

# Predicting the Businesses in London that offer Sustainable and Ethical Products and Services

**Phil Roussel**

June 2020

## 1. Introduction

### 1.1 Background

It has been more common for residents, workers and tourists in London to seek out supermarkets, cafes and restaurants that have sustainable and ethical offerings and practices. This has only become more prevalent on the turn of the decade with CoViD-19 highlighting the benefits of an environmentally friendly economy. This study aims to identify the density and clusters of venues with sustainable and ethical goods and services on offer. The study will leverage the Foursquare database of venues around London and users' comments. Using K-means clustering I clustered venues based on the venues that Foursquare users believe have a sustainable or ethical offering.

### 1.2 Problem

There is currently no easy way to identify Ethical or Sustainable venues in London, which for many is a topic of high priority. The study aims to help those who live, work and travel to London to identify the best areas in London to find sustainable and ethical businesses. This will be shown by showing the clusters of venues in London that have a sustainability attributed to them by Foursquare users.

### 1.3 Target Audience

The target audience in this study would include:

- Tourists - who will be interested in sustainable places to shops
- Residents - who will be interested on where to get ethically sourced food, drink and produce from in their local neighbourhood
- Local Authorities - who will be interested in potentially making London a go to place for sustainable and ethical businesses.

All the interested parties in this project would be able to encourage more sustainable practices by businesses in London and help market the city as an ethical destination for people looking to visit or move to the city.

## 2. Data

### 2.1 Data sources

- Foursquare Location Data - Restaurants, Cafes, Shops and Supermarkets in London
- Foursquare User Comments information - Comments from Foursquare users that highlight the keywords "sustainable" or "ethical"

## 2.2 Data cleaning

Data being pulled from the Foursquare API has its limitations in that there may be some missing venues that aren't included in the database or are too new as to contain tips or the appropriate data. This is more of an identification issue and is resolved in making the assumptions with the caveat that the database is not fully complete.

The data was also cleaned to remove variables that are not useful to this specific study of the data, so there were several values that were not categorised, these were removed from the study. There were also a number of columns that began with 'location' so to avoid confusion this term was dropped.

## 3. Methodology

### 3.1 Stage 1

Using the Foursquare database, I identified "London" and its coordinates as determined by the database using geocode. It then came to decide the search radius to operate in, for which I chose a 10-mile radius. This encompasses much of Zone 1 and Zone 2 in central London, the reason for choosing these is that these are the most common areas that people commute from and to. This also has the highest concentration of tourist spots.

The two string indicators used to identify my target data set was "sustainable" and "ethical", as these are widely recognised words for identifying companies that are having a positive environmental and social impact.

### 3.2 Stage 2

Using the Foursquare API script for pulling venue and tips data I pulled the data for the parameters set out above. The data was then normalised to put into a dataframe.

The 'name' and 'categories' columns were then filtered in order to ensure any data that was not categorised was removed from the study.

### 3.3 Stage 3

The dataframe was then converted onto a map of London using folium tool. The venues were identified on the map which allows for a visual understanding of the areas of London that are most focused on sustainability and ethical businesses.

## 4. Results

The results in Figure 4.1 show that there are 3 distinct clusters. Those clusters are in Mayfair, Shoreditch and Southwark. There were 2 outliers in Hammersmith and Old Ford, that would not be considered due to their proximity to other datapoints.

Figure 4.1



## 5. Discussion and Recommendations

### 5.1 Discussion

There is a bias inherent in studies like this from those carrying it out. This was not different in this situation. My expectations for this study was to see a high density of sustainable and ethical businesses in East London, which is certainly the case. What was less expected however was the clusters in Mayfair and Southwark.

Using my own inherent knowledge of tourist areas in London the concentration of venues in Mayfair are likely to be most relevant for tourists as this also has the closest proximity to destinations such as Buckingham Palace and Oxford Street.

Southwark is less known for tourists but certainly has a higher population density in terms of residents. It could then be inferred that residents living in this area are more concerned with shopping for sustainable or ethical produce.

The same can be said for East London and the higher population density there, but the added value here is the proximity to the City of London where a high percentage of the population commute to during the week.

### 5.2 Recommendations

I split my target audience into 3 distinct categories. The first was Tourists, who I would develop recommendations on the specific area to visit being near Mayfair and by the other tourist sites. But that there is an opportunity to grow the concentration of tourists in this spot and highlight the ethical venues in other areas of London that could be used.

For the Residents of London, or those looking to move to London and have a sustainable or ethical consideration as part of their requirement. I would suggest that the best areas to research are Shoreditch and Southwark. When doing their regular shopping or eating out at local restaurants these areas will offer the most opportunities. The reason for not mentioning Mayfair is the lower concentration of residential housing. As well as the higher cost of houses.

Finally, I would offer my recommendation to Local authorities to focus sustainable and ethical initiatives in areas such as Mayfair, Southwark or Shoreditch because there is likely to be a higher

adoption rate for any new schemes. It also allows a platform to build businesses that can influence the practices of other firms to make London being the place of choice if people are looking to move of travel.

## 6. Conclusion

This study analyses the clustering of venues in Central London based upon the key indicators of being “sustainable” or “ethical”. Using the Foursquare API I identified three key areas in London and gave my recommendations as to what area will be most appropriate for tourists, residents, and local authorities. This allows these three key groups to be able to identify and grow the business in these areas that will ultimately grow the requirement and drive demand for sustainable and ethical products and services.

## 7. Future Improvements

Any future studies would benefit from a wider variety of data to also assess other information such as house prices, the most popular tourist venues. This could provide my accurate recommendations of where specifically could be the most appropriate for tourists and residents to go.