IBM Applied Data Science Capstone

Opening a New Mall in Vancouver, British Columbia

By: Andrew Shrestha

Date: February 6th, 2021

Introduction:

Vancouver has been ranked consistently as one of the best cities to live in the world. This is due to its beautiful and accessible nature environments with mountains, ski resorts, hiking trails, beaches, parks, and shopping experience. These factors tend to be the driving factor for tourism in Vancouver since its ranking is number 2 most popular destination in Canada with over 15 million tourists per year.

Vancouver is highly successful in the retail industry, as it hosts 4 of the top 10 malls (in terms of sales per sqft.) in Canada. These numbers are a great indicator of the general population, where the public/tourists are very inclined to visit and spend time in malls to meet their growing demand for greater shopping experience. Malls are successful at generating great revenue for both property developers/ owners, as well as the city itself, and so it would be of great interest to look into pursuing and expanding this avenue.

Business Problem:

The objective of this Capstone Project will be to utilize data and select the best locations in the city of Vancouver to open a new Shopping Mall. Through the use of techniques such as foursquare analysis, clustering, and segmenting, we would like to solve the following business problem: "Where would be the optimum location to construct a new mall for property developers/owners in the city of Vancouver?"

Data:

- A list of neighborhoods in Vancouver. These will be the locational data that we will confine our analysis to
- Both the Latitude and Longitude coordinates of each of the neighborhood data points so that we may plot it on a map
- Data on Shopping Malls or related so that we may use this in order to perform clustering on the neighborhoods

Data source:

- Obtain a list of 22 neighborhoods from Wikipedia (https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Vancouver)
- Web scraping techniques to grab the data from Wikipedia (python and Beautifulsoup)
- Geocoder to extract longitude and latitude of the neighborhoods
- Foursquare API to obtain shopping mall venue data of each neighborhood
- Folium for map visualizations