Finding an optimal location for a new Japanese restaurant in London

Fabrizio Leo

6th June 2020

1. Introduction

London (UK) is a modern and vibrant city full of places and social venues. According to the Office for National Statistics (2015) there are already about 18110 restaurants in the city, including take away food shops and mobile food stands. Given this premise, one might wonder whether there are the conditions based on density to open a new Japanese restaurant in this city.

Certainly, a candidate area to open a new Japanese restaurant should satisfy at least the following criteria as defined following a meeting with stakeholders:

- not too crowded in terms of number of restaurants
- with no or very few Japanese restaurants
- as close as possible to the city center of London

More in detail, the stakeholders would like to open a restaurant in an area not farther than about 8 kilometers from the city center but the closer it is the better. Other conditions that were defined were the absence of other Japanese restaurants in a radius of 350 m and with very few restaurants of different type in the same radius. Data analysis and machine learning will be applied to find out whether these conditions exists. If they would, the potential neighborhoods will be discussed in terms of their characteristics to allow the stakeholders to take the best decision.

2. Data

In order to get the necessary data, I defined a square grid of coordinate locations with 15000 meters side and 500 meters of spacing between locations centered around the coordinates of the city center of London. The coordinates of the city center of London (latitude: 51.5073219; longitude: -0.1276474) were loaded with the module **GeoPy.** Whenever needed, I used the module **pyproj** to convert geographical coordinates to UTC (meters), and *vice versa*. Figure 1 shows the original grid superimposed on a map of London using **Folium** module. Each neighborhood was defined as the circular area of 350 meters radius centered in one of the center coordinates defined above. This

creates some overlapping of neighborhoods (see Figure 1). I preferred this solution to the alternative of neglecting entire streets because of non-overlapping areas.

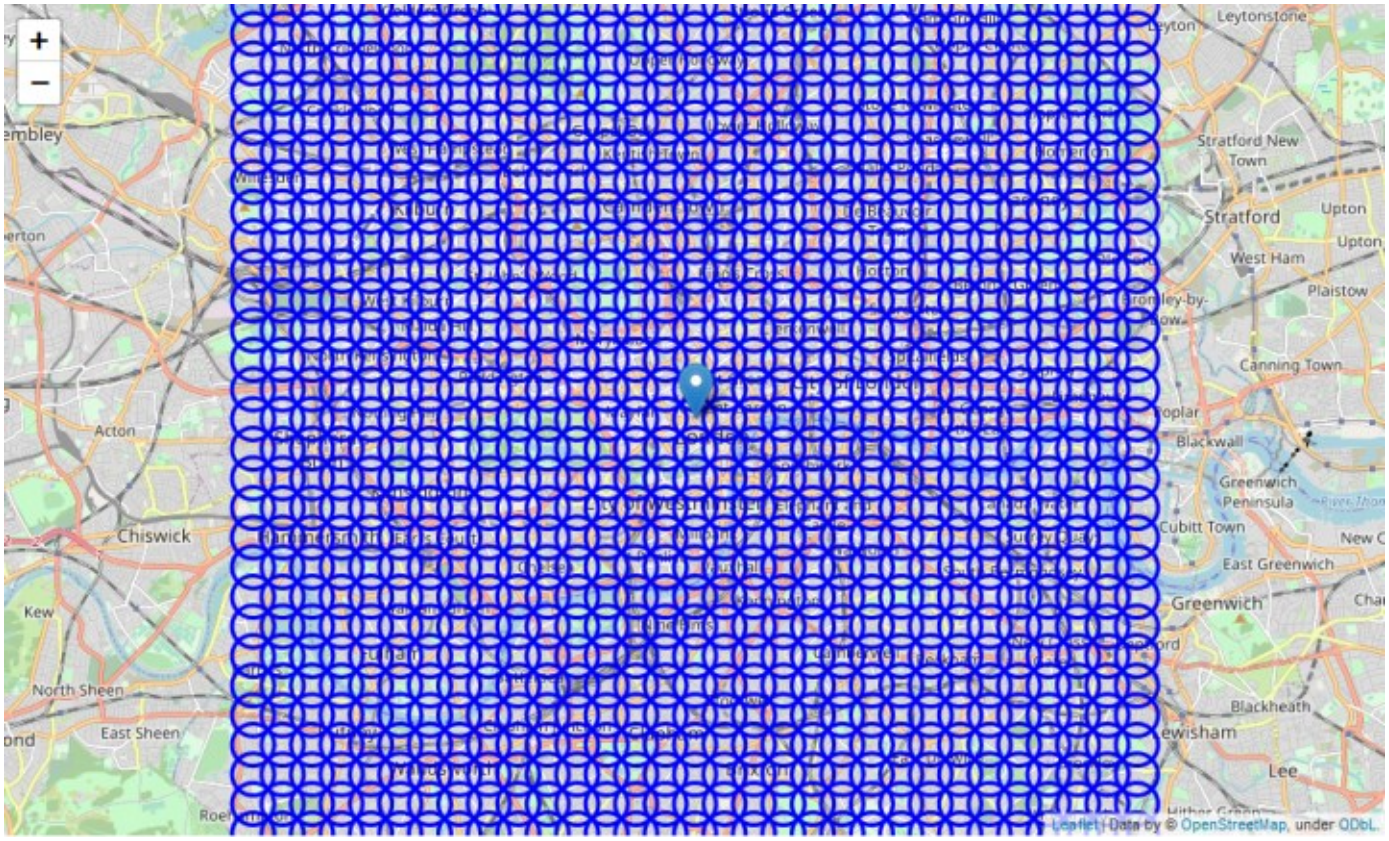

Figure 1. Grid of locations. The centered marker shows the location of the city center of London. The center of each circle represents the location of each neighborhood.

I then filtered the neighborhood retaining only those within 8 kilometers from the city center. To do so, I simply computed the Euclidean distance between each neighborhood and the city center of London and filtered out distances > 8000 meters. Figure 2 shows the resulting grid.

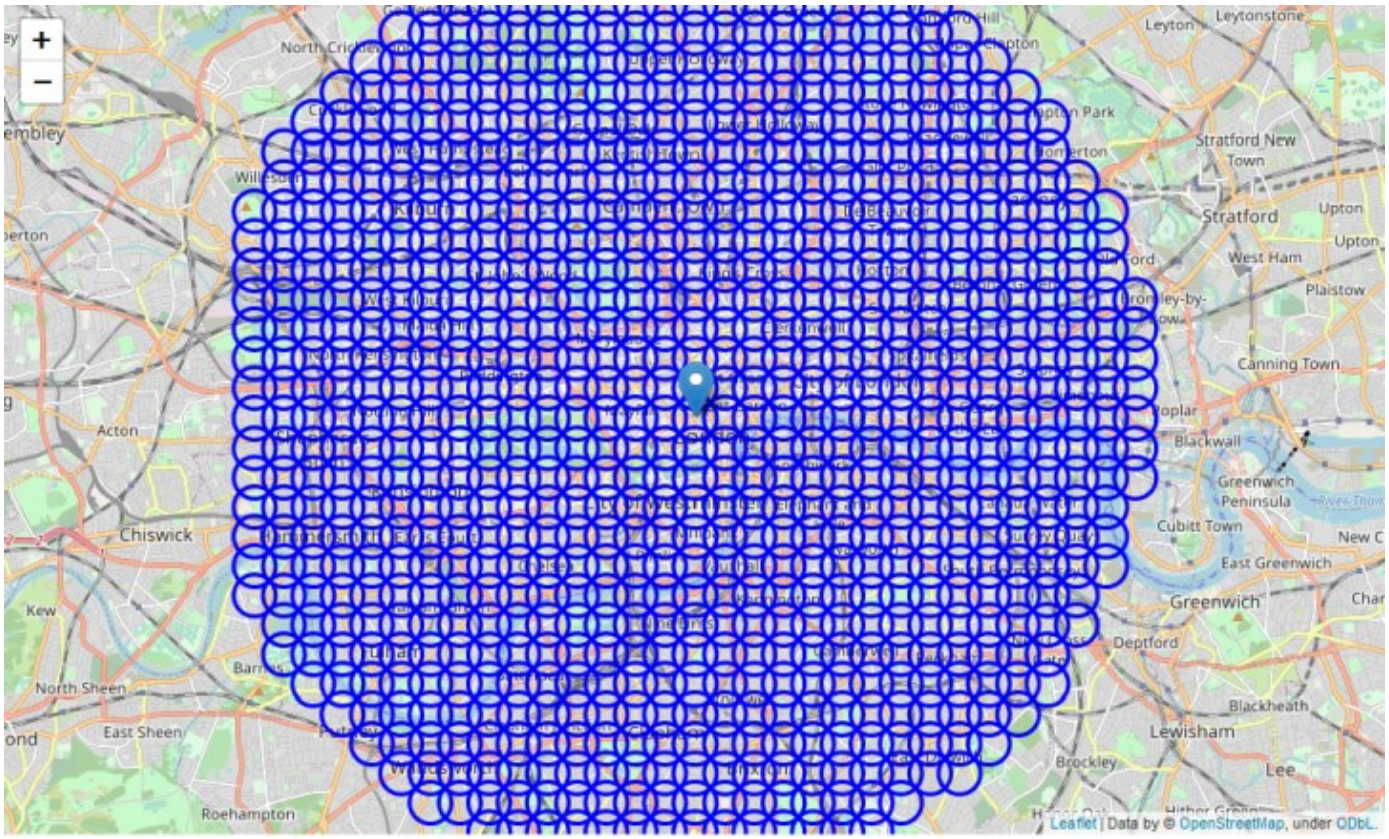

Figure 2. Grid of locations within 8 km from the city center.

Then, I retrieved the addresses of the center of each neighborhood using reverse geocoding with **GeoPy** (see Table 1).

	Address	Latitude	Longitude	Latitude_m	Longitude_m	Distance_to_center
0	Sarsfeld Road, Tooting Graveney, Upper Tooting...	51.442558	-0.160618	5.702655e+06	697327.162989	7959.650398
1	5, Balham High Road, Balham, London Borough of...	51.442558	-0.153432	5.702655e+06	697827.162989	7868.345241
2	65, Fernlea Road, London Borough of Lambeth, L...	51.442558	-0.146247	5.702655e+06	698327.162989	7808.052203
3	20, Glenfield Road, Abbeville Village, London ...	51.442558	-0.139063	5.702655e+06	698827.162989	7779.492373
4	Kirkstall Road, Streatham Hill, London Borough...	51.442558	-0.131878	5.702655e+06	699327.162989	7783.015094

Table 1. Address, geographical coordinates and distance to center of each neighborhood.

After that, I started to retrieve data about restaurants with **Foursquare API**. To retrieve the data, I used the keyword ‘restaurant’ since we were not interested in venues such as take away food shops and mobile food stands. In particular, I looked for restaurants in each neighborhood as defined before (i.e. 350 meters radius, see Fig. 2). For each venue, I also checked whether it was a Japanese

restaurant or not. Importantly, I also removed duplicates due to the fact I retrieved data based on partially overlapping areas. Table 2 shows a sample of retrieved and cleaned data.

	Name_restaurant	Categories	Latitude_restaurant	Longitude_restaurant	is_japanese
0	Curveball Restaurant & Bar	[{'id': '4bf58dd8d48988d1c4941735', 'name': 'R...	51.444472	-0.151164	False
1	The Exhibit	[{'id': '4bf58dd8d48988d116941735', 'name': 'B...	51.443153	-0.151080	False
2	Asafo Restaurant	[{'id': '4bf58dd8d48988d1c8941735', 'name': 'A...	51.445277	-0.124760	False
3	Troy Barbecue	[{'id': '4f04af1f2fb6e1c99f3db0bb', 'name': 'T...	51.442128	-0.124948	False
4	Bar 61	[{'id': '4bf58dd8d48988d1db931735', 'name': 'T...	51.440090	-0.125644	False

Table 2. Restaurant info including whether Japanese or not.

3. Methodology

In the following part of this report, I will be performing some basic statistics measuring density of restaurants and, particularly, Japanese restaurants. I will start performing these stats considering all the neighborhoods (i.e. within 8 km from the city center). Then, I will investigate how these statistics change as a function of the distance from the city center.

Then, I will perform a cluster analysis of the neighborhoods using **k-means**. The input of this analysis will be the distance from the city center, the number of restaurants and the number of Japanese restaurants. The purpose of this analysis will be to quickly identify clusters of neighborhoods which minimize the distance from the city center, the number of restaurants and the number of Japanese restaurants. Finally, I will select a sample of neighborhoods from the best clusters as a final sampling of candidate locations.

4. Results and Discussion

4.1 Restaurant distribution across London

Figure 3 shows the distribution of restaurants and Japanese restaurants in London.

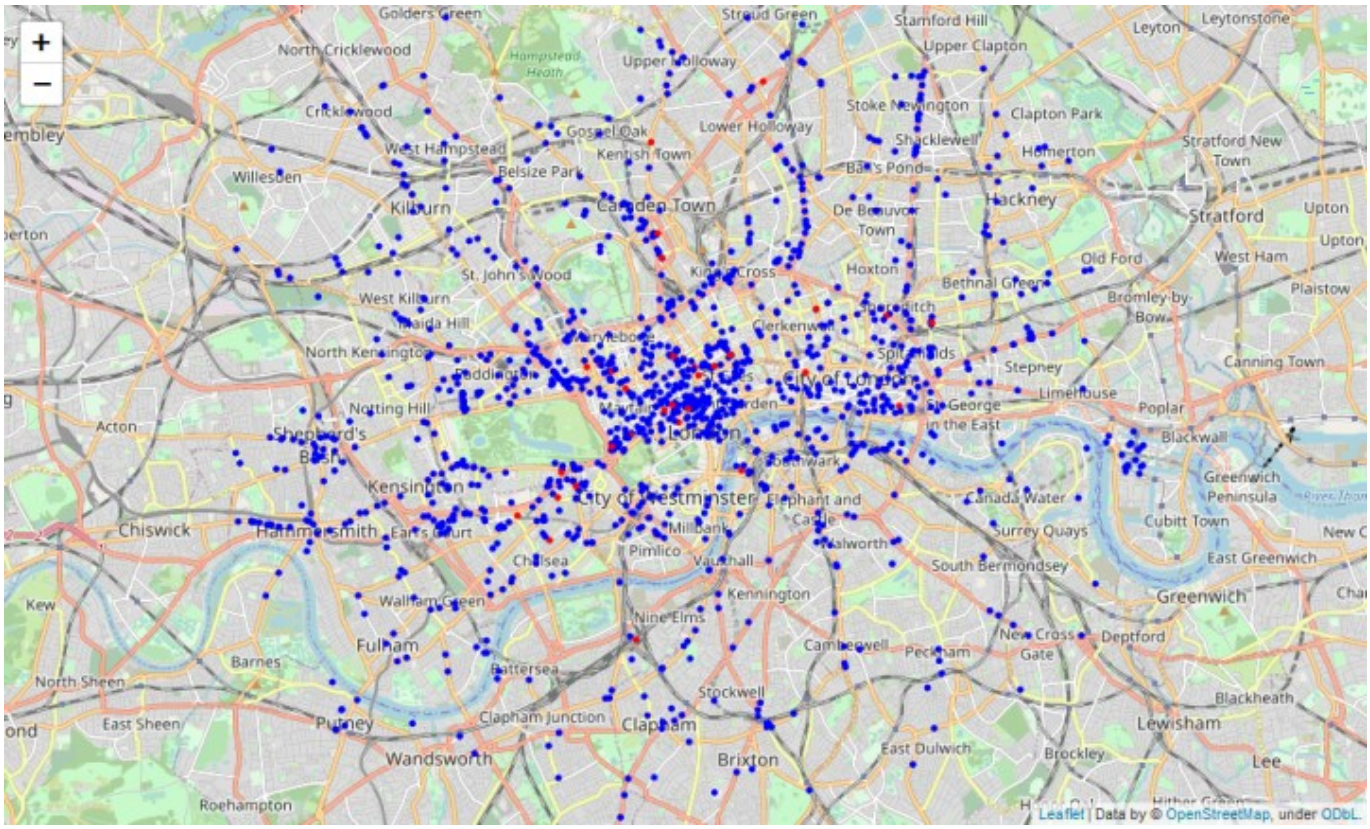

Figure 3. Distribution of restaurants (blue circles) and Japanese restaurants (red circles) within 8 km from the city center of London.

From the figure, it is already evident that Japanese restaurants represent a small minority of all restaurants. Furthermore, as expected there are more restaurants close to the city center and they get progressively less dense as the distance from the center increases.

More in detail, according to Foursquare there are 1315 restaurants within 8 km from the city center. 33 of these are Japanese restaurants which means 0.025% of all restaurants are Japanese. Table 3 shows how these statistics change when considering concentric areas progressively more distant from the city center.

Area	N° restaurants	N° Japanese restaurants	% Japanese restaurants
0-2667 meters	522	21	0.04
2667-5333 meters	477	10	0.02
5333-8000 meters	316	2	0.006

Table 3. Distribution of restaurants and Japanese restaurants in different areas.

As expected, the number of restaurants declines with increasing distance from the city center. Interestingly, the number of Japanese restaurants declines even more (from 0.04 to 0.006% of the total).

4.2 Clustering neighborhoods

In order to proceed with cluster analysis of neighborhoods, I first computed the number of restaurants and Japanese restaurants for each neighborhood (see again Figure 2 and Table 1). The updated neighborhood table is shown below.

Then, I normalized the data to cluster (i.e. Distance to city center, number of restaurants and number of Japanese restaurants). Using the unsupervised learning K-means algorithm I found the optimal number of clusters = 5 based on the elbow method (see Figure 4).

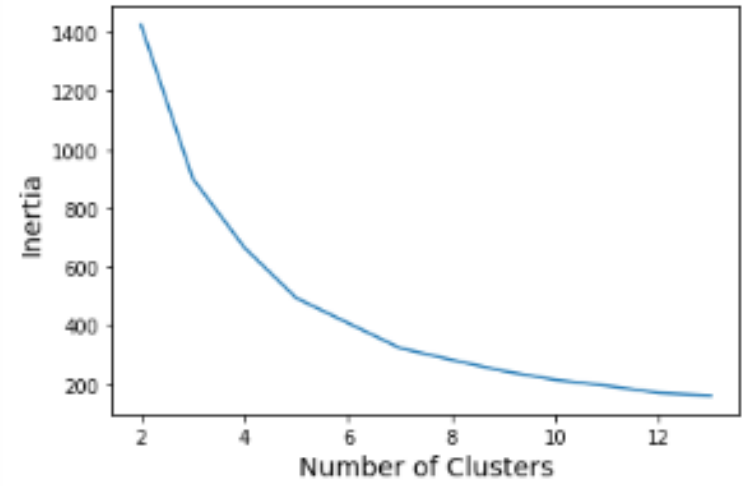

Figure 4. Inertia values for each number of clusters using K-means algorithm.

Table 4 shows the generated cluster labels added to the table of neighborhoods.

	Address	Latitude	Longitude	Latitude_m	Longitude_m	Distance_to_center	N_restaurants	N_japanese	Cluster
0	Sarsfeld Road, Tooting Graveney, Upper Tooting...	51.442558	-0.160618	5.702655e+06	697327.162989	7959.650398	0	0	1
1	5, Balham High Road, Balham, London Borough of...	51.442558	-0.153432	5.702655e+06	697827.162989	7868.345241	1	0	1
2	65, Fernlea Road, London Borough of Lambeth, L...	51.442558	-0.146247	5.702655e+06	698327.162989	7808.052203	0	0	1
3	20, Glenfield Road, Abbeville Village, London ...	51.442558	-0.139063	5.702655e+06	698827.162989	7779.492373	0	0	1
4	Kirkstall Road, Streatham Hill, London Borough...	51.442558	-0.131878	5.702655e+06	699327.162989	7783.015094	0	0	1

Table 4. Neighborhood info including number of restaurants, Japanese restaurants and cluster labels.

Table 5 shows the centroid values for each of the five clusters.

Cluster	Latitude	Longitude	Latitude_m	Longitude_m	Distance_to_center	N_restaurants	N_japanese
0	51.514467	-0.137267	5.710662e+06	698952.162989	2764.643007	9.723684	0.0
1	51.510641	-0.134038	5.710236e+06	699177.038919	6816.421153	0.843672	0.0
2	51.514408	-0.147146	5.710655e+06	698264.662989	1151.701883	28.750000	2.5
3	51.518257	-0.132392	5.711084e+06	699291.448703	2816.959587	10.535714	1.0
4	51.511430	-0.135523	5.710324e+06	699073.539800	4134.602990	1.362319	0.0

Table 5. Centroid values for each of the five clusters.

Cluster analysis did a good job as it identified three clusters without Japanese restaurants. These clusters are cluster 0, 1 and 4.

Figure 5 shows all the clustered neighborhoods superimposed on London's map.

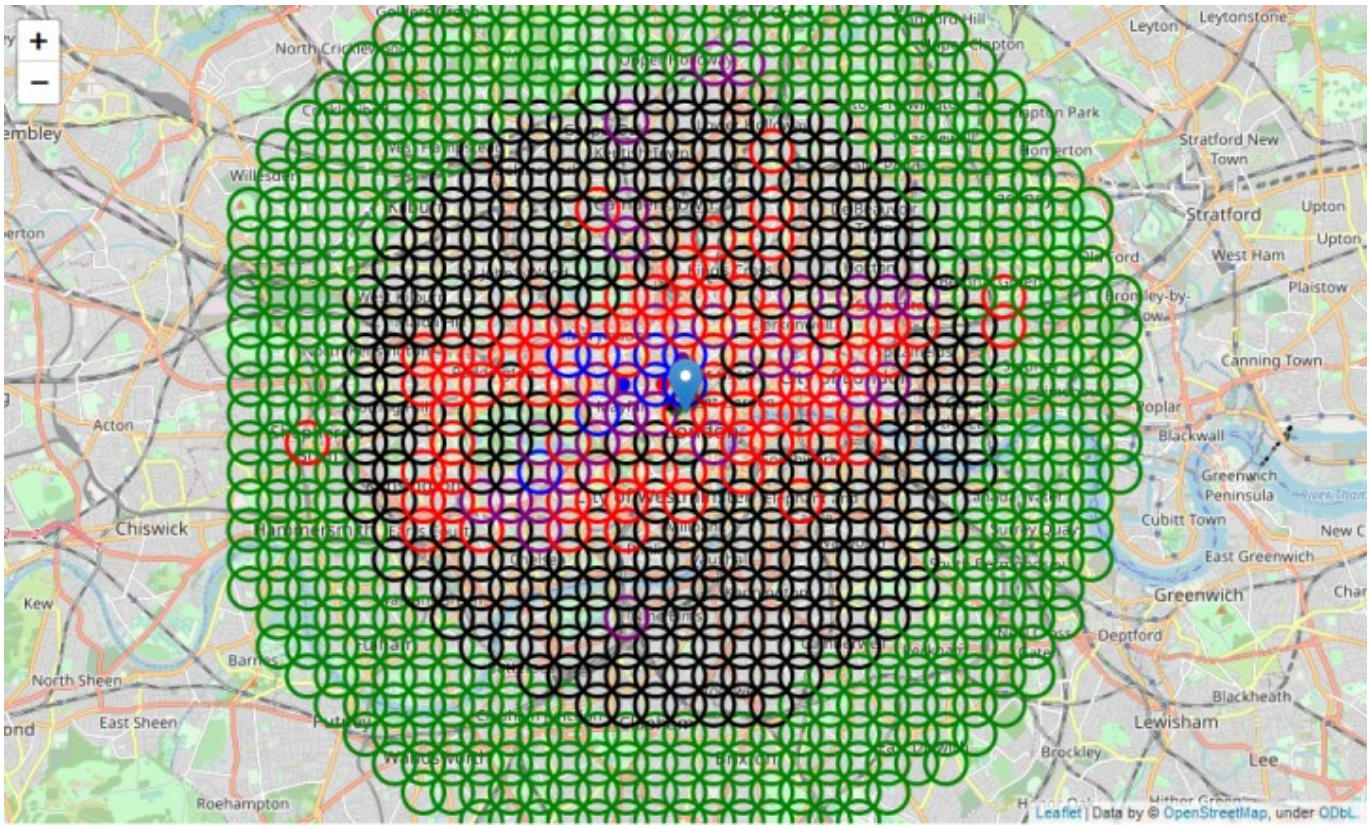

Figure 5. Color-coded clustered neighborhoods. Cluster 0 = red; cluster 1 = green; cluster 2 = blue; cluster 3 = purple; cluster 4 = black.

Cluster 2 (blue) and 3 (purple) although very close to the city center cannot be considered since they include areas with many restaurants and at least one Japanese restaurant for each neighborhood. Cluster 1 (green) contains areas too far from the city center. Hence, I will focus on cluster 0 and 4 (red and black). Since these clusters contain neighborhoods varying quite a bit in distance from the city center, I selected for each of these two clusters ten neighborhoods closest to the city center. Figure 6 shows these neighborhoods.

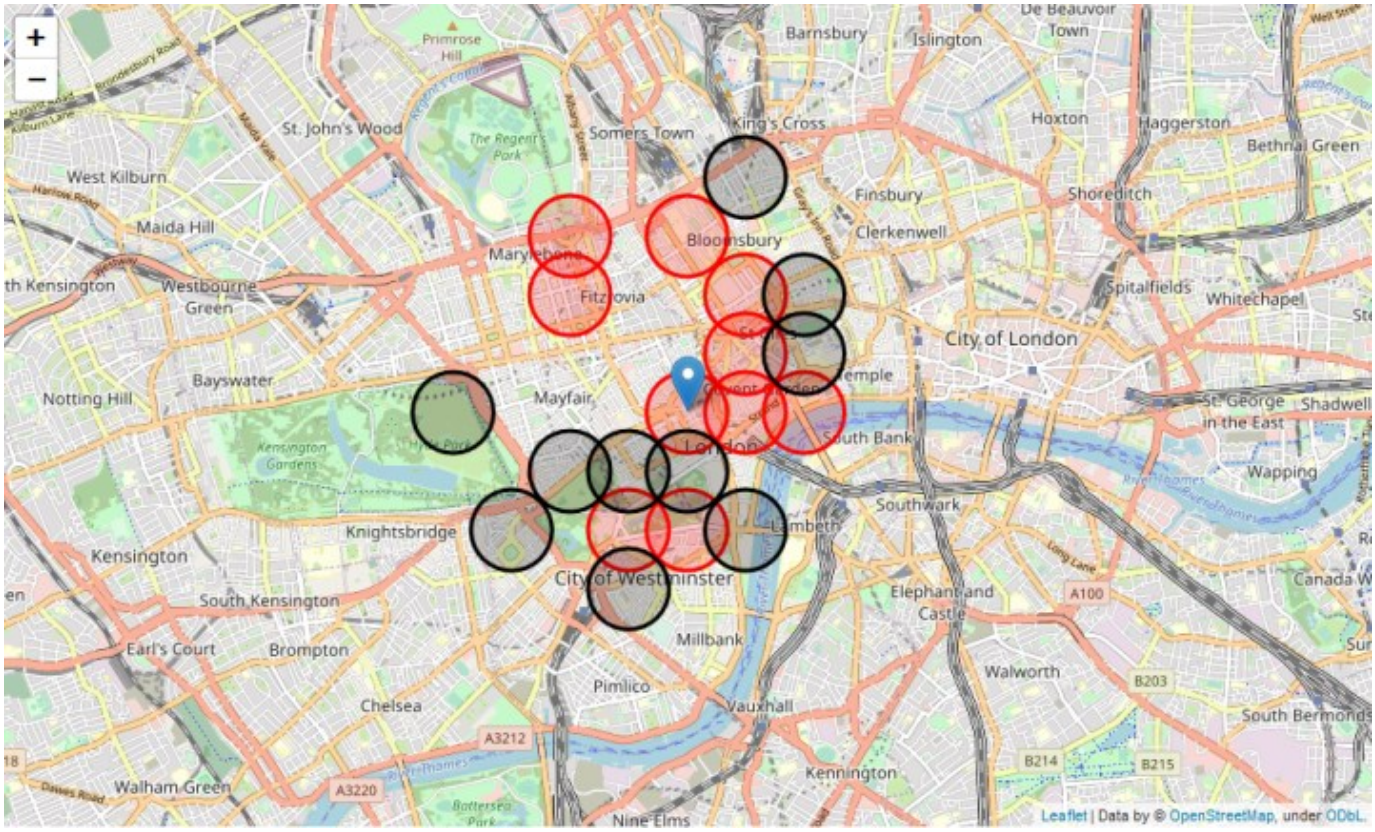

Figure 6. Ten neighborhoods closest to the city center for cluster 0 (red) and cluster 4 (black).

The neighborhoods of both clusters are very close to the city center. Both are also without Japanese restaurants inside. However, cluster 4 (black) has the advantage of having an average of 1.4 restaurants only in its areas whereas cluster 0 (red) has an average of 9.7 restaurants in each neighborhood. Hence, I restricted my selection to cluster 4. I also excluded the most western neighborhood since it includes an Hyde Park area which would might not allow to open a restaurant inside it. Figure 7 shows the final selection of candidate areas.

Figure 7. Final candidate areas for opening a new Japanese restaurant.

Thus, the final candidate areas are located in the **City of Westminster** and **between Covent Garden and King's Cross**. Table 6 reports the characteristics of each neighborhood including its center address , distance to center, number of restaurants and number of Japanese restaurants.

	Address	Latitude	Longitude	Latitude_m	Longitude_m	Distance_to_center	N_restaurants	N_japanese	Cluster
0	Chutney Mary, Little St. James's Street, St. J...	51.505427	-0.139063	5.709655e+06	698827.162989	801.181049	0	0	4
1	The Mall, St. James's, Victoria, City of Westm...	51.505427	-0.131878	5.709655e+06	699327.162989	834.693624	0	0	4
2	Embassy of Japan, 101-104, Piccadilly, St. Jam...	51.505427	-0.146247	5.709655e+06	698327.162989	1042.625868	1	0	4
3	St Clement Danes, Covent Garden, City of Westm...	51.514408	-0.117508	5.710655e+06	700327.162989	1323.733726	0	0	4
4	42, Eagle Street, Holborn, St Giles, London Bo...	51.518899	-0.117508	5.711155e+06	700327.162989	1491.719602	2	0	4
5	Caffè Nero, 70, Bridge Street, St. James's, Mi...	51.500936	-0.124693	5.709155e+06	699827.162989	1509.496415	1	0	4
6	Grosvenor Crescent, St. James's, Belgravia, Ci...	51.500936	-0.153432	5.709155e+06	697827.162989	1749.082598	1	0	4
7	Ashley Gardens, Ambrosden Avenue, St. James's,...	51.496445	-0.139063	5.708655e+06	698827.162989	1787.729926	2	0	4
8	Jessel House, Judd Street, Holborn, Bloomsbury...	51.527880	-0.124693	5.712155e+06	699827.162989	1901.661848	1	0	4

Table 6. Features of the final candidate areas.

To sum up, I identified several areas rather close to the city center that do not include Japanese restaurants within a radius of 350 meters and also with little number of non-Japanese restaurants

(mean = 1.4). Certainly, this analysis represents a preliminary selection of areas based on restaurant density. Therefore, it is possible that some of the selected areas might be not adapt to open a new restaurant because of some uncontrolled variables (e.g. high traffic or pollution, price, etc...). It should be considered as a first step towards an additional search based on additional parameters to find out the final location.

5. Conclusions

This project was aimed at verifying if there are the conditions based on density for opening a new Japanese restaurant in London. If these conditions exist another aim of the project was to identify good locations. This work identified nine areas which satisfied three constraints: 1) being close to the city center; 2) with very few restaurants; 3) with no Japanese restaurants. A relevant group of this areas is located in the City of Westminster. The other between Covent Garden and King's Cross. These areas seem optimal based on the criteria defined for this project.

A final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods not considered in this project such as the attractiveness of each location, real estate availability, prices, social and economic characteristics of neighborhoods, etc...