Coursera Capstone Project

The Battle of Neighborhoods

Nisa Momin

August 2020

In this project we will focus on using the location data to explore interesting problems in real word! The project title is: **Location recommendation for *Chinese Restaurant* in *Singapore*.**

## **Table of content**

1. Introduction/Business Problem
2. Solution (Methodology)
3. Result /Discussion /Conclusion
4. Acknowledgement

## **1. Introduction**

## **1.1 Background:**

Mobile device proliferation, lightning paced lifestyles and challenging working hours are adding more and more layers of complexity to consumers’ lives. Singaporeans are feeling more stretched than ever before, and are increasingly striving for solutions to help simplify their busy lives. Consequently, 44% of Singaporeans have ever purchased restaurant deliveries or meal-kit delivery service online. Eating out is no longer reserved just for special occasions; it is fast becoming a way of life for consumers. A general increase in the frequency of eating out at restaurants or other out-of-home dining establishments was observed among Singaporeans.

Keeping in view the increasing demand for good dine in restaurants, an entrepreneur would like to set up a restaurant that can be in the best location to provide food to a greater number of people and at the same time can profit well.

## **1.2Problem Description:**

The objective of this project is to find a suitable location for an entrepreneur to open a restaurant in Singapore. By using the data science methods and machine learning algorithms such as clustering, this project aims to provide a solution to answer this business problems: In Singapore if an entrepreneur wants to open a restaurant which location should be considered?

## **1.3Traget Audience:**

The Target audience for this project would be the working Singaporeans who want to eliminate hassle of cooking from their busy lives and at the same time can enjoy good food. Also being a leisure activity, we would target people who love to dine in restaurants.

## **2.Data:**

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

* List of Neighborhoods in Singapore
* Latitude and longitude of these districts
* Venue data related to Chinese restaurants in different districts

## **2.1Extracting Data:**

* The data would be scrapped from the districts of Singapore via website: https://www.ura.gov.sg/realEstateIIWeb/resources/misc/list\_of\_postal\_districts.htm
* Since the Latitude, Longitude are not available in any web page so I had to manually enter the data
* Getting the venues related data from Foursquare API

## **3.Methodology**

### **How can we learn?**

Therefore, we will use **Kmeans Clustering** to learn and predict number of the *Chinese restaurant* in a district by providing the venues information of **Singapore**. Venues information is discovered from [Foursquare](https://foursquare.com/) API.

* explore and cluster the Postal Districts in Singapore

Then, I will the build model to predict the number of *Chinese restaurants* in the city-based country and compare with exiting number of *Chinese restaurants* and a potential place to start the business!

**4.Results and discussions**

From our analysis, the most common venues are cafés and restaurants so cluster 1 seems to be the hotspot for opening a Chinese Restaurant as it’s a main residential and commercial area and a potential business place. This cluster includes most of the districts in Singapore as a potential place to open a business giving more options to the owner to look for their own convenience and comfort.

Further analysis is required on the following potential aspects

* Parking availability
* Zoning Licenses and liquor licenses
* Data regarding existing properties for opening the restaurant

**5.Conclusions**

This notebook can be used a s starting point for potential restaurant owner to shortlist areas for opening a restaurant. Due to extensive analysis and availability of data, the notebook can serve the owner to also look for the alternative cuisines instead of only looking to open a Chinese reaturant.