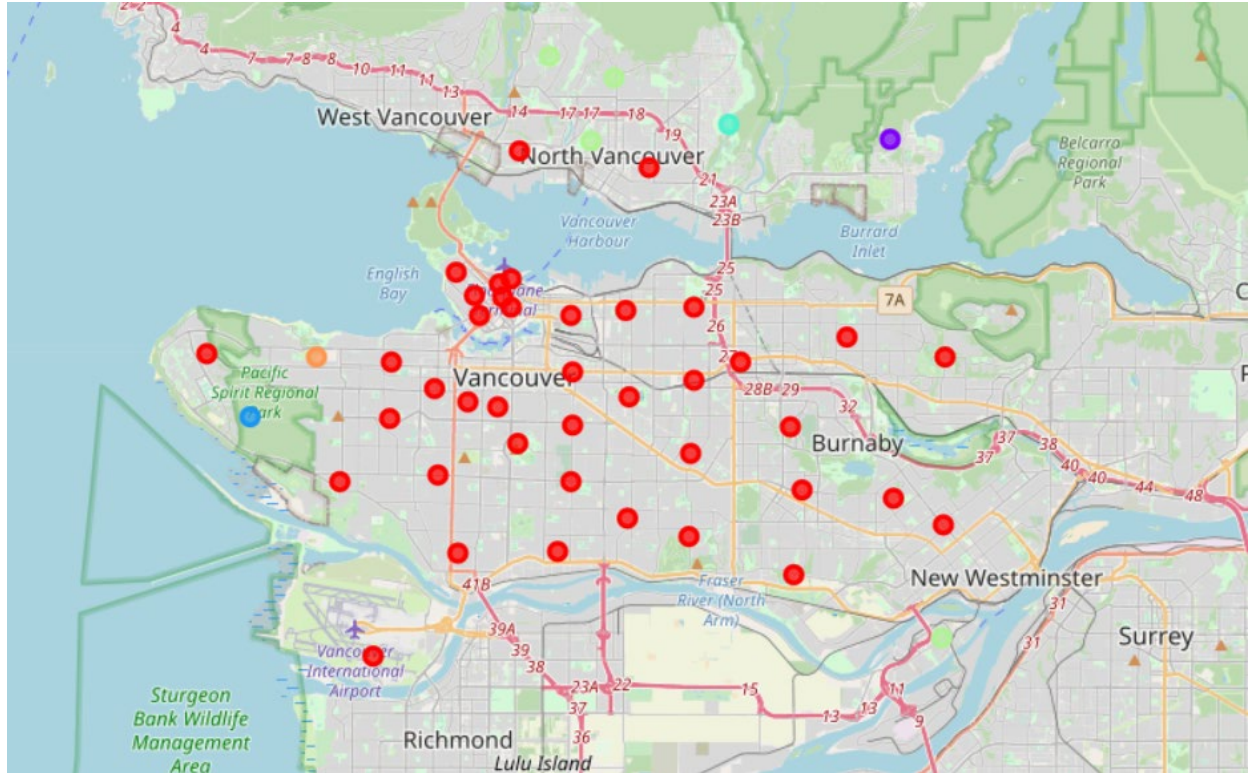# The Battle of Neighborhoods - Vancouver

Applied Data Science Capstone by IBM on Coursera

(Pooya Mirzabeygi)

1.   INTRODUCTION: BUSINESS PROBLEM

The purpose of this project is to use machine learning techniques to explore neighborhoods of Vancouver Canada in order to propose a potential recommendation for the location of a new restaurant or a business office. This project would help new entrepreneurs moving to Vancouver Canada establishing their business in a premium spot.

The k-mean algorithm which is an un-supervised clustering technique is used to divide the neighborhoods into meaningful categories. So that, the best neighborhood is recommended for the new location of the new restaurant. The Foursquare API is used to acquire data about the venue in each neighborhood.

For visualization purposes, Folium visualization library is utilized to map the location of each neighborhood and the other neighborhood residing in the same category or other categories. This information can be used by various new businesses to locate a prime spot for their new restaurant, office, etc. The major stakeholders for this project are small business owners and planning to start their business at in Vancouver. This project would help them find the optimal location based on the category of their business such as:

- What's a best spot to open up a restaurant in Vancouver?
- What type of restaurant is most likely to be successful in different parts of the city?

2.   DATA

Vancouver has many neighborhoods. The Wikipedia website has the list of all postal codes starting with V. Only postal codes in and around Vancouver are selected amongst this data

"https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_V"

	PostalCode	Borough	Neighborhood
0	V5A	Burnaby	Government Road, Lake City, SFU, Burnaby Mountain
1	V6A	Vancouver	Strathcona, Chinatown, Downtown Eastside
2	V5B	Burnaby	Parkcrest-Aubrey, Ardingley-Sprott
3	V6B	Vancouver	NE Downtown, Gastown, Harbour Centre, Internat...
4	V7B	Richmond	Sea Island, YVR

Unfortunately, the latitude and longitude values are missing for these postal codes in the website and geocode python module is very inaccurate. Therefore the latitude and longitude data has been procured from the google maps website manually and finally everything merged together.
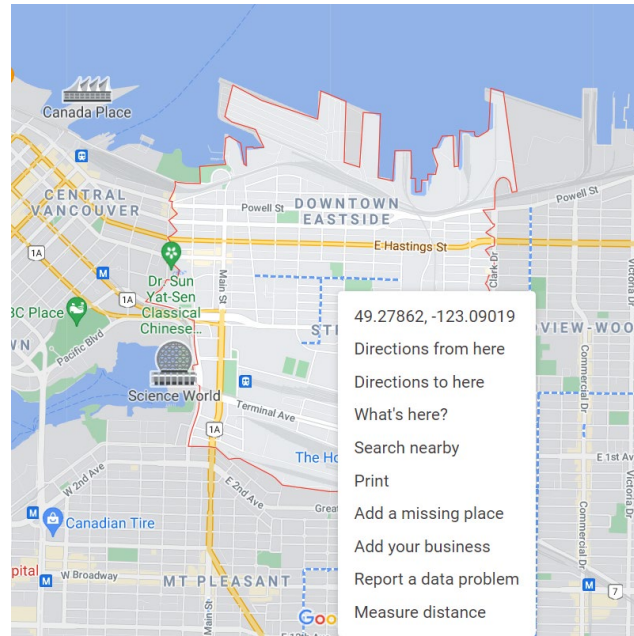
	PostalCode	Latitude	Longitude
0	V5A	49.266519	-122.936557
1	V6A	49.277722	-123.090575
2	V5B	49.271882	-122.976632
3	V6B	49.279990	-123.115413
4	V7B	49.185816	-123.172296

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	V5A	Burnaby	Government Road, Lake City, SFU, Burnaby Mountain	49.266519	-122.936557
1	V6A	Vancouver	Strathcona, Chinatown, Downtown Eastside	49.277722	-123.090575
2	V5B	Burnaby	Parkcrest-Aubrey, Ardingley-Sprott	49.271882	-122.976632
3	V6B	Vancouver	NE Downtown, Gastown, Harbour Centre, Internat...	49.279990	-123.115413
4	V7B	Richmond	Sea Island, YVR	49.185816	-123.172296

Next, the venues information data for each neighborhood is extracted from the foursquare data as follows using the latitude and longitude:

https://foursquare.com/

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Government Road, Lake City, SFU, Burnaby Mountain	49.266519	-122.936557	Burnaby Mountain Golf Course	49.264878	-122.942871	Golf Course
1	Government Road, Lake City, SFU, Burnaby Mountain	49.266519	-122.936557	Clubhouse at Burnaby Mountain	49.264949	-122.943104	Burger Joint
2	Government Road, Lake City, SFU, Burnaby Mountain	49.266519	-122.936557	Burnaby Mountain Driving Range	49.263959	-122.942353	Golf Driving Range
3	Strathcona, Chinatown, Downtown Eastside	49.277722	-123.090575	Union Market	49.277371	-123.086989	Deli / Bodega
4	Strathcona, Chinatown, Downtown Eastside	49.277722	-123.090575	MacLean Park	49.278809	-123.088546	Park