	-123.066842
3	V5P	SE Kensington	Victoria	49.223370	-123.067100
4	V5R	South Renfrew	Collingwood	49.239335	-123.041105
5	V5T	East Mount Pleasant	East Mount Pleasant	49.263410	-123.091214
6	V5Z	East Fairview	South Cambie	49.247102	-123.120980
7	V6E	South West	South West End	49.283870	-123.128981
8	V6K	Central Kitsilano	Central Kitsilano	49.267105	-123.165282
9	V6L	North West	Arbutus Ridge	49.249915	-123.165854

2.3 How the data will be manipulated.

The data on neighborhoods we have already advanced in what way have been handled in the previous section. These have been copied directly from the page provided and have been quickly manipulated on an Excel sheet. Since it is totally clean data, it has only been necessary to order the columans and save the file with .csv extension to facilitate later reading on our notebook.

The data will be used as follows: Use Foursquare and geopy data to assign the top 10 locations for all neighborhoods in the North Vancouver district as this is the designated place to locate the home and grouped into groups.

The foursquare and geopy data have been used to plot the location of the places of interest.

Using Geopy-distance and Nominatim, neighborhoods and places of interest have been geolocated.

2.4 Data Mapping.

Several maps were created for the analysis of the project:

- Map of the city of Vancouver with the districts.
- Map of the district of North Vancouver with its neighborhoods.
- Map with the distribution of points of interest.

3. Methodology :

3.1 Steps taken and strategy followed.

By direct appreciation of the neighborhoods of the city of Vancouver on the map, it is decided to perform the analysis on the district of North Vancouver as an objective.

	Code	Borough	Neighbourhood	Latitude	Longitude
0	V7G	North Vancouver	Outer East	49.388046	-122.934285
1	V7H	North Vancouver	Inner East	49.316885	-122.990073
2	V7J	North Vancouver	East Central	49.332998	-123.018990
3	V7K	North Vancouver	North Central	49.346020	-123.039847
4	V7L	North Vancouver	South Central	49.318325	-123.056241
5	V7M	North Vancouver	Southwest Central	49.320777	-123.082620
6	V7N	North Vancouver	Northwest Central	49.343520	-123.073501
7	V7P	North Vancouver	Southwest	49.320565	-123.115307
8	V7R	North Vancouver	Northwest	49.369327	-123.100700
9	V7G	North Vancouver	Deep Cove	49.388046	-122.934285
10	V7M	North Vancouver	Lower Lonsdale	49.320777	-123.082620
11	V7J	North Vancouver	Lynn Valley	49.332998	-123.018990

3.2 Data science methods, machine learning, mapping tools and exploratory data analysis.

Once the objective of the study has been determined, all geolocation, marking and grouping techniques are put into practice, which will be necessary to analyze the best options and make a decision based on the data obtained. The procedure followed will be the same as that already concluded in the case study of the city of New York. Points of interest will be searched in the designated neighborhood and a grouping will be made by points of interest that will help us decide which steps to follow.

Geolocate the district of North Vancouver with Nominatim.

It is time to register in the Foursquare service and for this we provide our access credentials.
We set one of the neighborhoods in the neighborhood and geolocalize this.
Now, let's get the top 100 venues that are in Horseshoe Bay within a radius of 500 meters.
First, let's create the GET request URL. Name your URL url.
We created a new data frame with the points of interest obtained in the GET of the API

	name	categories	lat	lng
0	The Edge Climbing Centre	Gym	49.317268	-123.114546
1	Mooyah	Burger Joint	49.323332	-123.113961
2	La Taqueria Pinche Taco Shop	Taco Place	49.317522	-123.111970
3	The Tomahawk	Diner	49.322817	-123.113150
4	Sushi Man Japanese Restaurant	Japanese Restaurant	49.323781	-123.111095

We check the number of venues returned by Foursquare.

Explore Neighborhoods in North Vancouver
We list the venues according to the neighborhoods of North Vancouver

Outer East
Inner East
East Central
North Central
South Central
Southwest Central
Northwest Central
Southwest
Northwest
Deep Cove
Lower Lonsdale
Lynn Valley
Dundarave Village
Edgemont Village
Horseshoe Bay
Ambleside

Let's check the size of the resulting dataframe

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Inner East	49.316885	-122.990073	Green Mountain Landscaping	49.316347	-122.986294	Construction & Landscaping
1	Inner East	49.316885	-122.990073	Mccartney Creek Park	49.320577	-122.993579	Baseball Field
2	North Central	49.346020	-123.039847	Endless Summer Landscapes	49.348430	-123.038549	Construction & Landscaping
3	North Central	49.346020	-123.039847	Mountain Market	49.341978	-123.037982	Convenience Store
4	South Central	49.318325	-123.056241	Filthy Cleaning	49.320605	-123.054664	Paper / Office Supplies Store

Let's check how many venues were returned for each neighborhood

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Ambleside	4	4	4	4	4	4
Dundarave Village	16	16	16	16	16	16
Horseshoe Bay	16	16	16	16	16	16
Inner East	2	2	2	2	2	2
Lower Lonsdale	4	4	4	4	4	4
North Central	2	2	2	2	2	2
Northwest	7	7	7	7	7	7
Northwest Central	1	1	1	1	1	1
South Central	2	2	2	2	2	2
Southwest	16	16	16	16	16	16
Southwest Central	4	4	4	4	4	4

Let's find out how many unique categories can be curated from all the returned venues

Analyze Each Neighborhood

	Neighborhood	Bank	Baseball Field	Burger Joint	Bus Stop	Coffee Shop	Construction & Landscaping	Convenience Store	Department Store	Diner	Discount Store	Greek Restaurant	Grocery Store	Gym	Japanese Restaurant	Liq St
0	Inner East	0	0	0	0	0	1	0	0	0	0	0	0	0	0	
1	Inner East	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
2	North Central	0	0	0	0	0	1	0	0	0	0	0	0	0	0	
3	North Central	0	0	0	0	0	0	1	0	0	0	0	0	0	0	
4	South Central	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

And let's examine the new dataframe size.

Next, let's group rows by neighborhood and by taking the mean of the frequency of occurrence of each category

	Neighborhood	Bank	Baseball Field	Burger Joint	Bus Stop	Coffee Shop	Construction & Landscaping	Convenience Store	Department Store	Diner	Discount Store	Greek Restaurant	Grocery Store	Gym	Japa Resta
0	Ambleside	0.25	0.0	0.0000	0.000000	0.000000	0.0	0.0	0.25	0.0000	0.0000	0.0000	0.0000	0.0000	0
1	Dundarave Village	0.00	0.0	0.0625	0.000000	0.062500	0.0	0.0	0.00	0.0625	0.0625	0.0625	0.0625	0.0625	0
2	Horseshoe Bay	0.00	0.0	0.0625	0.000000	0.062500	0.0	0.0	0.00	0.0625	0.0625	0.0625	0.0625	0.0625	0
3	Inner East	0.00	0.5	0.0000	0.000000	0.000000	0.5	0.0	0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0
4	Lower Lonsdale	0.00	0.0	0.0000	0.000000	0.000000	0.0	0.0	0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0
5	North Central	0.00	0.0	0.0000	0.000000	0.000000	0.5	0.5	0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0
6	Northwest	0.00	0.0	0.0000	0.142857	0.142857	0.0	0.0	0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0
7	Northwest Central	0.00	0.0	0.0000	0.000000	0.000000	0.0	0.0	0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0
8	South Central	0.00	0.0	0.0000	0.000000	0.000000	0.0	0.0	0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0
9	Southwest	0.00	0.0	0.0625	0.000000	0.062500	0.0	0.0	0.00	0.0625	0.0625	0.0625	0.0625	0.0625	0
10	Southwest Central	0.00	0.0	0.0000	0.000000	0.000000	0.0	0.0	0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0

Let's now look for the top 5 of venues.

----Ambleside---- venue freq 0 Bank 0.25 1 Mexican Restaurant 0.25 2 Department Store 0.25 3 Pharmacy 0.25 4 Toy / Game Store 0.00 5 Taco Place 0.00 6 Sporting Goods Shop 0.00 7 Ski Chairlift 0.00 8 Shopping Mall 0.00 9 Playground 0.00

----Dundarave Village---- venue freq 0 Grocery Store 0.06 1 Discount Store 0.06 2 Other Great Outdoors 0.06 3 Mobile Phone Shop 0.06 4 Shopping Mall 0.06 5 Liquor Store 0.06 6 Japanese Restaurant 0.06 7 Gym 0.06 8 Greek Restaurant 0.06 9 Outdoor Supply Store 0.06

----Horseshoe Bay---- venue freq 0 Grocery Store 0.06 1 Discount Store 0.06 2 Other Great Outdoors 0.06 3 Mobile Phone Shop 0.06 4 Shopping Mall 0.06 5 Liquor Store 0.06 6 Japanese Restaurant 0.06 7 Gym 0.06 8 Greek Restaurant 0.06 9 Outdoor Supply Store 0.06

----Inner East---- venue freq 0 Baseball Field 0.5 1 Construction & Landscaping 0.5 2 Pharmacy 0.0 3 Mobile Phone Shop 0.0 4 Mountain 0.0 5 Other Great Outdoors 0.0 6 Outdoor Supply Store 0.0 7 Paper / Office Supplies Store 0.0 8 Park 0.0 9 Bank 0.0

----Lower Lonsdale---- venue freq 0 Park 0.50 1 Playground 0.25 2 Market 0.25 3 Bank 0.00 4 Mexican Restaurant 0.00 5 Toy / Game Store 0.00 6 Taco Place 0.00 7 Sporting Goods Shop 0.00 8 Ski Chairlift 0.00 9 Shopping Mall 0.00

----North Central---- venue freq 0 Construction & Landscaping 0.5 1 Convenience Store 0.5 2 Bank 0.0 3 Pharmacy 0.0 4 Mountain 0.0 5 Other Great Outdoors 0.0 6 Outdoor Supply Store 0.0 7 Paper / Office Supplies Store 0.0 8 Park 0.0 9 Shopping Mall 0.0

----Northwest---- venue freq 0 Trail 0.43 1 Mountain 0.14 2 Bus Stop 0.14 3 Coffee Shop 0.14 4 Ski Chairlift 0.14 5 Pharmacy 0.00 6 Other Great Outdoors 0.00 7 Outdoor Supply Store 0.00 8 Paper / Office Supplies Store 0.00 9 Park 0.00

----Northwest Central---- venue freq 0 Park 1.0 1 Bank 0.0 2 Baseball Field 0.0 3 Toy / Game Store 0.0 4 Taco Place 0.0 5 Sporting Goods Shop 0.0 6 Ski Chairlift 0.0 7 Shopping Mall 0.0 8 Playground 0.0 9 Pharmacy 0.0

----South Central---- venue freq 0 Martial Arts Dojo 0.5 1 Paper / Office Supplies Store 0.5 2 Baseball Field 0.0 3 Toy / Game Store 0.0 4 Taco Place 0.0 5 Sporting Goods Shop 0.0 6 Ski Chairlift 0.0 7 Shopping Mall 0.0 8 Playground 0.0 9 Pharmacy 0.0

----Southwest---- venue freq 0 Grocery Store 0.06 1 Discount Store 0.06 2 Other Great Outdoors 0.06 3 Mobile Phone Shop 0.06 4 Shopping Mall 0.06 5 Liquor Store 0.06 6 Japanese Restaurant 0.06 7 Gym 0.06 8 Greek Restaurant 0.06 9 Outdoor Supply Store 0.06

----Southwest Central---- venue freq 0 Park 0.50 1 Playground 0.25 2 Market 0.25 3 Bank 0.00 4 Mexican Restaurant 0.00 5 Toy / Game Store 0.00 6 Taco Place 0.00 7 Sporting Goods Shop 0.00 8 Ski Chairlift 0.00 9 Shopping Mall 0.00

Let's put that into a pandas dataframe

**First, let's write a function to sort the venues in descending order.**

Now let's create the new dataframe and display the top 10 venues for each 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
0	Ambleside	Bank	Department Store	Mexican Restaurant	Pharmacy	Gym	Grocery Store	Greek Restaurant	Discount Store	Diner	Convenience Store
1	Dundarave Village	Other Great Outdoors	Outdoor Supply Store	Grocery Store	Gym	Japanese Restaurant	Liquor Store	Toy / Game Store	Diner	Mobile Phone Shop	Discount Store
2	Horseshoe Bay	Other Great Outdoors	Outdoor Supply Store	Grocery Store	Gym	Japanese Restaurant	Liquor Store	Toy / Game Store	Diner	Mobile Phone Shop	Discount Store
3	Inner East	Baseball Field	Construction & Landscaping	Trail	Toy / Game Store	Burger Joint	Bus Stop	Coffee Shop	Convenience Store	Department Store	Diner
4	Lower Lonsdale	Park	Playground	Market	Burger Joint	Bus Stop	Coffee Shop	Baseball Field	Construction & Landscaping	Liquor Store	Convenience Store

Cluster Neighborhoods

Run k-means to cluster the neighborhood into 5 clusters.

Let's create a new dataframe that includes the cluster as well as the top 10 venues for each neighborhood.
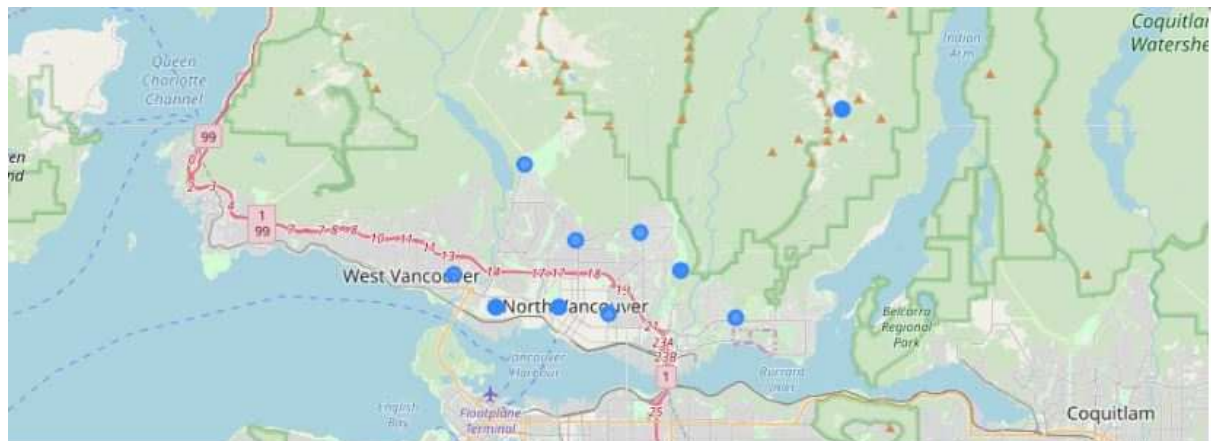
We examine the clusters below

Cluster 1

	Borough	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	North Vancouver	0.0	Construction & Landscaping	Convenience Store	Trail	Toy / Game Store	Baseball Field	Burger Joint	Bus Stop	Coffee Shop	Department Store	Diner

Cluster 2

	Borough	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	North Vancouver	1.0	Other Great Outdoors	Outdoor Supply Store	Grocery Store	Gym	Japanese Restaurant	Liquor Store	Toy / Game Store	Diner	Mobile Phone Shop	Discount Store
8	North Vancouver	1.0	Trail	Ski Chairlift	Bus Stop	Coffee Shop	Mountain	Diner	Gym	Grocery Store	Greek Restaurant	Discount Store
12	North Vancouver	1.0	Other Great Outdoors	Outdoor Supply Store	Grocery Store	Gym	Japanese Restaurant	Liquor Store	Toy / Game Store	Diner	Mobile Phone Shop	Discount Store
14	North Vancouver	1.0	Other Great Outdoors	Outdoor Supply Store	Grocery Store	Gym	Japanese Restaurant	Liquor Store	Toy / Game Store	Diner	Mobile Phone Shop	Discount Store
15	North Vancouver	1.0	Bank	Department Store	Mexican Restaurant	Pharmacy	Gym	Grocery Store	Greek Restaurant	Discount Store	Diner	Convenience Store

Cluster 3

	Borough	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	North Vancouver	2.0	Park	Playground	Market	Burger Joint	Bus Stop	Coffee Shop	Baseball Field	Construction & Landscaping	Liquor Store	Convenience Store
6	North Vancouver	2.0	Park	Trail	Liquor Store	Baseball Field	Burger Joint	Bus Stop	Coffee Shop	Construction & Landscaping	Convenience Store	Department Store
10	North Vancouver	2.0	Park	Playground	Market	Burger Joint	Bus Stop	Coffee Shop	Baseball Field	Construction & Landscaping	Liquor Store	Convenience Store

Cluster 4

	Borough	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	North Vancouver	3.0	Baseball Field	Construction & Landscaping	Trail	Toy / Game Store	Burger Joint	Bus Stop	Coffee Shop	Convenience Store	Department Store	Diner

Cluster 5

	Borough	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
4	North Vancouver	4.0	Paper / Office Supplies Store	Martial Arts Dojo	Trail	Gym	Grocery Store	Greek Restaurant	Discount Store	Diner	Department Store	Liquor Store

4. Results obtained:

Discussion of the results and how they help to make a decision.
The grouping according to interests is a tool that will allow to make a decision based on the interests of the project.

5. Discussion about observations:¶

Observations about the obtained data.
In view of the results obtained in the grouping according to the groups of venues, it is appreciated that the second one is the most interesting according to the preferences of the client, where he looks for a wide range of services ranging from leisure to transport.

6. Last conclusions:

Decide taken and conclusion report.
As already explained in the previous section, the client will be provided with the information and the report will show the best options that have been presented to us thanks to the geolocation of venues and the grouping of them according to the interests of the project.