IBM Data Science Capstone Project

Picking the right location for a new restaurant in Colombo, Sri Lanka

By Tobi Taiwo

January, 2020

Introduction

This project is about using data science tool set on a real-life problem and demonstrating the creation of value by applying the learned skills. It particularly addresses the business decision of “Where is the best location to set up a new restaurant in the city of Colombo, Sri Lanka that will yield high profitability”.

Business Problem

The main objective behind this project is to select the ideal locations in the city of Colombo in Sri Lanka to to open up a new restaurant. The project aims at using Data Science methodology and machine learning techniques to find out the best possible neighborhood for starting up a new restaurant that will lead to a high amount of sales and increased profitability compared to other neighborhoods in Colombo.

Target Audience

This project is beneficial to individuals or investors looking at starting up or investing money into building new restaurants in the city of Colombo in Sri Lanka. The project serves as a form of guideline when it comes to choosing the most viable location to open up a new restaurant that will potentially lead to high amounts of sales and increased profitability for the restaurant stakeholders.

Data

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

• List of neighbourhoods in Colombo, Sri Lanka. This defines the scope of this project which is confined to the city of Colombo, the capital city of the country of Sri Lanka.
• Geo-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 restaurants. This will be obtained from Foursquare through an API.

Sources of data and methods to extract them

This Wikipedia page (https://en.wikipedia.org/wiki/Category:Suburbs_of_Colombo) contains a total of 67 neighbourhoods in Colombo. Web scraping techniques will be used to extract the data, as well as Python requests and beautifulsoup packages. Python Geocoder package will be used to get the geographical coordinates of the neighbourhoods. Then, we will use Foursquare API to get the venue data for those neighbourhoods.