Coursera Capstone Project…

Opening a Gym in Chennai City, India.

By: Sumedha

Introduction

A proverb I am listening from my childhood “Health is Wealth” that means even you earn billion but if your health is not good, then this money will be worthless. This era people are that much busy that they don’t have time to take care of there body. Gym is a place where people can get both fit body and mental peace by doing the exercise using the equipment available. This problem statement is all about opening a gym in Chennai.

Business Problem

The objective of this capstone project is to analyse and select the best locations in the city of Chennai, India to open a new shopping mall. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the business

question: In the city of Chennai, India, if a property developer is looking to open a new Gym, where would you recommend that they open it?

Target Audience

This project is particularly useful to property developers and investors looking to open or invest in new Gym in the capital city of Chennai, India.

Data

- List of neighborhoods in Chennai.
- Latitude and longitude coordinates of those neighbourhoods. This is required in order to plot the map and also to get the venue data.
- Venue data, particularly data related to Gym. We will use this data to perform clustering on the neighborhoods.

Sources

- Wikipedia web scrapping: https://en.wikipedia.org/wiki/Category:Suburbs_of_Chennai
  It contains the neighborhood and all the areas in Chennai.
- After that we will get the Latitude and Longitude of all the areas of Chennai using geocoder Python library.
- Using Foursquare API to get the venue data for the neighborhood.

Foursquare API

- Foursquare has one of the largest databases of 105+ million places and is used by over 125,000 developers.

- Foursquare API will provide many categories of the venue data, we are particularly interested in the Gym category in order to help us to solve the business problem put forward. This is a project that will make use of many data science skills, from web scraping (Wikipedia), working with API (Foursquare), data cleaning, data wrangling, to machine learning (K-means clustering) and map visualization (Folium). In the next section, we will present the Methodology section where we will discuss the steps taken in this project, the data analysis that we did and the machine learning technique that was used.