Business in Vancouver

Where to settle and start a business!!

Bett, Iyan

21 November 2020

Capstone Project: The Battle of Neighbourhoods (Week 2)

1 Introduction

1.1 Background Information

Almost everyone needs to move at some point in their lifetime. Be it moving neighbourhoods, cities or even continents. As you move to a new city there are several factors that may influence where you settle in and where you are likely to set up a business within the region you have moved to. Besides the price of the housing, access to facilities and proximity to important resources, one very important factor is safety. Even though different people put emphasis on different things, your safety mostly comes first. Therefore, most people moving to a new city have a common problem: How to find a neighbourhood to settle in that is safe and secure? In this report we will attempt to answer this question using Data Science.

1.2 Problem statement: Where to settle and start a business in Vancouver, Canada

In this report, we are trying to find the safest borough in Vancouver for Alex, where he would most likely would feel comfortable, safe and secure. Alex is 30 years old and has just moved to Vancouver. He has had good stories about living in Vancouver, but he recently came through an article that indicated increased burglaries and crime rates in Vancouver. Alex has been running a grocery store in Toronto where he used to stay before moving and intends to open a similar store in his new city.

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 in Vancouver City, Canada.

The first task would be to choose the safest borough by analysing crime data for opening a grocery store and short listing a neighbourhood, 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 analyse data and focus on the safest borough and explore its neighbourhoods and the 10 most common venues in each neighbourhood so that the best neighbourhood where grocery store is not amongst the most common venue can be selected.

1.3 Target audience for this report

This report is an analysis of the boroughs in Vancouver to establish the safest area to open commercial establishments in and specifically grocery stores. The information gathered from Foursquare in combination with data science methods form a good basis to derive data driven decisions regarding boroughs that best fit the specific needs at hand. It would even be possible for any newcomers to Vancouver to use some similar approaches to find the perfect home for their business in the City.

2 Data

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

- Finding the safest borough in Vancouver based on crime statistics
- Finding the most common venues
- Choosing the right neighbourhood within the borough

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

Following data sources will be needed to extract/generate the required information:

1. *Part 1*: 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 Neighbourhood in Vancouver along with type of crime, recorded year, month and hour.
2. *Part 2*: Gathering additional information of the list of officially categorized boroughs in Vancouver from Wikipedia.: Borough information will be used to map the existing data where each neighbourhood can be assigned with the right borough.
3. *Part 3*: Creating a new consolidated dataset of the Neighbourhoods, along with their boroughs, crime data and the respective Neighbourhood's co-ordinates.: This data will be fetched using Open Cage Geocoder to find the safest borough and explore the neighbourhood by plotting it on maps using Folium and perform exploratory data analysis.
4. *Part 4*: Creating a new consolidated dataset of the Neighbourhoods, boroughs, and the most common venues and the respective Neighbourhood along with co-ordinates.: This data will be fetched using Foursquare API to explore the neighbourhood venues and to apply machine learning algorithm to cluster the neighbourhoods and present the findings by plotting it on maps using Folium.

2.1 Methodology

Categorized the methodology section into two parts:

- **Exploratory Data Analysis**: Visualize the crime reports in different Vancouver boroughs to identity the safest borough and normalize the neighborhoods of that borough. We will Use the resulting data and find 10 most common venues in each neighborhood.
- **Modelling**: To help stakeholders choose the right neighborhood within a borough we will be clustering similar neighborhoods using K - means clustering which is a form of unsupervised machine learning algorithm that clusters data based on predefined cluster size. We will use K-Means clustering to address this problem so as to group data based on existing venues which will help in the decision-making process.

Results and Discussion

The objective of the business problem was to help stakeholders identify one of the safest boroughs in Vancouver, and an appropriate neighbourhood within the borough to set up a commercial establishment especially a Grocery store. This has been achieved by first making use of Vancouver crime data to identify a safe bough with considerable number of neighbourhoods for any business to be viable. After selecting the borough, it was imperative to choose the right neighbourhood where grocery shops were not among venues in a close proximity to each other. We achieved this by grouping the neighbourhoods into clusters to assist the stakeholders by providing them with relevant data about venues and safety of a given neighbourhood.

Conclusion

We have explored the crime data to understand different types of crimes in all neighbourhoods of Vancouver and later categorized them into different boroughs, this helped us group the neighbourhoods into boroughs and choose the safest borough first. Once we confirmed the borough the number of neighbourhoods for consideration also comes down, we further shortlist the neighbourhoods based on the common venues, to choose a neighbourhood which best suits the business problem.