Capstone Project
The Battle of Neighbourhoods

Applied Data Science Capstone by IBM/Coursera

Introduction

- The aim of this project is to find a safe and secure location for opening of commercial establishments in Vancouver, Canada. Specifically, this report will be targeted to stakeholders interested in opening any business place like Grocery Store in Vancouver City, Canada.
- The first task would be to choose the safest borough by analyzing crime data for opening a grocery store and short listing a neighborhood, where grocery stores are not amongst the most common venues, and yet as close to the city as possible.
- We will make use of our data science tools to analyze data and focus on the safest borough and explore its neighborhoods and the 10 most common venues in each neighborhood so that the best neighborhood where grocery store is not amongst the most common venue can be selected.

Data

Based on definition of our problem, factors that will influence our decision are:

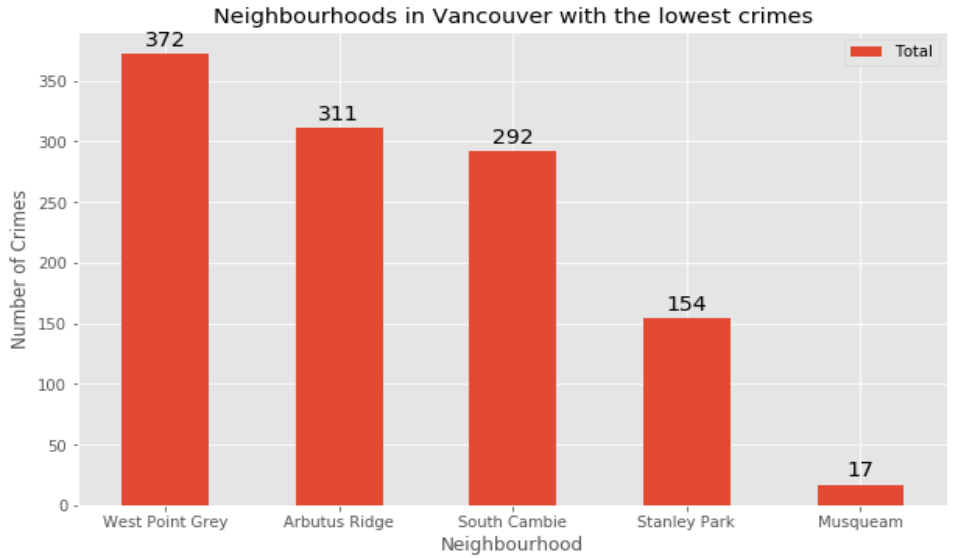
- Finding the safest borough based on crime statistics
- Finding the most common venues
- Choosing the right neighborhood within the borough

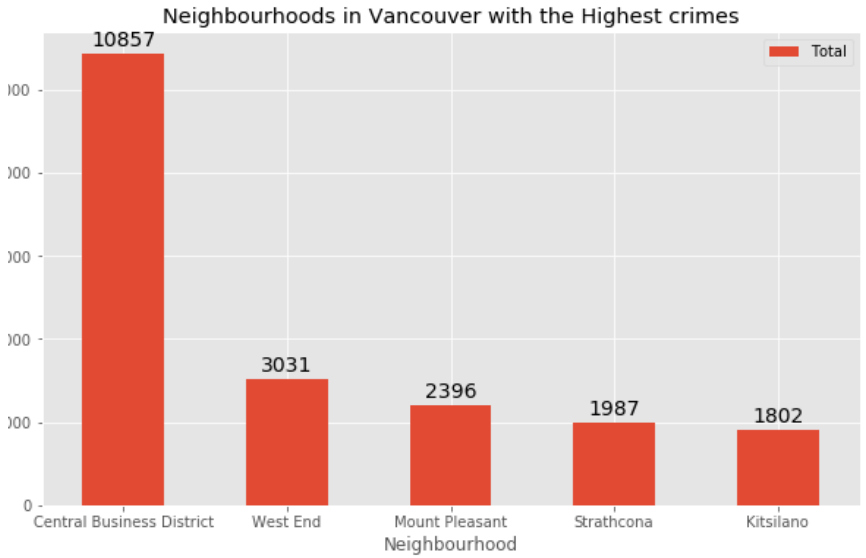
We will be using the geographical coordinates of Vancouver to plot neighborhoods in a borough that is safe and in the city's vicinity, and finally cluster our neighborhoods and present our findings.

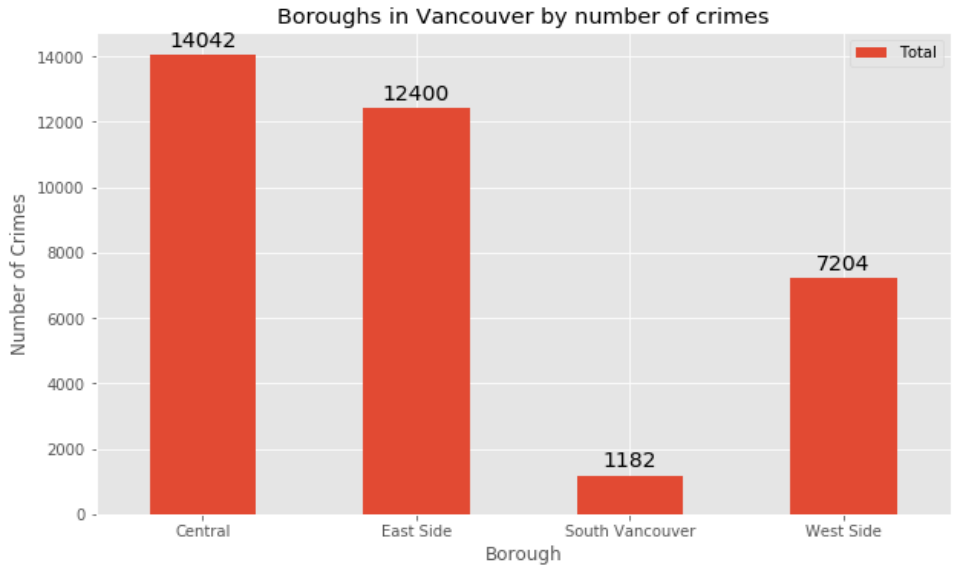
Data -Working the dataset

- 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.
- From the Vancouver Crime Report, year 2018 is considered as it is the latest year to be complete in the dataset. From Data set URL: https://www.kaggle.com/agilesifaka/vancouver-crime-report/version/2
- Additional data is collected 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.
- Creating a new consolidated dataset of the Neighborhoods, along with their boroughs, crime data and the respective Neighborhood's coordinates.
- Data is also collected from Foursquare API to find the most common venues and the respective neighborhood along with co-ordinates

Crimes Exploratory analysis

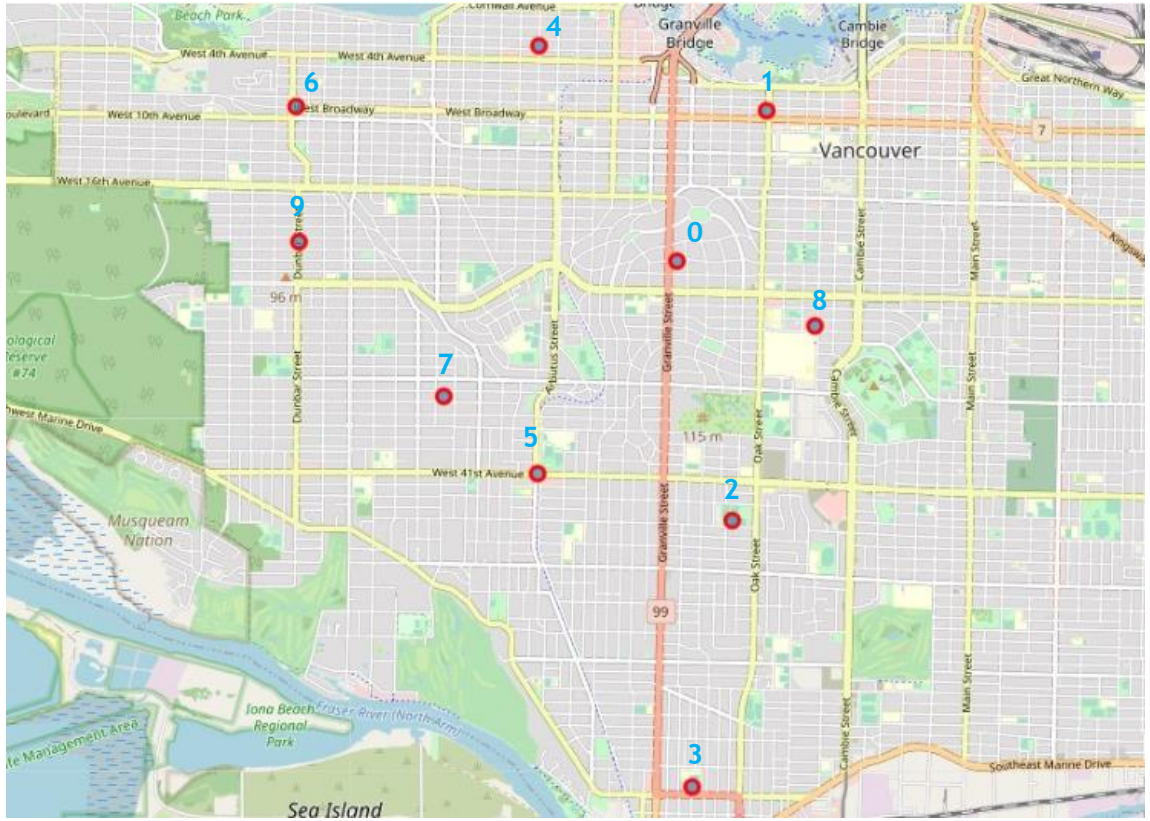






- Based on the data collected West Side Borough would be the optimal choice excluding South Vancouver for limited amount of neighborhoods.

West Side Vancouver - plotted neighbourhoods map



	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

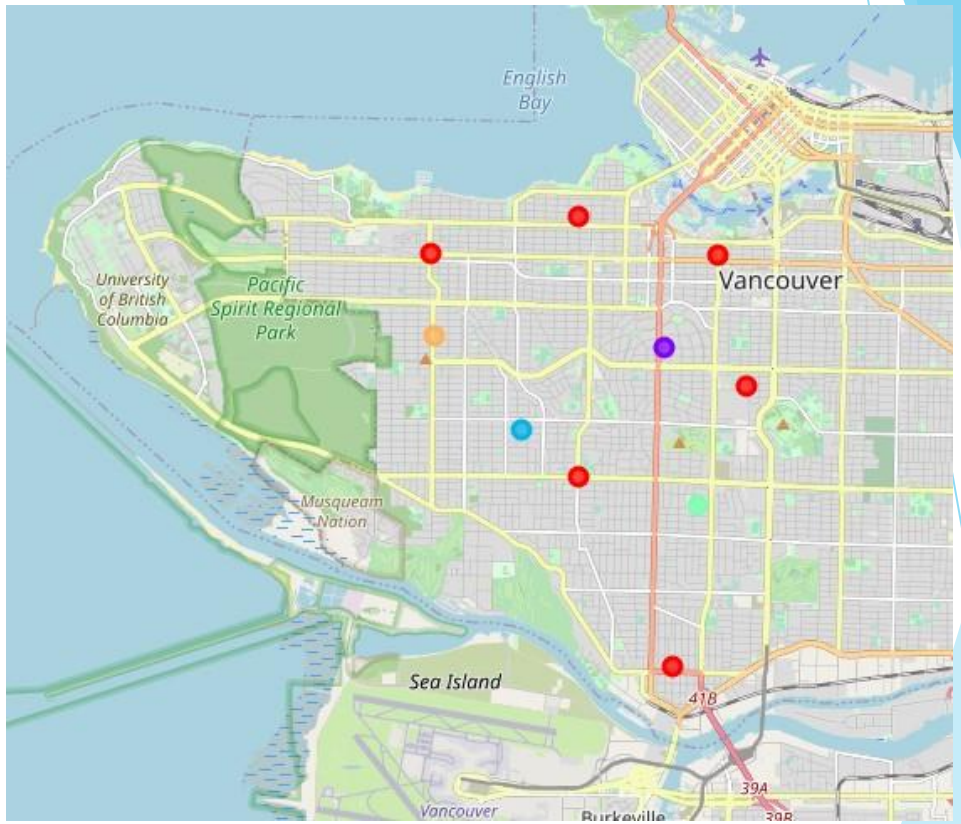
Venues Exploratory Analysis

- Common venues are put together with latitude and longitude, so that geospatial clusters can be created, and an analysis is possible.

	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
0	Shaughnessy	West Side	49.251863	-123.138023	1	Park	French Restaurant	Yoga Studio	Dim Sum Restaurant	Diner	Falafel Restaurant	Fast Food Restaurant	Food Truck
1	Fairview	West Side	49.264113	-123.126835	0	Coffee Shop	Asian Restaurant	Park	Chinese Restaurant	Sandwich Place	Japanese Restaurant	Korean Restaurant	Malay Restaurant
2	Oakridge	West Side	49.230829	-123.131134	3	Sushi Restaurant	Vietnamese Restaurant	Pharmacy	Sandwich Place	Fast Food Restaurant	Park	Convenience Store	French Restaurant
3	Marpole	West Side	49.209223	-123.136150	0	Vietnamese Restaurant	Pizza Place	Sushi Restaurant	Chinese Restaurant	Liquor Store	Shanghai Restaurant	Park	Dessert Shop
4	Kitsilano	West Side	49.269410	-123.155267	0	Bakery	American Restaurant	Coffee Shop	Japanese Restaurant	Ice Cream Shop	French Restaurant	Food Truck	Sushi Restaurant

Formed Clusters

	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Cluster 0	Coffee Shop	Asian Restaurant	Park	Chinese Restaurant	Sandwich Place
	Vietnamese Restaurant	Pizza Place	Sushi Restaurant	Chinese Restaurant	Liquor Store
	Bakery	American Restaurant	Coffee Shop	Japanese Restaurant	Ice Cream Shop
	Coffee Shop	Chinese Restaurant	Sandwich Place	Sushi Restaurant	Pharmacy
	Sushi Restaurant	Japanese Restaurant	Coffee Shop	Café	Bookstore
	Coffee Shop	Bus Stop	Sushi Restaurant	Malay Restaurant	Grocery Store
Cluster 1	Park	French Restaurant	Yoga Studio	Dim Sum Restaurant	Diner
Cluster 2	Nightlife Spot	Bakery	Pet Store	Grocery Store	Spa
Cluster 3	Sushi Restaurant	Vietnamese Restaurant	Pharmacy	Sandwich Place	Fast Food Restaurant
Cluster 4	Coffee Shop	Italian Restaurant	Indian Restaurant	Ice Cream Shop	Sushi Restaurant



Results – Discussion - Conclusion

- Based on the data if stakeholders in the area were to open a Grocery store, the recommendation would be to open it in cluster 1-3-4 since Grocery store does not make the top 5 list in the area.
- Data could have been expanded and made more precise by using crimes per population, so that small boroughs can be analysed more effectively against larger ones. However, this is made harder since every few years the limits and shape of neighbourhoods and boroughs are changed making comparison using population ineffective.
- Research is also limited by the scope of assignment.