Coursera Capstone

IBM Applied Data Science Capstone

Opening a New Restaurant in Auckland, New Zealand

By: Hardik Modha
July 2020

Introduction

The population migration toward urban areas, tourism and busy lifestyle in the financial hub of New Zealand leads to surge in food serving business. The restaurants provide ease to singles, working parents, tourists and elderly people who have some inconvenience with cooking. This shows that this food industry is of bright future in terms of monetary return. As a result there are many restaurants in the Auckland suburbs and more will come. The reality is that many such restaurants fail in few years of their business due to poor planning. The prime factor of this failure is the location of the business. As this industry is highly competitive, the serious consideration should be given when deciding the location of business.

Problem Statement

The main objective of this project is to analyze the location of existing restaurants and to select the best location for new restaurant in the Auckland. By using the data science methodology and machine learning techniques, this project aims at providing solution to the following question:

In the Auckland, if an investor is looking to establish new restaurant, where would you recommend it?

New Zealand is the popular tourist destination in the world whose tourism industry and hospitality sector are one of the major contributors in the GDP. Also, it is a venue for the international events, and concerts, which boost these sectors. All such factors attract the investors and local entrepreneur toward hospitality sector and get benefit from it. So, this project is particularly helpful to the local entrepreneur who are novice in this industry and also to the financial investors in making investment decision for opening a new restaurant.

Data Collection

To solve this problem the following data will be needed:

- List of suburbs of Auckland, New Zealand. These suburbs are termed as "Neighborhoods".
- Latitude and Longitude of those neighborhoods. This coordinates will help in plotting the map and getting the venue data.
- Venue data for the restaurants. This data will be used to perform clusters on these neighborhoods.

Data Source and Extraction Method

The Wikipedia page (https://en.wikipedia.org/wiki/Category:Suburbs_of_Auckland) contains a list of neighborhoods in Auckland. This will be extracted by using the method of web scraping techniques with the help of Python 'Pandas' packages. After that we will get the geographical coordinates of the neighborhoods using Python 'Geocoder' package which will give us the latitude and longitude of the neighborhoods.

Once the coordinate are collected we will use Foursquare API to get the venue data for those neighborhoods. The reason for using Foursquare API is that, it 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 restaurant category in order to help us to solve the business problem put forward.

This project 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 for the data analysis that we did and the machine learning technique that was used.