**Building a Japanese restaurant in London**

Mary Mavrokapnidou

23rd August 2020

1. **Introduction**

**Background**

London is a city where restaurants and night life are very highly regarded as a means of entertainment and are very frequently visited by locals and tourists. With more and more nationalities joining the group of Londoners every year, international cuisine is at the heart of preferences for both English and foreigners, as people start having more refined taste and look for new recipes combined with a more exotic environment.

As a result, a lot of investors in HoReCa are looking for opportunities in opening restaurants based on cuisine from a foreign country. Londoners get more and more interested in sushi and Japanese culture in general, so the option to open a Japanese restaurant is considered a move which can prove to be very successful if location is chosen wisely. In this project, we assume that investors are interested in the growing area of South East London.

**Problem statement**

The purpose of the project is to provide advice to stakeholders as to which neighbour in South East London is the most appropriate to open a Japanese restaurant by answering the following questions:

1. What are the areas that do not have a Japanese restaurant, so that there is not direct competition?
2. What are the areas that do have not a lot of restaurants but have some, so that they offer some options for dining and attract customers?
3. What are the areas where tourists usually stay?
4. **Data**

In order to answer the questions in the problem statement, we need to identify

1. Areas where successful Japanese restaurants exist (based on postcode)
2. Areas where there are either one or two restaurants (no more, no less) within the top 10 visited places of the neighbourhood
3. Areas that have a hotel

The sources that will be used are:

1. **Wikipedia**, to identify London’s municipalities and the first digits of their postcode

Link: https://en.wikipedia.org/wiki/List\_of\_areas\_of\_London

1. The **library Geopy, function Nominatim**, that will provide the latitude and longitude of a central point in each municipality, based on the first digits of the postcode
2. The number of restaurants and their type and location in every neighbourhood, as well as information on other venues, will be obtained using **Foursquare API**
3. **Methodology**

The first stage is to download all London neighbourhood names and add them in a dataframe. The actual names of the columns do not correspond to the names displayed, so they are substituted with the right names for convenience in referencing and analysis. We then filter the neighbourhoods whose postcode start with SE as they comprise the South East London and create an address column that will be used as identifier for the latitude and longitude of the neighbourhoods’ central locations.

The second stage is to identify the coordinates of these central points using the library Geopy, function Nominatim and add them in a dataframe.

The third step is to remove all locations from the second dataframe where Latitude and Longitude is ‘None’. We identify the coordinates of central London and place them along with neighbourhoods’ coordinates on a map.

In the next phase, we identify the venues in each neighbourhood from Foursquare using the Foursquare credentials of our personal account. We download all venues within a distance of 500 metres from central locations and save them in a dataframe.

Then, we create a table with the venue number as row and venue type as column and set the table cells equal to 1 when there is a comment for the venue. We calculate how many comments on average there are per venue type in each neighbourhood and create another table with this information, putting this time the neighbourhoods as rows. Next, we split this table into two, one with neighbourhoods with a Japanese restaurant and the other without one.

These tables are then transformed into two ones, where the top 10 venue categories are identified per each neighbourhood, based on the number of comments per venue type. Looking closer at the table with the neighbourhoods having Japanese restaurants within their most popular options, their characteristics are:

1. There is a frequently visited hotel
2. There are one or two frequently visited restaurants (no more, no less)
3. There are pubs, bars and coffee places

In the next phase, after adding the latitude and longitude details in the corresponding top 10 venues table, we cluster the neighbourhoods without a Japanese restaurant using the kmeans method in 5 clusters.

1. **Results and discussion**

We will use the criterion 'Hotel existence' to define if a neighborhood is good to create a Japanese restaurant. Checking the most common venues, the neighborhoods with a hotel appear only in clusters 1 and 2 and these are:

* Bankside
* New Eitham
* Peckham
* Penge
* Tulse\_Hill

In these areas, the neighborohoods that have either one or two restaurants (not more, not less) are Peckham and Penge, so they are chosen as the most popular neighborhoods to build a Japanese restaurant and are recommended to the investors.

1. **Conclusion**

Using data from Wikipedia and Foursquare, we identified the South East London neighborhoods and the top 10 venue categories based on foursquare comments. This information was used to trace the neighbourhoods with Japanese restaurants where it wouldn’t be commercially sensible to build a new one due to direct competition reasons. By observing the popular venue types of these neighborhoods, we confirmed that a successful Japanese restaurant is usually places in relatively touristic regions (ie. where a hotel is located) and in places with medium presence of popular restaurants (either one or two). These characteristics were used to filter the neighborhoods without Japanese restaurants and leaded to our recommendation of Peckham and Penge as the best areas for investment.