The Battle Of The Neighbourhoods

Opening a restaurant in London

Haoyun Zhang

August 2020

1.Introduction

- **Background:**

Due to the pandemic, London restaurants faced many challenges. Now in the late stage, people's willingness to dine out is higher than ever. With incumbent restaurants struggling to recover, relaxation of regulations and drinkers and dinner returning to the street, some may consider now an opportunity to open a restaurant in London. As a former student in London, it is also a place of my interest to look into.

- **Business problem:**

This project is relevant to those who are interested in opening a restaurant in London, especially, in looking for an **ideal location** of the new restaurant. Location is one of the crucial determinants of a restaurant's success, therefore it is important to make an informed decision on choosing one.

In this project:

- We seek to find a borough with the highest level of activity based on the density of the restaurants and the total turnover of restaurants in the borough.
- We focus on one cluster of the restaurants in the selected borough; identify locations with low level of competition but are also close to the centre of the cluster.

2.Data

Data sources:

- Table 'London population by borough', 'London postal codes', 'London areas', JSON file of London boroughs. We extract information of postal codes and population of London by borough from websites. Then transform them into dataframe and visualise using folium. We draw boundaries of each borough and choropleth map of the population by borough to identify them visually.

  - http://www.internetpk.com/postal-codes/united-kingdom/london.php - **'London postal codes(currently offline)'**
  - https://en.wikipedia.org/wiki/List_of_areas_of_London - **'List of London areas'**

- **Foursquare API**: we will first get geographical data for the list of postal codes using geopy, which provides good coverage of the city. Once we have the co-ordinates, will get nearby venues and restaurants using Foursquare API. Extract restaurant-related data and transform them into dataframes. Then use folium to plot and generate heat maps to extract insight into the distribution of restaurants in London and in each borough.

- Dataset 'Restaurants in London 1st August 2019' by the office for National Statistics. The data set contains a total turnover of restaurants in each borough. We will visualise the numbers using histogram. We will only be using the total turnover for the purpose of this project.

  - https://www.ons.gov.uk/businessindustryandtrade/business/activitysizeandlocation/adhocs/010295restaurantsinlondon - **'Restaurants in London'**

We will then use these data to:

- Cluster restaurants using **DBSCAN** with **ball-tree algorithm** and **haversine metric** for geographical data.
- Transform latitudes and longitudes into Xs and Ys using **pyproj**, in order to calculate distance between locations.
- Create equally-spaced candidate locations within a radius of **~2km** of the cluster centre. Ideal locations will have **1** to **5** restaurants within a radius of **250m**.

3.Methodology

3.1 London overview - Selecting a borough:

Numeber of restaurants by borough

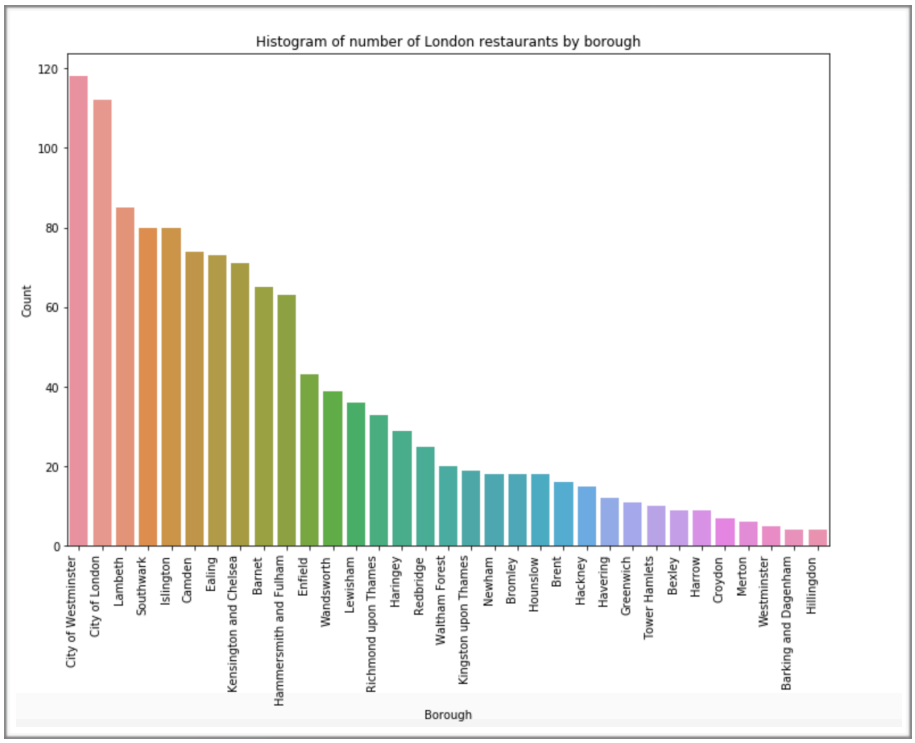

Figure 1

Figure 1 is a histogram of the number of restaurants in each borough.
With a radius of 500 meters and limitation of items of 100, Foursquare returned a dataset containing a total number of 1227 restaurants with a Maximum of 123 restaurants in City of Westminster. *(Although it is not full coverage of London restaurants, in an equally constrained search, City of Westminster and City of London clearly out numbered other boroughs.)*

Visualisation of restaurant locations

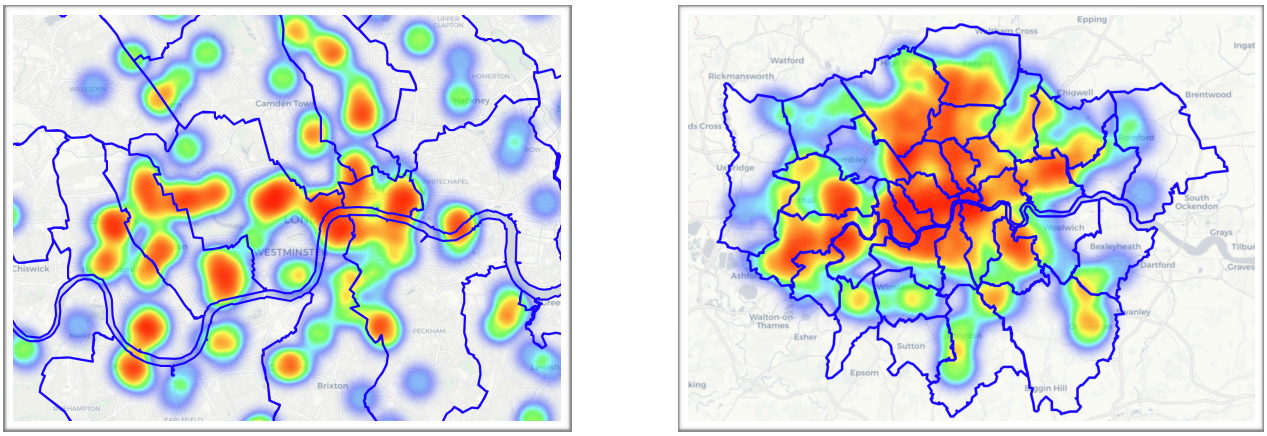

Figure 2

Figure 2 contains two heat maps, one that zooms at the City of Westminster and the other at London. We see that the ‘hottest’ spots are in **City of Westminster**, **City of London** and **Kensington and Chelsea**. And the periphery of London has a ‘low temperature’.

Total turnover of restaurants by borough(Aug 2019)

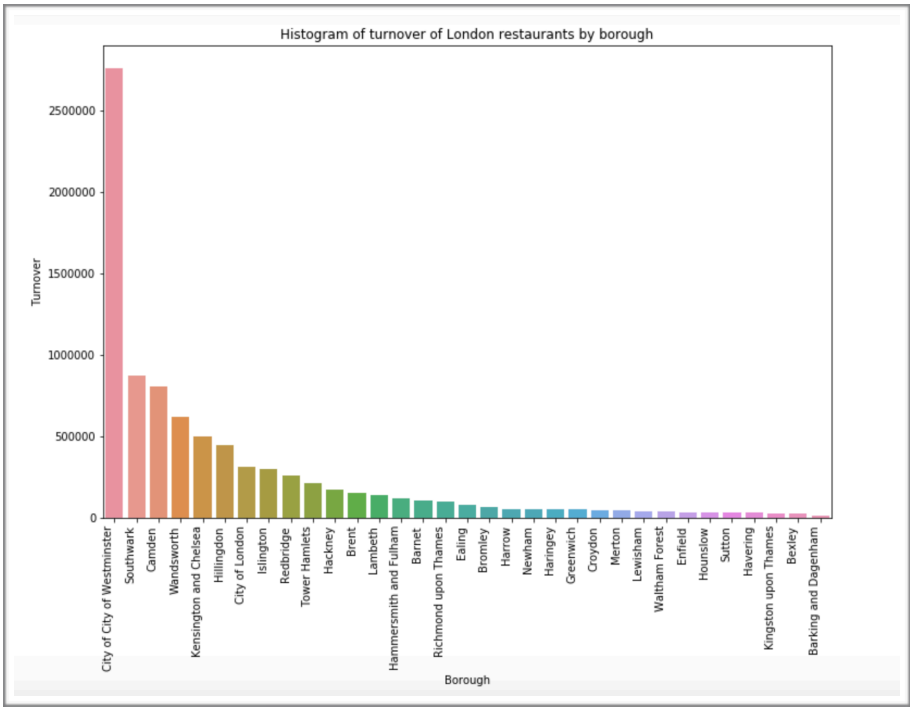

Figure 3

Figure 3 exhibits total turnover of restaurants of each borough. According to the graph, turnover of City of West minster more than doubled that of the second highest borough, Southwark, reaching 2.8 million pounds in Aug 2019. **With its remarkably high turnover, City of Westminster will be our candidate for analysis.**

3.2 City of Westminster

- **Visualisation and exploratory analysis**

To get a better picture of restaurants in City of Westminster, we allocate all available calls to the search of restaurants in the borough. With a radius of 500 meters, limit of 500 venues, we search according to the '**List of London areas**'. Foursquare returned 505 restaurants within City of Westminster.

Figure 4 & 5 below are visualization of restaurant locations and scatter plot of restaurant locations. In **Figure 4** we see two major clusters of restaurants at north west(Paddington, Bayswater) and south east(Covent Garden, Mayfair, Knightsbride) of the borough respectively. (*Although some restaurants are not in the boundary, they are of close neighborhood and worthwhile of consideration. City of London and Lambeth are also competitive candidate boroughs.)* **Figure 5**, the scatter plot of all latitudes and longitudes further clarifies the cluster.

Visualisation of restaurant locations in City of Westminster

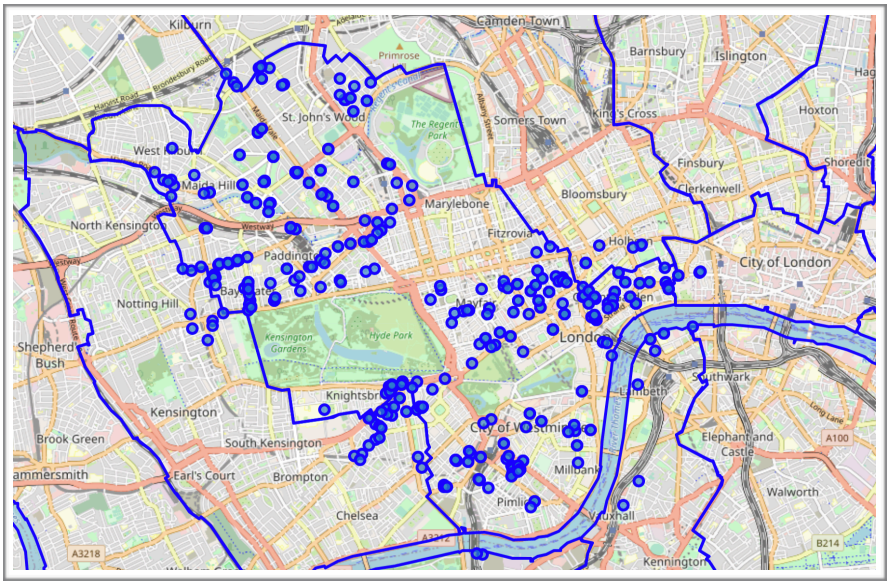

Figure 4

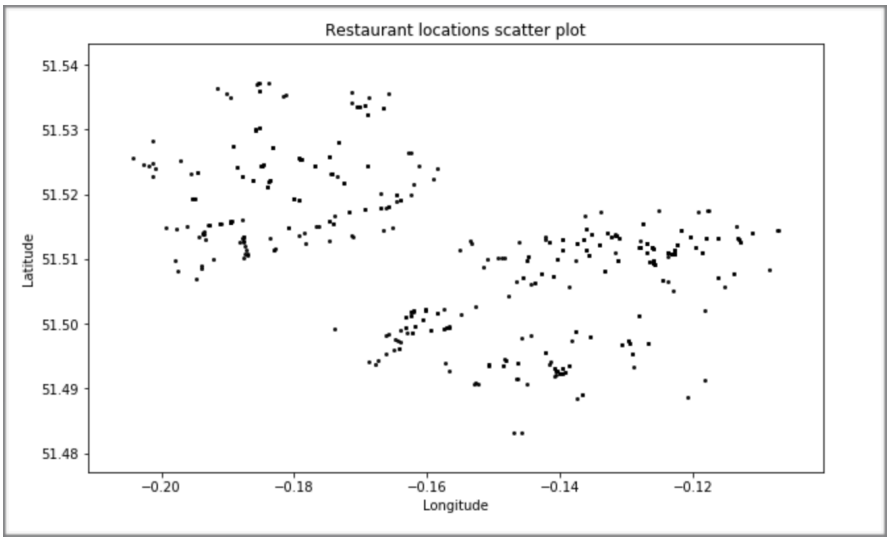

Figure 5

- **Clustering, Centroids - DBSCAN**

To find the centroids of those two major clusters, DBSCAN was used to cluster the geographical data. It clusters a spatial data set based on two parameters: a physical distance from each point, and a minimum cluster size.

- Parameters: *Epsilon* of 0.78km, *minimum sample* of 1(this leaves out noises)
- Haversine metric, ball tree algorithm were used to calculate great circle distances between points. (It may seem overkill, but k-means with 2 number of clusters would return the same result and we are only looking for a centroid.)
- With each cluster, we visualise a coverage of radius of 1.8km

DBSCAN returned 3 clusters. **Cluster 1** cover the south east cluster, **Cluster 2** south east cluster and Cluster 3 is formed by two restaurants outside the borough.

Visualisation of DBSCAN result

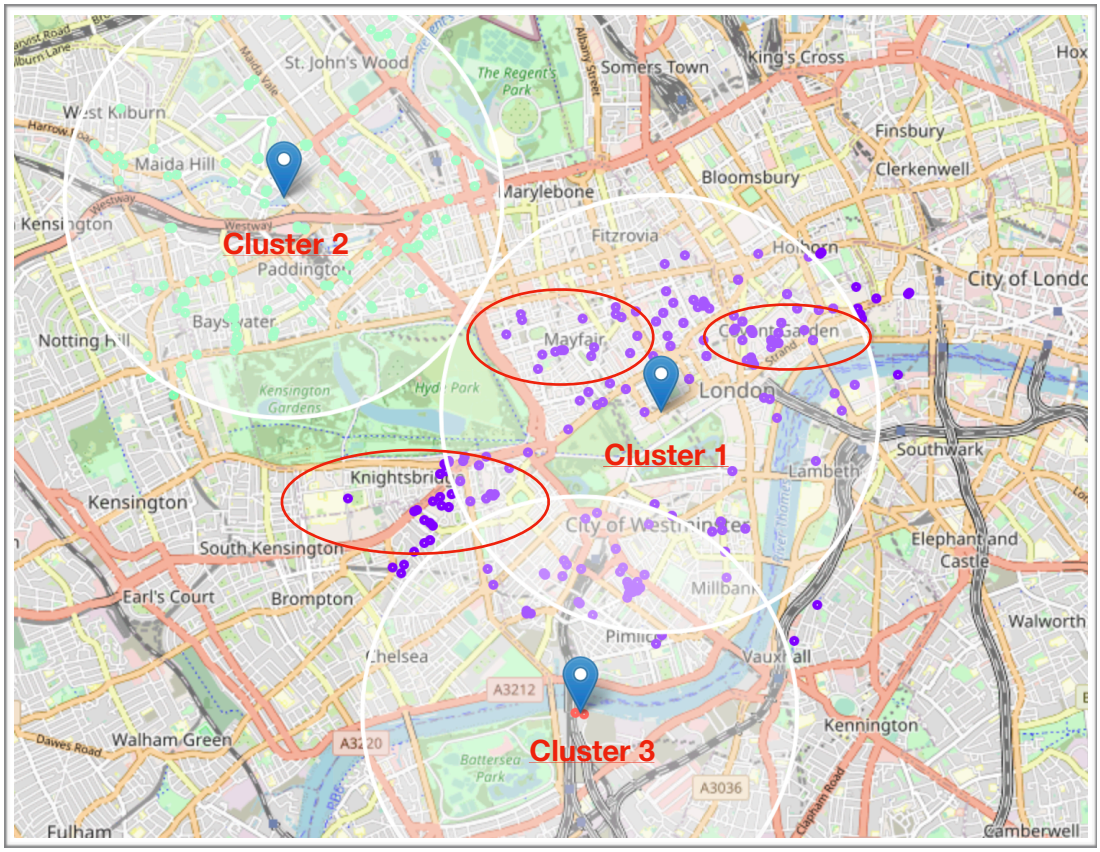

Figure 6

In **Figure 6**, **Cluster 1** covers areas such as *Knightsbridge, Mayfair, Covent Garden, Soho, and Oxford circus* which are centers of fine dining, tourist attractions with dense foot traffic and popular hangouts. **This is undoubtedly the golden cluster for restaurants or any other retailers.**

- **Finding and clustering ideal locations**

To find ideal locations, we first generate equally-spaced(50 meters apart) locations that covers entire **Cluster 1** and select only locations that has 1 to 5 restaurants nearby within a radius of 250 meters.

Candidate loctaions

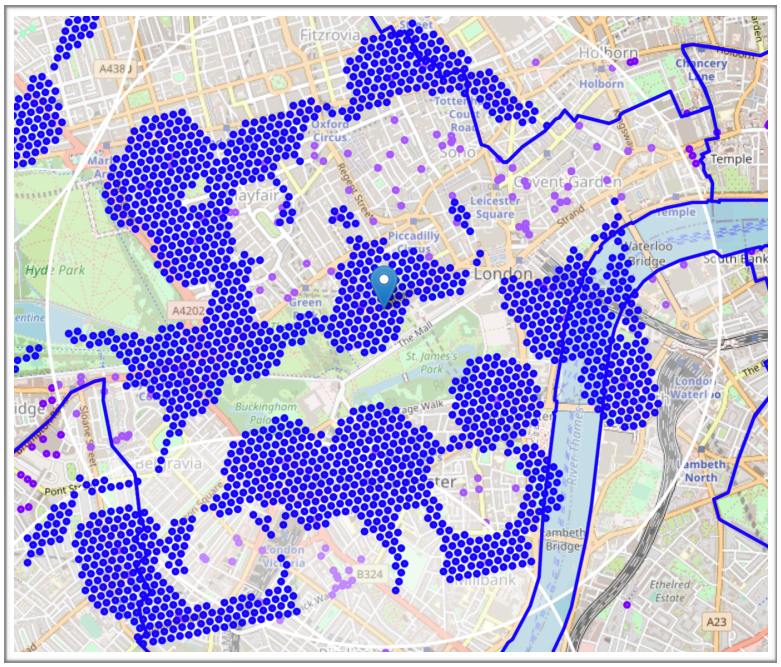

Figure 7

Blue dots in **Figure 7** above are 1970 locations that satisfy previous requirements. They capture the periphery of sub clusters(Covent Garden, Mayfair) within Cluster 1 and covers area with low density of restaurants as well.

Figure 8 below shows results of K-means clustering. With a specified number of 20 clusters, we reduce the set to a manageable size. The blue markers represents centroid of each cluster; white circles has a radius of 500 meters which covers all candidate locations and the red dots are existing restaurants. We see that the areas covered by the white circles have relatively lower number of red dots. Therefore, there are still vacancy in this extremely competitive center of culinary attractions.

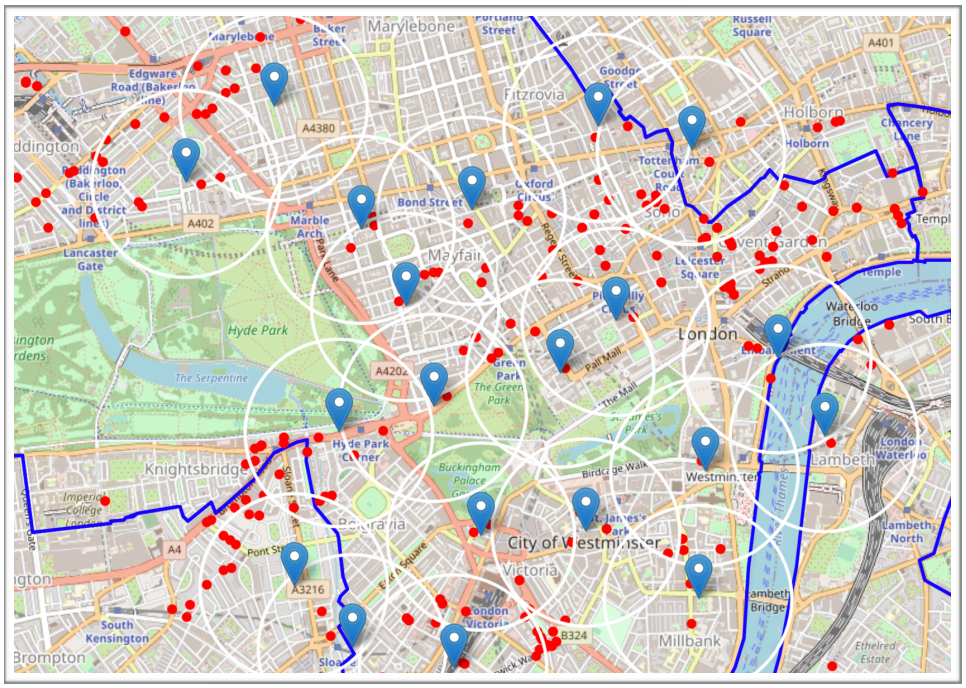

Figure 8

4.Results and Discussion

Table of all ideal locations clusters centroids

Address	No. Res. within 250m	No. Res. within 500m
Chutney Mary, Little St. James's Street, St. James's, Victoria	1	10
Lees Court, 22, Lees Place, St. James's, Mayfair	3	8
Sri Suwoon Thai Restaurant, 44, Hugh Street, St. James's, Victoria	3	22
Joseph Bazalgette, Victoria Embankment, Covent Garden	4	21
York House, 10-12, Berners Street, Cavendish Square & Oxford Market	2	14
62 Buckingham Gate, Spenser Street, Westminster, Victoria	5	11
2, Hyde Park Square, Paddington	2	16
Little George Street, Westminster, Millbank	2	10
Parkside, Knightsbridge, 28-52, Knightsbridge, Belgravia	5	27
16-22, Bourne Street, Belgravia	3	11
Bryanston Square, Marylebone	3	15
20a, Brook Street, St. James's, Mayfair	1	17
County Hall, Belvedere Road, South Bank, Lambeth	1	3
A4, St. James's, Mayfair	1	9
Victoria Square, St. James's, Victoria	2	12
Buchanan House, 3, St James's Square, St. James's, Covent Garden	3	20
St John's Building, 79, Marsham Street, Westminster, Millbank	3	11
Pavilion Road, Royal Borough of Kensington and Chelsea	1	16
Thomas Goode China and Glass, 19, South Audley Street, St. James's, Mayfair	4	14
St Giles, Bloomsbury, London Borough of Camden	3	29
	Mean = 2.6	**Mean = 14.8**

- Above is the table of all ideal locations clusters centroids. Number of restaurants within 250 meters is set between 1 to 5 and has a mean of **2.6**. Number of restaurants within 500 meters has a mean of **14.8**. For a radius of 500 meters, *St Giles, Bloomsbury,* highlighted in red, has the highest number 29 restaurants nearby and *County halls,* highlighted in green, has only 3 restaurants nearby.

  - *County hall* is a residential building next to Waterloo train station and London eye, both of which contribute significantly to foot traffic, thereby bringing possible high volumes of customers. It is a promising area for a new restaurant.
  - *St Giles, Bloomsbury* is next to Tottenham court road tube station and minutes away from Covent Garden and Soho. Surrounded by tube station and center of popular hangouts, the area is no less appealing than *County hall*.

- It is crucial to note that opening a restaurant is ultimately a comprehensive task with many dimensions that need to be considered, location is only one piece of the puzzle. Depending on the actual financial budget, cuisine style, targeting audience, stakeholders may adjust their choices among all suggested centroids. This project only aims to provide information in the aspect of location and practice implementation of studied codes.

5.Conclusion

According to the overview, with its high total turnover and diverse culinary culture, we found City of Westminster to be a promising borough to open a restaurant in. All existing restaurants in the borough were clustered using DBSCAN and Cluster 1 which covers more culinary attractions was focused on. Ideal locations were selected within a radius of 1.8 km from the centroid of Cluster 1 and were clustered using K-means with a number of 20 clusters.

It is demonstrated that, despite fierce competition, there are promising areas in the periphery of sub clusters of existing restaurants and next to transport junctions, tourist attractions such as Waterloo train station and London eye. All analysis resulted in 20 centroids scattered mostly inside Cluster 1. A final decision may requires more information on other aspects of the restaurant due to the nature of the task of opening a restaurant.