

Coursera Capstone Project

IBM Data Science Professional Certificate

Best location for opening a gym accessory retail outlet in Singapore

By: Yang Lin

Date: January 2021



A. Introduction and backgrounds

With the concepts of wellness and mindfulness becoming ever more important for consumers, fitness has become a more relevant goal for many people (Forbes, 2019). Nowadays, more people are going to health clubs and gyms in order to curb the side-effects that come with the hectic urban lifestyle. Holding a gym membership and spending money on fitness used to be seen as a luxury, but today it has become a part of people's lifestyle. According to Entrepreneur, "growing urbanization, rising middle class, and increasing disposable incomes are boosting the need for gyms (F45, 2020).

The high demand for gym and fitness also boom the retails business for the gym peripheral products. Both the gym operators and their consumers can easily access the gym accessory shop around the gym centers. Therefore, beside more and more gym and fitness centers opened island-wide in Singapore, the gym accessory retailers and potential investors also are more interested to know the numbers of gym centers located in a different neighborhood in Singapore. Then they can make a wise investment decision that opening their new retail outlets in a place with more gym centers.



1. Objective and Business Problems

The objective of this project is to help the retailers and potential investors to find the answers for the question: where are the best neighborhood locations to open new gym accessory retail outlets in Singapore? This study will be done by using the data science methodologies and machine learning techniques and concluded in a business solution and advice.

2. Target Audience of this project

This research project is particularly useful for gym accessory retailers who want to grow their business by opening more gym accessory retail outlets and the potential new investors in these industry sectors. As Singapore is quite a small country, the limited land space caused very crowd in the neighborhood areas. Therefore, high rental and labor cost is a big burden to run an offline

business. Although there are increasing numbers of gym and fitness centers over the recent years, choose a location with low potential customers surrounding may easily result in business failure.

B. Data

To find the solution for the problems, the following data will be required for this project:

- List of neighborhoods in Singapore. This defines the scope of this project that only within the Country of Singapore.
- Latitude and longitude coordinates of the neighborhoods. This is required in order to plot on the map as well as getting the venue data.
- Venue data, particularly data related to gym and fitness centers. The venue data will be used in the machine learning process to cluster the labels on neighborhoods.

1. Data sources and how it helps to solve the problem

A list of neighborhoods in Singapore can be obtained from the Wikipedia web page at https://en.wikipedia.org/wiki/Category:Places_in_Singapore. The data contains 191 neighborhoods in total over the island-wide in Singapore. Then we can retrieve the geographical coordinates for each neighborhood through the python package called Geocoder. With the latitude and longitude of each neighborhood, it can plot the neighborhood as circle markers on the map. It can be visualized where the neighborhoods are located.

Next, the most important data in this project is the venue data. It can be obtained by using the Foursquare API in python which can help to provide many categories of venue data by giving the neighborhood geographical coordinates. But we will only filter out the Gym and Fitness category from the venue data and then use it to train the k-means model for clustering. With the final clustering labels and map visualization, we are able to advise the best location for opening new gym accessory outlets.