

Opening a New Gym in Pune,India

Omkar Nagarkar

April 30,2018



1.Introduction

A gymnasium, also known as a gym, is a covered location for athletics. The word is derived from the ancient Greek gymnasium. They are commonly found in athletic and fitness centers, and as activity and learning spaces in educational institutions. "Gym" is also slang for "fitness centre ", which is often an area for indoor recreation. A gym may be open air as well. Gym apparatus such as barbells, jumping board, running path, tennis-balls, cricket field, and fencing area are used as exercises. In safe weather, outdoor locations are the most conducive to health. Their curricula included self-defense, gymnastic a medical, or physical therapy to help the sick and injured, and for physical fitness and sports, from boxing to dancing to skipping rope. Today, gym are commonplace in the around. The number of gyms in the India has more than doubled since the late 1980s.Gym has been popular among the youth. Property developers are also taking advantage of this trend to build more gyms to cater to the demand. As a result, there are many gyms in the city of Pune and many more are being built. Opening gym allows property developers to earn consistent rental income. Of course, as with any business decision, opening a new gym requires serious consideration and is a lot more complicated than it seems. Particularly, the location of the gym is one of the most important decisions that will determine whether the mall will be a success or a failure.

2.Business Problem

The objective of this capstone project is to analyze and select the best locations in the city of Pune, India to open a new gym. 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 Pune, India, if a property developer is looking to open a new gym, where would you recommend that they open it?

3.Target Audience of this project

This project is particularly useful to property developers and investors looking to open or invest in new gym in the Pune,India.

4.Data

To solve the problem, we will need the following data:

- List of neighborhoods in Pune. This defines the scope of this project which is confined to the city of Pune, India.
- Latitude and longitude coordinates of those neighborhoods. This is required in order to plot the map and also to get the venue data.
- Venue data, particularly data related to Gyms. We will use this data to perform clustering on the neighborhoods.

Sources of data and methods to extract them:

This Wikipedia page (https://en.wikipedia.org/wiki/Template:Neighbourhoods_of_Pune) contains a list of neighborhoods in Pune, with a total of 42 neighborhoods. We will use web scraping techniques to extract the data from the Wikipedia page, with the help of Python requests and BeautifulSoup packages. Then we will get the geographical coordinates of the neighborhoods using Python Geocoder package which will give us the latitude and longitude coordinates of the neighborhoods.

After that, we will use Foursquare API to get the venue data for those neighborhoods. Foursquare has one of the largest database 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.