Customer Base Segmentation for Opening new Gym Chain

1. Introduction

1.1 Background

The aim of this project is to help an entrepreneur who is planning to open Cross fit gyms in different areas of Canada. Due to this corona crisis people have become more focused on health issues and recent study has shown that there is a surge in new subscriptions of gyms. CrossFit is a lifestyle characterized by safe, effective exercise and sound nutrition. CrossFit can be used to accomplish any goal, from improved health to weight loss to better performance. The program works for everyone—people who are just starting out and people who have trained for years. So this project will highlight those places where number of gyms are less so that ROI would be more and Investment risk will be less.

1.2 Business Problem

The objective of this capstone project is to find the best locations to open gyms in Toronto. If there are already of gyms available, then people may be less likely to join the new gym because existing gyms already have their own customer base and initially people may reluctant to move to new gym.

1.3 Target Audience

The entrepreneur who wants to open the gym chain.

2. Data

2.1 Data Definition

Link: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

We will extract data from the above mentioned link to get the neighborhood data of Canada.

Foursquare Data Provider:

We will need data about different venues in different neighborhoods of that specific borough. In order to gain that information, we will use "Foursquare" locational information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API.

After finding the list of neighborhoods, we then connect to the Foursquare API to gather information about venues inside each and every neighborhood. For each neighborhood, we have chosen the radius to be 100 meter.

The data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

1. Neighborhood
2. Neighborhood Latitude
3. Neighborhood Longitude
4. Venue
5. Name of the venue e.g. the name of a store or restaurant
6. Venue Latitude
7. Venue Longitude
8. Venue Category