Opening a new restaurant in London

Jazz Bhangu

3 July 2020

Introduction

The client for this project is a restaurant owner in West Hampstead, a busy borough in central London. Here, the client runs a successful Vegetarian restaurant. He now plans to open a second restaurant with the same offering and needs data driven insights on the best location. He is clear about the business strategy that he believes have led to his success:

- Busy, cosmopolitan area of London. The next location must be similar to West Hampstead in terms of overall amenity availability.
- Vegetarian food is an upcoming health trend but it is still not the norm in British society. Hence the owner believes people will only give this a try if there are limited other restaurant options in the vicinity. Once they try it, they love it – having discovered it tastes as good as meat….and hence they will keep coming back for more.
- Location:
    - Only interested in locations within very quick (500m max) walking distance of the main tube station.
    - Can open anywhere in London, but all else being equal, it would be ideal if the restaurant was on the same side of London as West Hampstead – for ease of managing both simultaneously.

Data acquisition and cleaning

There are distinct requirements for data provision and cleaning. Sources and methods have been identified to address these.

Geo-referencing

The primary geo-reference will be the postcode area eg. W1. This offers a reasonable segmentation of London; approx. 120 postcode areas. Alternative was to look at each Borough but there are only approx.. 30 so each covers too large an area for meaningful analysis.

- **London postcodes and associated borough names**: available from https://www.milesfaster.co.uk/london-postcodes-list.htm

- **Geo-spatial coordinates for centre of each postcode area in London**: available from https://www.freemaptools.com/download-uk-postcode-lat-lng.htm

London postcodes/borough names list and postcode geo-spatial coordinates were downloaded from the internet, into CSV files. These were then uploaded to Watson Studio which helpfully provides the Python code to read the data into the primary project Jupyter notebook

We do not have geo-locations of each tube-station. However, we can use the postcode area centre coordinates as an initial proxy, given that tube stations tend to be in the centre of the districts, from historical legacy. This can then be confirmed for the final choices using **Folium maps**.

Location data for local amenities

Location data is available using the 'explore' endpoint of Foursquare API. A standard sandbox account is sufficient for this. Client wants a max distance of 500m walking. Accordingly, a radius of 300m 'as the crow flies' will be used in the get queries.