Finding the Optimal Location for Opening a New Restaurant in South West London

IBM Data Science Capstone Project

Ken Young, July 2020

Introduction

Our client is a successful restaurant business with a long established presence in the West End and City in Central London.
They have built a reputation for excellent service in the casual dining market, offering quality innovative menus blending dishes from East and South East Asia.
Their recent addition of fixed price menus appealing to time-sensitive customers has proven popular.
This is valued, for example, by office workers on lunch breaks and diners heading to the cinema or theatre.
Our client is looking to expand by opening a new restaurant in South West London.
From the client's initial research, which involves limited detailed analysis, there are a number of locations that are potentially suitable.

Business Problem

In the restaurant business, location is one of the key factors in determining whether a new venture will be profitable.
The objective of this project is to identify, using data science techniques and the Foursquare location API, optimal locations for the proposed restaurant in South West London.
We will analyse and compare neighbourhoods with the following factors which we believe are critical in identifying a suitable location:-

- Existing restaurants particularly potential competitors in the vicinity.
- Spending power of local population.
- Proximity to transportation hubs.
- Proximity to entertainment venues such as cinemas.

Audience:-
The findings of this project will be presented to the client's senior management and we believe our analytical approach will identify suitable locations that will support the profitability and long term competitive advantage of the venture.

Data

To solve our business problem, we will use the following data:-

- **List of neighbourhoods in South West London.**
  This defines the scope of the project.
  As agreed with our client, this definition includes the London Boroughs of Merton, Richmond Upon Thames and Wandsworth.
  We will use the following Wikipedia page to extract a list of neighbourhoods in the desired Borough list.
  List of areas of London (https://en.wikipedia.org/wiki/List_of_areas_of_London)
- **Latitude and Longitude data for these neighbourhoods.**
  This will be obtained from the Python Geocoder package and is necessary to plot maps and obtain venue data.
- **Venue data, particularly cinemas and competitor restaurants.**
  We will use the Foursquare API and we will be using this data to perform clustering on neighbourhoods.
- **Locations of railway stations.**
  Locations close to railway stations, in particular tube stations, will generally have heavier footfall and potential customers.
  We will use the following Wikipedia page to obtain railway station locations.
  List of London railway stations (https://en.wikipedia.org/wiki/List_of_London_railway_stations)
- **Median household income for each neighbourhood.**
  As the restaurant is in the casual dining sector, information on potential disposable income is an important consideration.
  We will use the Excel spreadsheet obtained from the following website.
  Household income (https://data.london.gov.uk/dataset/household-income-estimates-small-areas)

The data from the above sources will require checking and cleaning, with machine learning and map visualisations being utilised to help solve our problem.
This is explored further in the following section.