Capstone Final Project

Professional Data Science Certificate by IBM/Coursera

Table of contents

Introduction: Description and Discussion of the Background:

1.1 Description:

Project Title:

Finding the best location and neighborhood in Vancouver to establish Daily Convenient/Franchisee store which will be operated 24 Hours a Day (24/7).

Project Goal:

This Project goal to make an analysis based on the neighborhoods , features and crime record of the borough in Vancouver for the Business Investor/Stakeholders/Small Business Owner / Franchisee Opener to invest money to open business to the best borough with neighborhood as a comparative analysis between neighborhoods. The features include according to ratings, crime rates of that particular area, road connectivity, weather conditions, good management for emergency, water resources both fresh and waste water and excrement conveyed in sewers and recreational facilities.

This will guide to the investor to get awareness of the area and neighborhood before establishing the business to a new city, state, country or place for their daily convenient store to run 24/7.

1.2) Discussion of the Background:

It will help Business Investor/Stakeholders/Small Business Owner / Franchisee Opener making smart and efficient decision on selecting great neighborhood out of numbers of other neighborhoods and borough in Vancouver , Canada where they can run the business safely 24/7 as per crime records and analysis.

1.3) Project Prepared For :

Business Investor/Stakeholders/Small Business Owner / Franchisee Opener.

Many Business Investor/Stakeholders/Small Business Owner / Franchisee Opener are investing and opening a retail business to various states of Canada they needed lots of research for good borough and neighborhoods for their

investment. This project is for those Business Investor/Stakeholders/Small Business Owner / Franchisee Opener who are looking for better neighborhoods and safest borough in Vancouver.

1.4) Data Scientist role:

Explore and analysis all neighborhood and borough in Vancouver, Canada and find out the best and safe location for open Retail business/Franchisee for establishment.

1.5) Data Science tools and Techniques are used for Problem Analysis:

As part of analysis to choose the safest borough from crime data record, for opening a daily convenient store(24/7) and short listing a neighborhood, where this type of stores are not amongst in the most common venues, and as close to the city as possible.

Following data science tools, techniques and libraries are going to be use to analyze data and focus on the safest borough and explore its neighborhoods, sorting top ten common venues in each neighborhood, where daily convenient stores (24/7) are not amongst in the most common venues.

1.6) The Location:

Vancouver is a popular destination for Business Investor/Stakeholders/Small Business Owner / Franchisee Opener in Canada. As a result, it is one of the most diverse and multicultural area.

1.7) Foursquare API: Four-square API going to be use as its prime data gathering source as it has a database of millions of places, especially their places API which provides the ability to perform location search, location sharing and details about the business involvement.

1.8) Work Flow:

Foursquare API using for sorting features of near-by places of the neighborhoods. Due to http request limitations from Foursquare API, the number of places per neighborhood parameter would reasonably be set to 100 and the radius parameter would be set to 500.

1.9) Clustering Approach:

Unsupervised machine learning : k-means clustering algorithm going to be used for comparing the similarities of two cities, we decided to explore neighborhoods, segment them, and group them into clusters to find similar neighborhoods in a big city like Toronto and Montreal.

2) Data Description Acquisition and Cleaning :

Based on the background and problem analysis, following issues are going to be resolved by collection data, cleaning, clustering and analysis.

- To find the safest borough based on crime statistics
- finding the top ten common venues
- choosing the right neighborhood within the top borough.

As per geographical coordinates of Vancouver to plot neighborhoods in a borough that is safe and in the city's vicinity, and finally cluster the neighborhoods and present the research findings.

Below steps are followed to source the data file and processed to extract and generate the dataset as per project requirements:

2.1)Step1:

Using a real world data set from Kaggle containing the Vancouver Crimes from 2003 to 2019: A dataset consisting of the crime statistics of each Neighborhood in Vancouver along with type of crime, recorded year, month and hour. Since found the huge data (Around 600,000 rows) from the source data file from 2003 to 2019, this is considered to process dataset only from the year 2018.
Data set URL: https://www.kaggle.com/agilesifaka/vancouver-crime-report/version/2

2.2)Step2:

Sourcing more information of the list of officially categorized boroughs in Vancouver from Wikipedia. Borough information will be used to map the existing data where each neighborhood can be assigned with the right borough. As part of data set Borough, the neighborhood was not categorized, so created a dictionary of Neighborhood and based on data in the following Wikipedia page.

https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Vancouver

2.3)Step3:

From the source data, Creating a new consolidated dataset of the Neighborhoods, along with their boroughs, crime data and the respective Neighborhood's co-ordinates. This data will be fetched using OpenCage Geocoder to find the safest borough and explore the neighborhood by plotting it on maps using Folium library and perform exploratory data analysis.

2.4) Step 4:

After finding the list of neighborhoods, Foursquare API used for collecting data about different venues in different neighborhoods of that specific borough. Foursquare API used to get location data with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. we then connect to the Foursquare API to gather information about venues inside each and every neighborhood. For each neighborhood, chosen the radius from 100 to 500 meters.

Creating a new consolidated dataset of the Neighborhoods, boroughs, and the most common venues and the respective Neighborhood along with co-ordinates.: This data will be fetched using Four Square API to explore the neighborhood venues and to apply machine learning algorithm to cluster the neighborhoods and present the findings by plotting it on maps using Folium.

3) Methodology :

Methodology Categorized the by two parts:

3.1) Exploratory Data Analysis:

To identify the safest borough and normalize the neighborhoods of that borough visualize the crime reports in different Vancouver boroughs . By using the data result finding 10 most common venues in each neighborhood.

3.2) Modeling:

By using modeling Business Investor/Stakeholders/Small Business Owner / Franchisee Opener will have advantage to choose the right neighborhood within a borough which will be clustered similar neighborhoods using K - means clustering which is a form of unsupervised machine learning algorithm as per predefined cluster size. This K - means clustering will help in the decision-making process.

3.1.A) Exploratory Data Analysis

Pivoting the table to better understand the data by crimes per borough Table 1:

Type	Break and Enter Commercial	Break and Enter Residential/Other	Mischief	Other Theft	Theft from Vehicle	Theft of Bicycle	Theft of Vehicle	Vehicle Collision or Pedestrian Struck (with Fatality)	Vehicle Collision or Pedestrian Struck (with Injury)	Year All
Borough										
Central	787	198	2280	2489	6871	857	245	1	314	14042
East Side	786	1043	2192	1674	4754	678	605	8	660	12400
South Vancouver	49	156	187	88	483	36	71	1	111	1182
West Side	403	1000	1062	696	2838	588	225	3	389	7204
All	2025	2397	5721	4947	14946	2159	1146	13	1474	34828

Merging the Pivoted Column with other columns . Table 2:

	Borough	YearBreak and Enter Commercial	YearBreak and Enter Residential/Other	YearMischief	YearOther Theft	YearTheft from Vehicle	YearTheft of Bicycle	YearTheft of Vehicle	YearVehicle Collision or Pedestrian Struck (with Fatality)	YearVehicle Collision or Pedestrian Struck (with Injury)	Total
0	Central	787	198	2280	2489	6871	857	245	1	314	14042
1	East Side	786	1043	2192	1674	4754	678	605	8	660	12400
2	South Vancouver	49	156	187	88	483	36	71	1	111	1182
3	West Side	403	1000	1062	696	2838	588	225	3	389	7204

Pivoting the table to better understand the data by crimes per neighborhood Table 3:

Type	Break and Enter Commercial	Break and Enter Residential/Other	Mischief	Other Theft	Theft from Vehicle	Theft of Bicycle	Theft of Vehicle	Vehicle Collision or Pedestrian Struck (with Fatality)	Vehicle Collision or Pedestrian Struck (with Injury)	Year All
Neighbourhood										
Arbutus Ridge	12	78	49	18	111	12	12	1	18	311
Central Business District	551	124	1812	2034	5301	640	165	0	230	10857
Dunbar-Southlands	8	106	81	31	199	16	9	1	23	474
Fairview	138	73	233	297	692	245	55	0	62	1795
Grandview-Woodland	148	162	304	215	634	110	123	0	65	1761
Hastings-Sunrise	48	117	195	107	607	52	74	0	70	1270
Kensington-Cedar Cottage	62	145	255	148	541	69	71	3	97	1391
Kerrisdale	24	97	49	9	172	13	11	0	42	417
Killarney	34	72	90	31	240	19	33	0	46	565
Kitsilano	106	165	320	154	755	189	51	1	61	1802
Marpole	44	125	134	75	290	34	39	0	87	828
Mount Pleasant	205	124	353	493	822	232	67	0	100	2396
Musqueam	0	4	3	0	4	2	2	0	2	17
Oakridge	19	123	64	63	164	18	18	0	30	499
Renfrew-Collingwood	91	156	243	472	569	37	92	0	102	1762
Riley Park	35	122	140	53	378	52	39	2	45	866
Shaughnessy	12	120	41	0	187	10	11	0	33	414
South Cambie	22	42	41	38	111	19	8	0	11	292
Stanley Park	6	2	8	0	109	14	3	0	12	154
Strathcona	160	124	527	81	821	108	76	2	88	1987
Sunset	37	93	175	105	382	18	63	1	93	967
Victoria-Fraserview	15	80	94	57	239	15	36	1	63	600
West End	230	72	460	455	1461	203	77	1	72	3031
West Point Grey	18	71	50	11	157	32	11	0	22	372
All	2025	2397	5721	4947	14946	2159	1146	13	1474	34828

Merging the Pivoted Column with other columns Table 4:

	Neighbourhood	YearBreak and Enter Commercial	YearBreak and Enter Residential/Other	YearMischief	YearOther Theft	YearTheft from Vehicle	YearTheft of Bicycle	YearTheft of Vehicle	YearVehicle Collision or Pedestrian Struck (with Fatality)	YearVehicle Collision or Pedestrian Struck (with Injury)	Total
0	Arbutus Ridge	12	78	49	18	111	12	12	1	18	311
1	Central Business District	551	124	1812	2034	5301	640	165	0	230	10857
2	Dunbar-Southlands	8	106	81	31	199	16	9	1	23	474
3	Fairview	138	73	233	297	692	245	55	0	62	1795
4	Grandview-Woodland	148	162	304	215	634	110	123	0	65	1761

Pandas describe() is used to view some basic statistical details like percentile, mean, std etc. of a data frame or a series of numeric values. Table 5

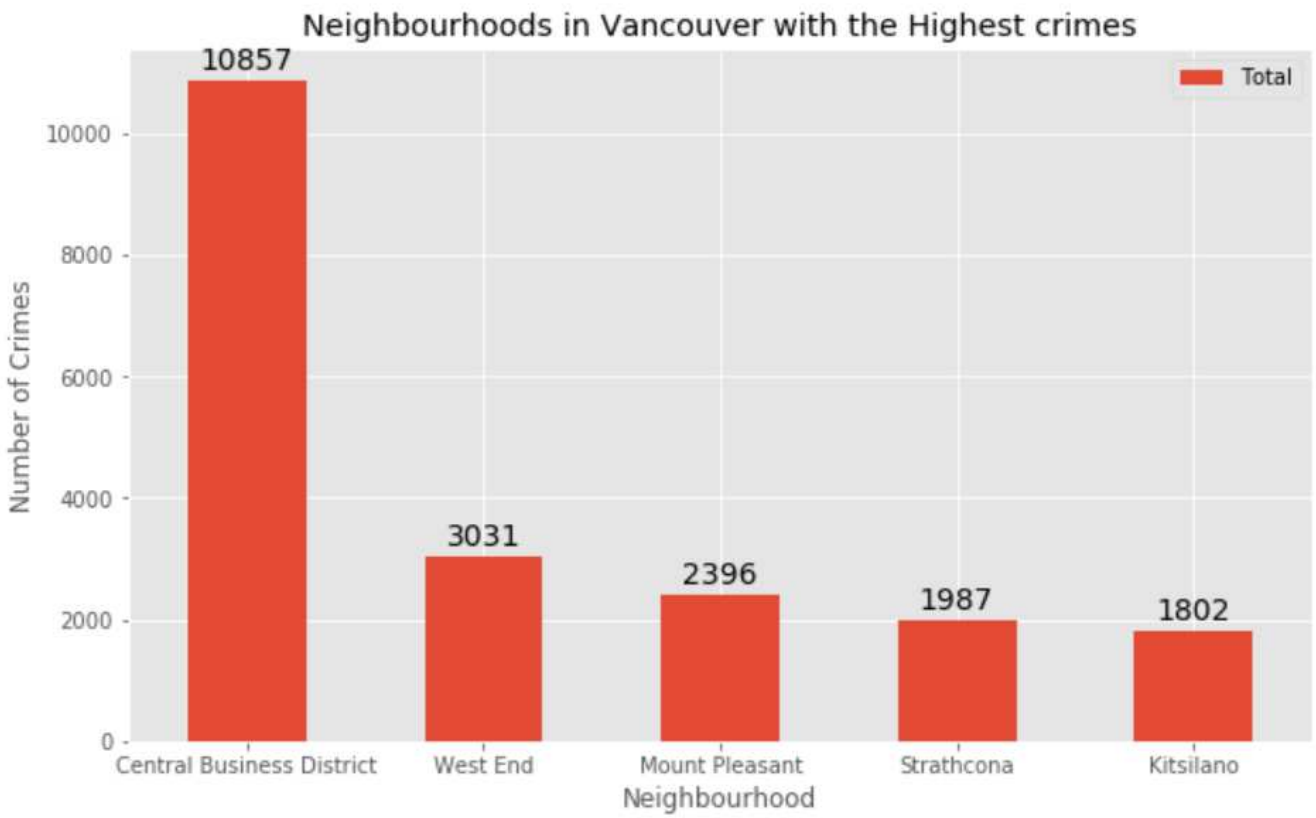
	YearBreak and Enter Commercial	YearBreak and Enter Residential/Other	YearMischief	YearOther Theft	YearTheft from Vehicle	YearTheft of Bicycle	YearTheft of Vehicle	YearVehicle Collision or Pedestrian Struck (with Fatality)	YearVehicle Collision or Pedestrian Struck (with Injury)	Total
count	4.000000	4.000000	4.00000	4.000000	4.000000	4.000000	4.000000	4.000000	4.000000	4.000000
mean	506.250000	599.250000	1430.25000	1236.750000	3736.500000	539.750000	286.500000	3.250000	368.500000	8707.000000
std	354.409721	488.189427	997.26572	1060.087221	2723.536977	353.955153	226.117226	3.304038	227.060198	5801.870618
min	49.000000	156.000000	187.00000	88.000000	483.000000	36.000000	71.000000	1.000000	111.000000	1182.000000
25%	314.500000	187.500000	843.25000	544.000000	2249.250000	450.000000	186.500000	1.000000	263.250000	5698.500000
50%	594.500000	599.000000	1627.00000	1185.000000	3796.000000	633.000000	235.000000	2.000000	351.500000	9802.000000
75%	786.250000	1010.750000	2214.00000	1877.750000	5283.250000	722.750000	335.000000	4.250000	456.750000	12810.500000
max	787.000000	1043.000000	2280.00000	2489.000000	6871.000000	857.000000	605.000000	8.000000	660.000000	14042.000000

Exploring the data by Visualizing:

Sorting the data by crimes per neighborhood. Table 6:

	Neighbourhood	YearBreak and Enter Commercial	YearBreak and Enter Residential/Other	YearMischief	YearOther Theft	YearTheft from Vehicle	YearTheft of Bicycle	YearTheft of Vehicle	YearVehicle Collision or Pedestrian Struck (with Fatality)	YearVehicle Collision or Pedestrian Struck (with Injury)	Total
1	Central Business District	551	124	1812	2034	5301	640	165	0	230	10857
22	West End	230	72	460	455	1461	203	77	1	72	3031
11	Mount Pleasant	205	124	353	493	822	232	67	0	100	2396
19	Strathcona	160	124	527	81	821	108	76	2	88	1987
9	Kitsilano	106	165	320	154	755	189	51	1	61	1802

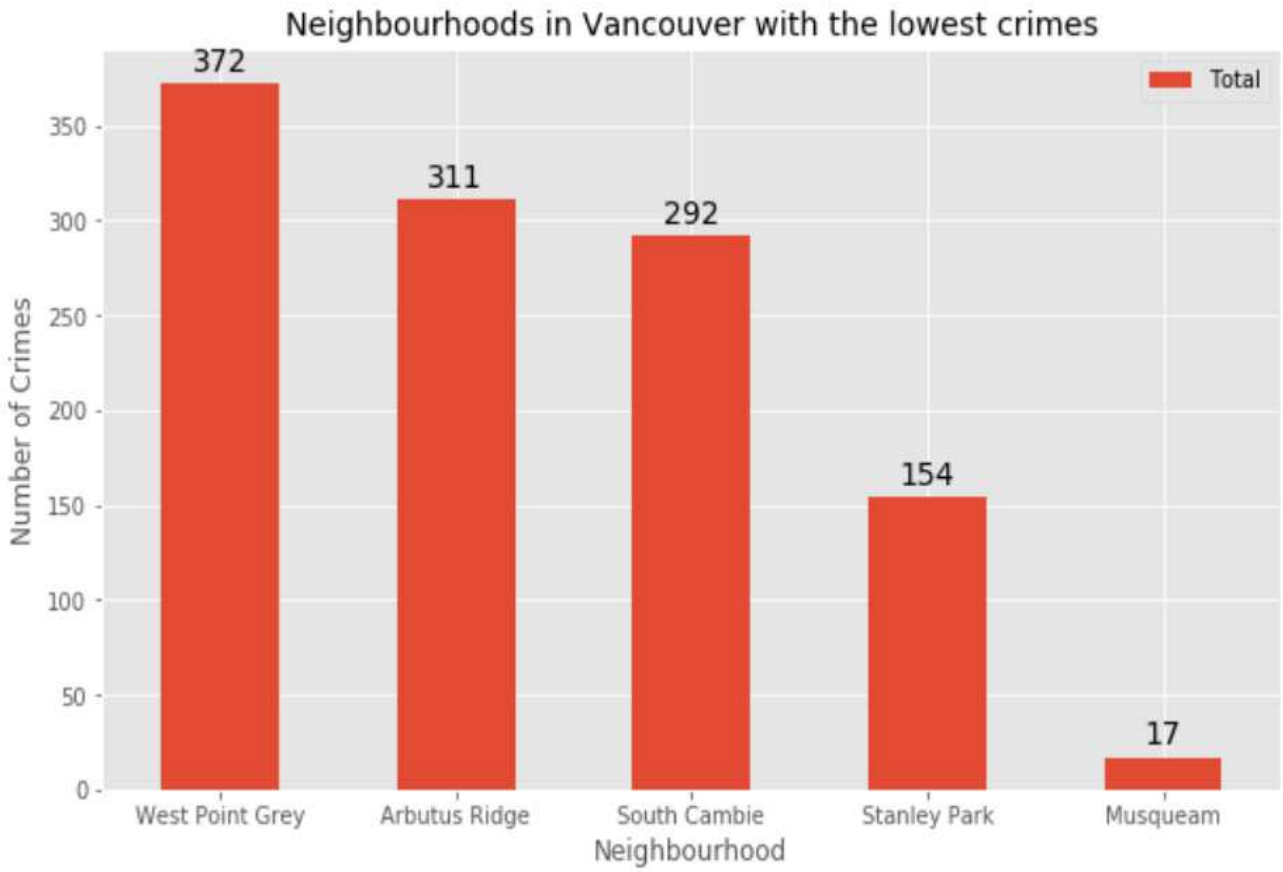
Five Neighborhoods with highest crime. Figure 1:



Five Neighborhoods with lowest crime Table 7:

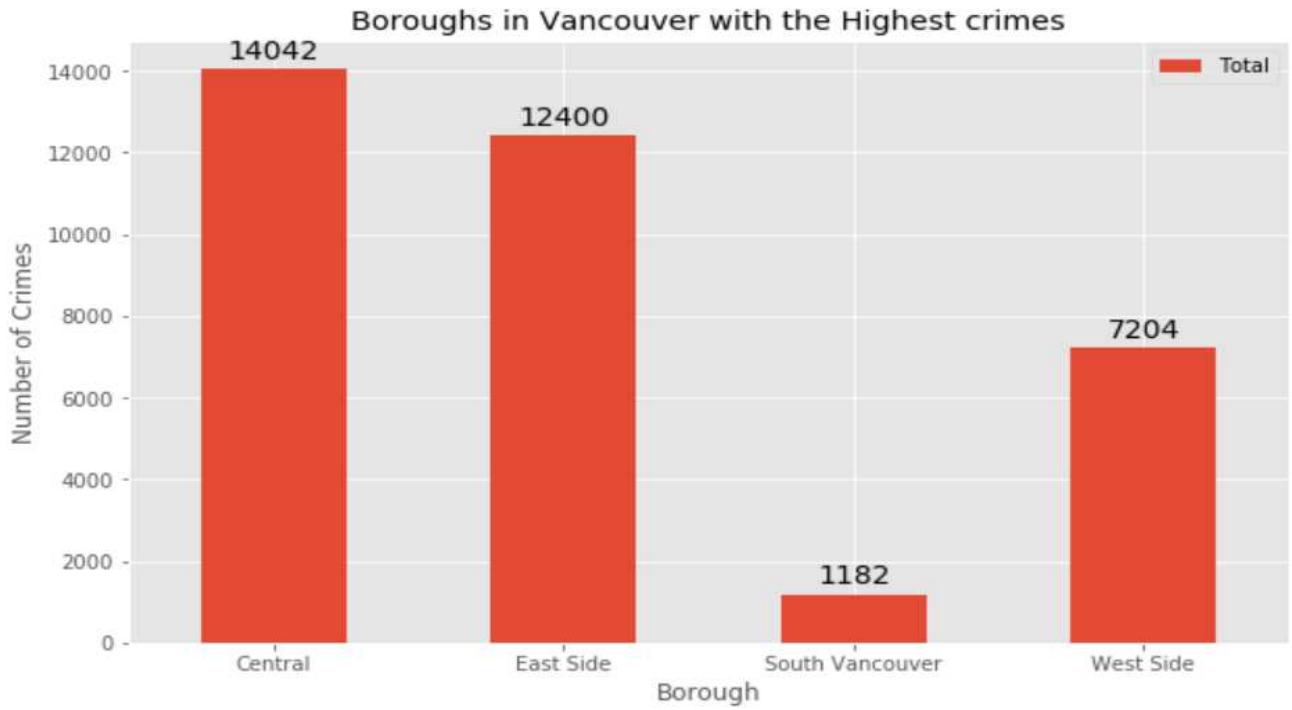
	Neighbourhood	YearBreak and Enter Commercial	YearBreak and Enter Residential/Other	YearMischief	YearOther Theft	YearTheft from Vehicle	YearTheft of Bicycle	YearTheft of Vehicle	YearVehicle Collision or Pedestrian Struck (with Fatality)	YearVehicle Collision or Pedestrian Struck (with Injury)	Total
23	West Point Grey	18	71	50	11	157	32	11	0	22	372
0	Arbutus Ridge	12	78	49	18	111	12	12	1	18	311
17	South Cambie	22	42	41	38	111	19	8	0	11	292
18	Stanley Park	6	2	8	0	109	14	3	0	12	154
12	Musqueam	0	4	3	0	4	2	2	0	2	17

Five Neighborhoods with lowest crime Table Figure 2:



Borough in Vancouver with Highest Crime Table 8 :

	level_0	index	Borough	Break and Enter Commercial	Break and Enter Residential	Mischief	Other	Theft from Vehicle	Theft of Bicycle	Theft of Vehicle	Vehicle Collision or Pedestrian Struck (with Fatality)	Vehicle Collision or Pedestrian Struck (with Injury)	Total
0	0	0	Central	787	198	2280	2489	6871	857	245	1	314	14042
1	1	1	East Side	786	1043	2192	1674	4754	678	605	8	660	12400
2	2	2	South Vancouver	49	156	187	88	483	36	71	1	111	1182
3	3	3	West Side	403	1000	1062	696	2838	588	225	3	389	7204

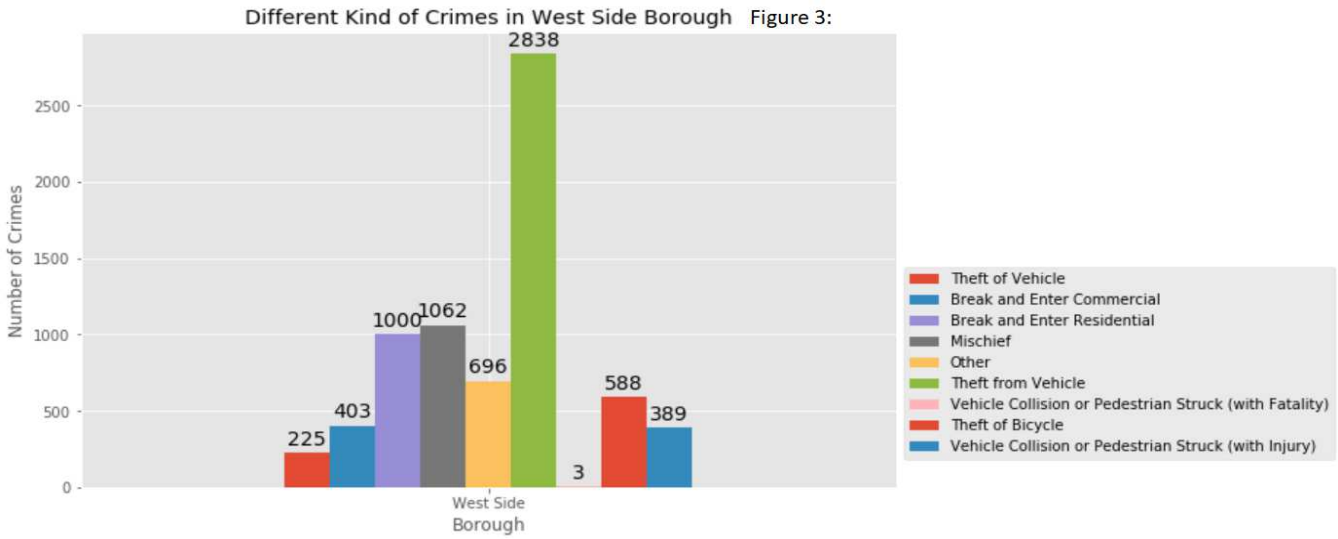


Through the exploratory data analysis, it is clear that South Vancouver has the lowest crimes.

Even though South Vancouver has the lowest crimes but very little number of neighborhoods which doesn't support opening a daily convenient store 24/7 for establishment, so next choice is 2nd lowest crime which is West Side.

Different types of crimes recorded in the West Side Borough Figure 4:

West side was chosen because crime type Break and enter Commercial is also low amongst other crimes types which makes West Side ideal destination for opening of commercial establishments



Step3:

Creating a new consolidated dataset of the Neighborhoods, along with their boroughs, crime data and the respective Neighborhood's co-ordinates.

This data will be fetched using Opencage Geocoder to find the safest borough and explore the neighborhood by plotting it on maps using Folium and perform exploratory data analysis.

Restricting the rows in the data frame to only those with West side as Borough.

Creating a new Data frame with Lat, Lng being fetched from OpenCage geocoder.

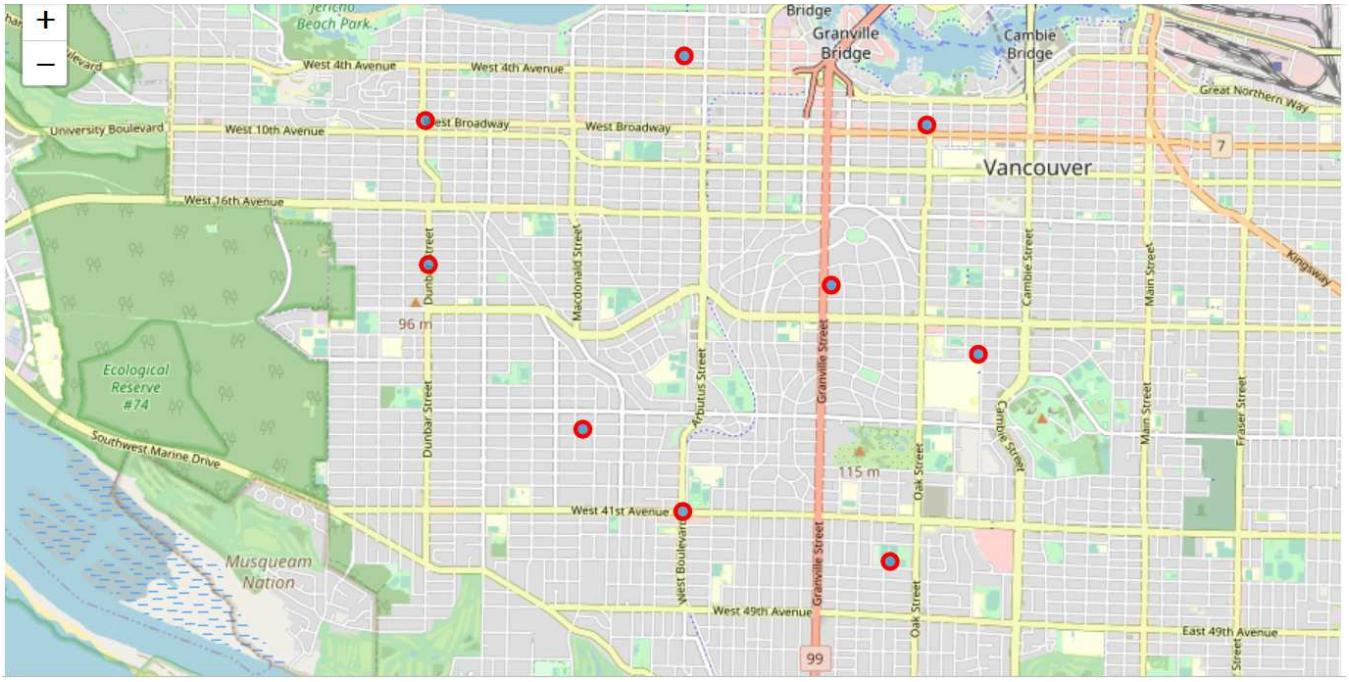
Glimpse of the new Data Frame with Neighborhoods in West Side Borough of Vancouver along with centroid of their co-ordinates Table 10:

	Neighbourhood	Borough	Latitude	Longitude
0	Shaughnessy	West Side	49.251863	-123.138023
1	Fairview	West Side	49.264113	-123.126835
2	Oakridge	West Side	49.230829	-123.131134
3	Marpole	West Side	49.209223	-123.136150
4	Kitsilano	West Side	49.269410	-123.155267
5	Kerrisdale	West Side	49.234673	-123.155389
6	West Point Grey	West Side	49.264484	-123.185433
7	Arbutus Ridge	West Side	49.240968	-123.167001
8	South Cambie	West Side	49.246685	-123.120915
9	Dunbar-Southlands	West Side	49.253460	-123.185044

Fetching the Geographical co-ordinates of Vancouver to plot on Map 1 :

The geographical coordinate of Vancouver, Canada are 49.2608724, -123.1139529.

Using Folium to plot Vancouver City's West Side Borough and it's Neighborhoods



Step 4:

Creating a new consolidated dataset of the Neighborhoods, boroughs, and the most common venues and the respective Neighborhood along with co-ordinates. This data will be fetched using Four Square API to explore the neighborhood venues and to apply machine learning algorithm to cluster the neighborhoods and present the findings by plotting it on maps using Folium.

Setting Up Foursquare Credentials

Your credentials:

Defining a function to fetch top 10 venues around a given neighborhood

Generating Venues

Shaughnessy
Fairview
Oakridge
Marpole
Kitsilano
Kerrisdale
West Point Grey
Arbutus Ridge
South Cambie

Data frame containing venues for each neighborhood in West Side:

Venue Count per neighborhood Table 11:

Neighbourhood	Venue
Arbutus Ridge	5
Dunbar-Southlands	9
Fairview	26
Kerrisdale	37
Kitsilano	47
Marpole	30
Oakridge	7
Shaughnessy	2
South Cambie	18
West Point Grey	42

There are 85 unique categories.

3.2) Modeling :

One Hot Encoding to Analyze Each Neighborhood Table 12:

	Neighbourhood	American Restaurant	Asian Restaurant	BBQ Joint	Bakery	Bank	Bar	Beach	Bookstore	Boutique	...	Taiwanese Restaurant	Tea Room	Tennis Court	Thai Restaurant	Thrift / Vintage Store	Vegetarian / Vegan Restaurant	Video Store	Vietnamese Restaurant	Wine Shop	Yoga Studio
0	Shaughnessy	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
1	Shaughnessy	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
2	Fairview	0	0	1	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
3	Fairview	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
4	Fairview	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0

5 rows × 86 columns

Group By Table 13:

	Neighbourhood	American Restaurant	Asian Restaurant	BBQ Joint	Bakery	Bank	Bar	Beach	Bookstore	Boutique	...	Taiwanese Restaurant	Tea Room	Tennis Court	Thai Restaurant	Thrift / Vintage Store	Vegetarian / Vegan Restaurant	Video Store	Vietnamese Restaurant	Wine Shop	Yoga Studio
0	Arbutus Ridge	0.000000	0.000000	0.000000	0.200000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.00000	0.000000	0.00000	0.000000	0.00000	0.000000
1	Dunbar-Southlands	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.00000	0.000000	0.00000	0.000000	0.00000	0.000000
2	Fairview	0.000000	0.076923	0.038462	0.000000	0.038462	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.00000	0.000000	0.00000	0.038462	0.00000	0.000000
3	Kerrisdale	0.000000	0.027027	0.000000	0.027027	0.027027	0.000000	0.000000	0.000000	0.054054	...	0.000000	0.054054	0.000000	0.027027	0.00000	0.000000	0.00000	0.027027	0.00000	0.000000
4	Kitsilano	0.042553	0.021277	0.000000	0.063830	0.000000	0.000000	0.021277	0.000000	0.000000	...	0.000000	0.021277	0.021277	0.042553	0.00000	0.021277	0.00000	0.000000	0.00000	0.021277
5	Marpole	0.000000	0.000000	0.000000	0.000000	0.033333	0.033333	0.000000	0.000000	0.000000	...	0.033333	0.000000	0.000000	0.033333	0.00000	0.000000	0.00000	0.033333	0.00000	0.000000
6	Oakridge	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.00000	0.000000	0.00000	0.142857	0.00000	0.000000
7	Shaughnessy	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.00000	0.000000	0.00000	0.000000	0.00000	0.000000
8	South Cambie	0.000000	0.000000	0.000000	0.000000	0.055556	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.00000	0.000000	0.00000	0.055556	0.00000	0.000000
9	West Point Grey	0.000000	0.047619	0.000000	0.023810	0.023810	0.023810	0.000000	0.047619	0.000000	...	0.000000	0.000000	0.000000	0.000000	0.02381	0.047619	0.02381	0.000000	0.02381	0.023810

10 rows × 86 columns

Top 5 most common venues across neighborhoods :

----Arbutus Ridge----

	venue	freq
0	Grocery Store	0.2
1	Pet Store	0.2
2	Bakery	0.2
3	Nightlife Spot	0.2
4	Spa	0.2

----Dunbar-Southlands----

	venue	freq
0	Sushi Restaurant	0.22
1	Italian Restaurant	0.11
2	Coffee Shop	0.11
3	Sporting Goods Shop	0.11
4	Ice Cream Shop	0.11

----Fairview----

	venue	freq
0	Coffee Shop	0.15
1	Park	0.08
2	Asian Restaurant	0.08
3	Sandwich Place	0.04
4	Camera Store	0.04

----Kerrisdale----

	venue	freq
0	Coffee Shop	0.11
1	Chinese Restaurant	0.08
2	Sandwich Place	0.05
3	Pharmacy	0.05

4 Boutique 0.05

----Kitsilano----

	venue	freq
0	Bakery	0.06
1	American Restaurant	0.04
2	Restaurant	0.04
3	Ice Cream Shop	0.04
4	Japanese Restaurant	0.04

----Marpole----

	venue	freq
0	Sushi Restaurant	0.10
1	Bubble Tea Shop	0.07
2	Pizza Place	0.07
3	Bus Stop	0.07
4	Chinese Restaurant	0.07

----Oakridge----

	venue	freq
0	Sushi Restaurant	0.14
1	Pharmacy	0.14
2	Fast Food Restaurant	0.14
3	Vietnamese Restaurant	0.14
4	Convenience Store	0.14

----Shaughnessy----

	venue	freq
0	French Restaurant	0.5
1	Park	0.5
2	American Restaurant	0.0
3	Pet Store	0.0
4	Record Shop	0.0

----South Cambie----

	venue	freq
0	Coffee Shop	0.28
1	Park	0.11
2	Liquor Store	0.06
3	Bus Stop	0.06
4	Light Rail Station	0.06

----West Point Grey----

	venue	freq
0	Coffee Shop	0.10
1	Café	0.07
2	Sushi Restaurant	0.07
3	Japanese Restaurant	0.07
4	Vegetarian / Vegan Restaurant	0.05

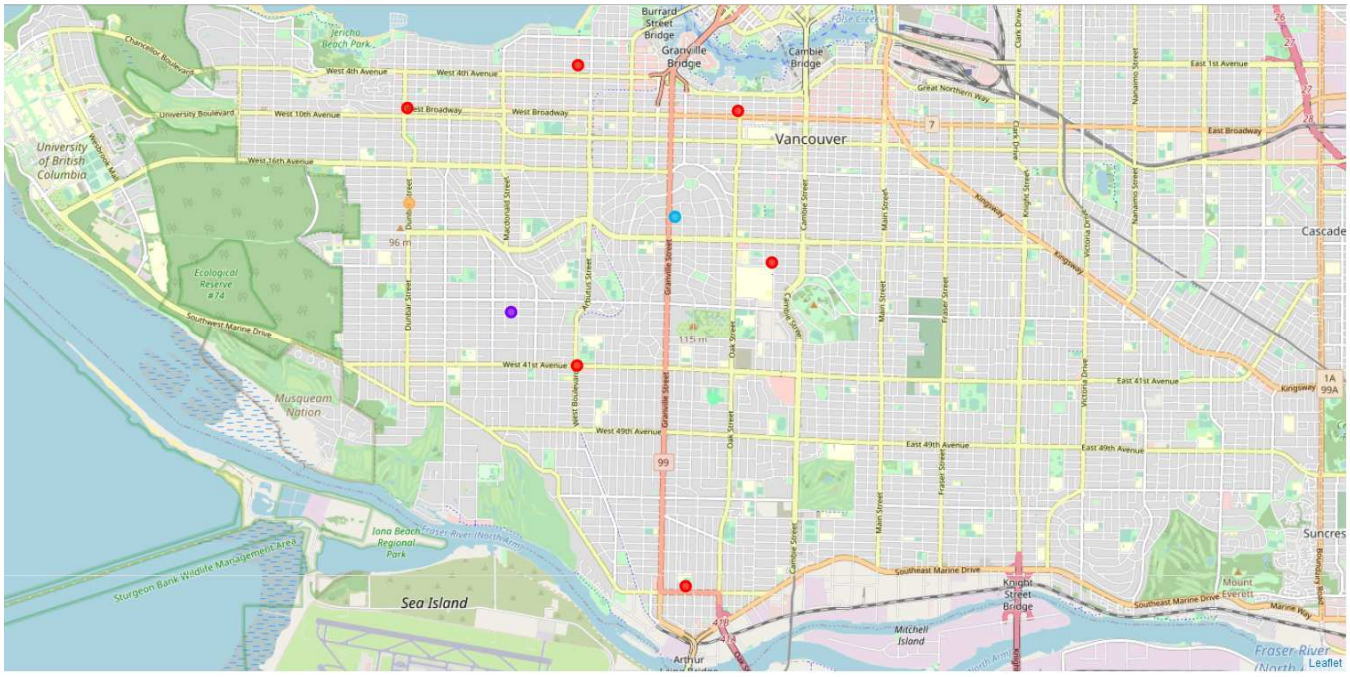
Now let's create the new data frame and display the top 10 venues for each neighborhood. Table 14

	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	Arbutus Ridge	Nightlife Spot	Spa	Bakery	Grocery Store	Pet Store	Dessert Shop	Dim Sum Restaurant	Diner	Falafel Restaurant	Fast Food Restaurant
1	Dunbar-Southlands	Sushi Restaurant	Liquor Store	Italian Restaurant	Indian Restaurant	Ice Cream Shop	Coffee Shop	Sporting Goods Shop	Japanese Restaurant	Bookstore	Gastropub
2	Fairview	Coffee Shop	Park	Asian Restaurant	Chinese Restaurant	Malay Restaurant	Diner	Falafel Restaurant	Nail Salon	Camera Store	Restaurant
3	Kerrisdale	Coffee Shop	Chinese Restaurant	Sandwich Place	Pharmacy	Boutique	Tea Room	Sushi Restaurant	Hobby Shop	Portuguese Restaurant	Pizza Place
4	Kitsilano	Bakery	American Restaurant	Coffee Shop	Restaurant	Japanese Restaurant	Ice Cream Shop	French Restaurant	Sushi Restaurant	Food Truck	Thai Restaurant

Clustering Neighborhoods. Table 15, Map 2:

(set number of clusters = 5)

	Neighbourhood	Borough	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	Shaughnessy	West Side	49.251863	-123.138023	2	Park	French Restaurant	Yoga Studio	Dessert Shop	Dim Sum Restaurant	Diner	Falafel Restaurant	Fast Food Restaurant	Food Truck	Gas Station
1	Fairview	West Side	49.264113	-123.126835	0	Coffee Shop	Park	Asian Restaurant	Chinese Restaurant	Malay Restaurant	Diner	Falafel Restaurant	Nail Salon	Camera Store	Restaurant
2	Oakridge	West Side	49.230829	-123.131134	3	Sushi Restaurant	Concert Hall	Sandwich Place	Pharmacy	Fast Food Restaurant	Convenience Store	Vietnamese Restaurant	Gym Pool	Falafel Restaurant	Deli / Bodega
3	Marpole	West Side	49.209223	-123.136150	0	Sushi Restaurant	Pizza Place	Chinese Restaurant	Bus Stop	Bubble Tea Shop	Liquor Store	Gas Station	Dessert Shop	Coffee Shop	Plaza
4	Kitsilano	West Side	49.269410	-123.155267	0	Bakery	American Restaurant	Coffee Shop	Restaurant	Japanese Restaurant	Ice Cream Shop	French Restaurant	Sushi Restaurant	Food Truck	Thai Restaurant



4) Analysis: Examining the result through 5 Clusters Table 16:

Cluster 1

	Borough	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 Side	Coffee Shop	Park	Asian Restaurant	Chinese Restaurant	Malay Restaurant	Diner	Falafel Restaurant	Nail Salon	Camera Store	Restaurant
3	West Side	Sushi Restaurant	Pizza Place	Chinese Restaurant	Bus Stop	Bubble Tea Shop	Liquor Store	Gas Station	Dessert Shop	Coffee Shop	Plaza
4	West Side	Bakery	American Restaurant	Coffee Shop	Restaurant	Japanese Restaurant	Ice Cream Shop	French Restaurant	Sushi Restaurant	Food Truck	Thai Restaurant
5	West Side	Coffee Shop	Chinese Restaurant	Sandwich Place	Pharmacy	Boutique	Tea Room	Sushi Restaurant	Hobby Shop	Portuguese Restaurant	Pizza Place
6	West Side	Coffee Shop	Japanese Restaurant	Café	Sushi Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant	Bookstore	Pub	Yoga Studio	Gastropub
8	West Side	Coffee Shop	Park	Liquor Store	Bus Stop	Vietnamese Restaurant	Light Rail Station	Bank	Malay Restaurant	Grocery Store	Sushi Restaurant

Cluster 2:

	Borough	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 Side	Nightlife Spot	Spa	Bakery	Grocery Store	Pet Store	Dessert Shop	Dim Sum Restaurant	Diner	Falafel Restaurant	Fast Food Restaurant

Cluster 3:

	Borough	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	West Side	Park	French Restaurant	Yoga Studio	Dessert Shop	Dim Sum Restaurant	Diner	Falafel Restaurant	Fast Food Restaurant	Food Truck	Gas Station

Cluster 4:

	Borough	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	West Side	Sushi Restaurant	Concert Hall	Sandwich Place	Pharmacy	Fast Food Restaurant	Convenience Store	Vietnamese Restaurant	Gym Pool	Falafel Restaurant	Deli / Bodega

Cluster 5:

	Borough	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
9	West Side	Sushi Restaurant	Liquor Store	Italian Restaurant	Indian Restaurant	Ice Cream Shop	Coffee Shop	Sporting Goods Shop	Japanese Restaurant	Bookstore	Gastropub

5) Results and Discussion:

Purpose of the project to find out the safest borough in Vancouver, and an appropriate neighborhood within the borough for Business Investor/Stakeholders/Small Business Owner / Franchisee Opener to establish Daily Convenient Store/Retail store/Franchisee which will be operated 24 Hours a Day.

It was obtained by analyzing Vancouver crime data Neighborhood to identify a safe borough with considerable number of neighborhoods for any business to be viable. After selecting the borough, it was imperative to choose the right neighborhood where daily convenient store 24/7 were not among venues in a close proximity to each other. Successfully achieved this purpose by using necessary data science tools by grouping the neighborhoods into clusters to assist the Business Investor/Stakeholders/Small Business Owner / Franchisee Opener by providing them with relevant data about venues and safety of a given neighborhood. West Side as an ideal destination for opening and establishing Daily Convenient Store/Retail store/Franchisee which will be operated 24 hours a day (24/7).

6) Conclusion :

Through this analysis of the crime data it is clear that different types of crimes happening in all neighborhoods of Vancouver due to project requirements categorized them into different boroughs, which helped to group the neighborhoods into boroughs and choose the safest borough first. Once we confirmed the borough the number of neighborhoods for consideration also comes down, then further shortlist the neighborhoods based on the less crime, common venues, sufficient neighborhood which is **West Side Borough** the best suits for Business Investor/Stakeholders/Small Business Owner / Franchisee Opener to establish Daily Convenient Store/Retail store/Franchisee which will be operated 24 hours a day (24/7).

West side was chosen because crime type Break and enter Commercial is also low amongst other crimes types, which makes West Side as an ideal destination for opening and establishing Daily Convenient Store/Retail store/Franchisee which will be operated 24 hours a day (24/7).

Project Completed at 3rd July 2020.

Project Prepared By:
Wahid Rahman
Data Scientist.

References:

- [1] https://www.kaggle.com/agilesifaka/vancouver-crime-report/version/2
- [2] https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Vancouver
- [3] https://developer.foursquare.com/
- [4] https://opencagedata.com
- [5] https://www.google.com/maps/