Coursera Capstone

IBM Applied Data Science Capstone

Opening a Fitness Centre, Nairobi Kenya.

Omaore Turuka

July 2020

Nairobi

Introduction

The fitness industry has exploded in recent years. Data from the International Health, Racquet & Sportsclub Association (IHRSA) shows there were "over 6B Visits to 39,570 Gyms in 2018" in the US. This was up almost 3% from the previous year. Membership to fitness clubs is on an upward trajectory, so are regular visits. This is replicated in most urban areas and global cities. Nairobi is no exception.

This global trend is driven by obesity awareness as well as trends toward improving health and fitness. This will have more significance post the current Covid-19 period as more people seek to become active, emerging from the sedentary lifestyle in the existing lockdown.

Small speciality gyms and independently-owned fitness businesses have majority stake in the Kenyan fitness market.

Choosing the right location is one of the key factors in ensuring the success of a business.

A good location needs to be easily accessible, visible to target market, competitively located and complies with the local zoning laws. There's need to determine if the locality has the target market readily available and what competitors may exist near that locality. A preference to a neighbourhood may also be a factor. Safety for the clients needs consideration as well as is access to public transport.

Business Problem

The objective of this capstone project is to carry out an analysis and select the best possible locations to set up a fitness centre. This will be guided by the neighbourhood analysis and competitor presence.

Using data science methodology and machine learning techniques, the project intends to answer the question: Where is the best possible location to open a sustainable fitness centre with the best chance of success?

Target Audience of this project

The information to be garnered from the data analysis carried out can be used by individuals or business owners keen on setting up a fitness centre. It will inform them on the low risk areas with a high chance of return on investment and that meets their target audience.

Data

The data required to solve the problem, includes the following:

- List of Suburbs in Nairobi, Kenya. This will be the scope of this project.

- Latitude and longitude coordinates of the suburbs to plot on a map and also to get the venue data.

- Venue data near the suburbs. This data will help in identifying competitors in a locality and access to nearby recreational facilities.

Data Sources and Extraction Methods

The web page (https://en.wikipedia.org/wiki/Category:Suburbs_of_Nairobi) contains a list of suburbs in Nairobi. There is a total of 58 suburbs, at the time of this writing.

To extract the data from the webpage, I shall employ the Python speciality packages (requests & beautifulsoup).

The Python Geocoder package will be used in getting the latitude and longitude coordinates of the suburbs.

The Foursquare API will get the venue data for the suburbs, utilising the location coordinates from above. From the Foursquare categories, we can list the most popular venues and try to classify suburbs. We can also search for fitness centres near the listed suburbs and identify gaps that may exist for a potential new fitness centre.

We will also make use of many data science skills such as data manipulation, data cleaning, data wrangling & machine learning (K-means clustering). To visualise the data, we shall use a map visualisation tool (Folium).

Reference:

International Health, Racquet & Sportsclub Association (IHRSA). (2019, March 28). Latest IHRSA Data.

https://www.ihrsa.org/about/media-center/press-releases/latest-ihrsa-data-over-6b-visits-to-39-570-gyms-in-2018/#