Coursera Applied Data Science Capstone

IBM Applied Data Science Specialization

Opening a New Gym in Dubai, United Arab Emirates

Introduction

For so many people, going to the gym is an amazing way to relax, build muscle, lose weight, and many other health benefits, but a lot of times it can be a bit annoying for people when there are no close gyms, and they have to travel good distance to reach there, hence it can be a good idea to start a gym/ Fitness center that is situated in a place that can cater to people in the surrounding area. But starting a gym is a big business decision, it's a lot more complicated than it seems, in particular, as mentioned the location of the gym is a very important decision that can determine if it will be a success or a failure

Business Problem

The objective of this project is to analyse and select the best locations in Dubai, United Arab Emirates, to open a new Gym. Using data science methodology and machine learning, this project will provide an answer to someone who wants to open a gym, the best recommended place to open.

Target Audience of this project

This project will be useful to property developers and investors looking to open a new gym in Dubai, United Arab Emirates. Which can be strategically placed that it can attract as much business as possible, and increase the property value.

Data

The following data will be used in this project:

• List of communities in Dubai.

• Latitude and longitude coordinates of those communities.

• Venue data, particularly data related to Gyms. Which will be used to perform clustering on the communities.

Data Sources and methods of extraction

The Wikipedia page ("https://en.wikipedia.org/wiki/Category:Communities_in_Dubai") contains a list of the communities in the city of Dubai, with 79 communities, with the help of beautifulsoup package, the data will be extracted, and then by using Python Geocoder package, will get the latitude and longitude coordinates of the communities.

After that, Foursquare API will be used to acquire the venue data for each of the communities, while it will provide a lot of categories, only the gym category will be the most useful, to solve the business problem.