

STORE LOCATION IN LONDON



Claudio Ikeda

Table of Contents

| | | |
|-------------|---|-----------|
| I. | INTRODUCTION | 1 |
| A. | DESCRIPTION AND DISCUSSION OF BACKGROUND | 1 |
| B. | PROBLEM DESCRIPTION | 2 |
| II. | DATA | 3 |
| III. | METHODOLOGY | 4 |
| A. | EXPLORATORY DATA ANALYSIS | 4 |
| B. | PROBLEM APPROACH & USING K-MEANS CLUSTERING | 5 |
| IV. | RESULTS | 8 |
| V. | DISCUSSION | 17 |
| VI. | CONCLUSION | 17 |

I. INTRODUCTION

A. DESCRIPTION & DISCUSSION OF THE BACKGROUND

London is a megacity, the capital of the United Kingdom and one of the oldest of the world's great cities, with its history spanning nearly two millennia. By far Britain's largest metropolis, it is also the country's economic, transportation, and cultural center. The population of London had already exceeded one million by 1800 and it reached 9 million in 2018. London and the UK's population has one of the most diverse groups of origin countries in the world. Also, London is one of the world's leading tourism destinations with 21 million international visitors in 2018.

In recent years the restaurant industry in the United Kingdom and specifically London has undergone a period of growth. Consumer expenditure on restaurants and cafes reached close to £ 88 billion in 2017. The number of enterprises has been steadily increasing, contributing £ 17.9 billion to the British Economy.

Considering London's diversity and ethnicity it is evident that starting a restaurant business would earn you more money comparatively than most of the other businesses. However this diversity of nations and culinaries possibilities may be an issue for a WW brand with a Fast Food Behavior similarity. Although, with more profitable business there comes the most competition. This article can serve as one of the guides to start a store with different identities Indian, Italian, American, with different food sort variation based on the cultural tendency we have by providing a specific location. The number of restaurants in a specific location categorized based on cuisine and population distribution based on ethnicity and culture are some of the features considered for analysis.

B. PROBLEM DESCRIPTION

A restaurant is a business which prepares and serves food and drink to customers in return for money, either paid before the meal, after the meal or with an open account. London is famous for its excellent cuisine. Its food culture includes an array of international cuisines influenced by the city's immigrant history.

So, it is evident that to survive in such competitive market it is very important to strategically plan. Various factors need to be studied in order to decide on location such as;

1. London Population and demographics
2. Who are the competitors in that location?
3. Cuisine served / Menu of the competitors
4. Are there any venues like Tourist attractions, Entertainment zones, Parks etc., nearby where floating population is high.
5. Segmentation of the Borough
6. Untapped markets
7. Saturated markets etc.

And the list goes on...

Even though well-funded STARBUCKS Company Ltd. needs to choose the correct location to start its first venture. If this is successful, they can replicate the same in other locations. First move is very important, thereby choice of location is very important.

Target Audience:

To recommend the correct location, STARBUCKS Company Ltd has appointed me to lead of the Data Science team. The objective is to locate and recommend to the management which neighborhood of London will be best choice to start a store based in the customized nationality. The management also expects to understand the rationale of the recommendations made.

Success Criteria:

The success criteria of the project will be a good recommendation of neighborhood choice to STARBUCKS Company Ltd based on Lack of such restaurants considering cuisine as a factor in that location.

II. DATA

To build a recommendation model, following datasets and information are considered for analysis;

1. Scrapped Wikipedia using BeautifulSoup, to extract information about [32 London boroughs](#) also known as local authority districts. Also, considered [local areas or neighborhoods](#) for each borough for detailed analysis.
2. I used Foursquare API to get information about available restaurants for a given neighborhood and borough in London. The API also provided information about restaurant styles based on cuisine.
3. Employed data provided by the British Government from [data.london.gov.uk](#) to get more insights about London boroughs.

The data provided knowledge about the population density, immigrants' country of origin and many more.

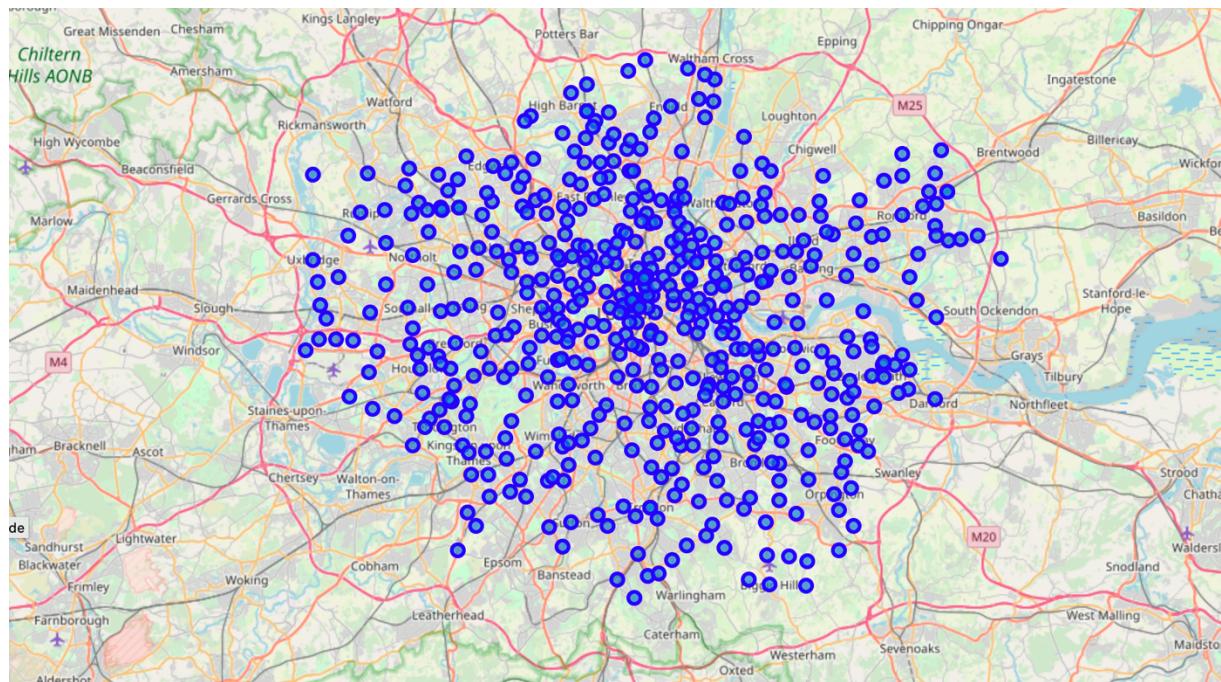
III. METHODOLOGY

Business Understanding:

Our main goal is to get optimum location for new restaurant business in London for STARBUCKS Company based on cuisine.

A. EXPLORATORY DATA ANALYSIS

London is mainly divided into 32 boroughs also known as local authority districts with 526 neighborhoods. Geopy and folium libraries to create a map to visualize neighborhoods of London.

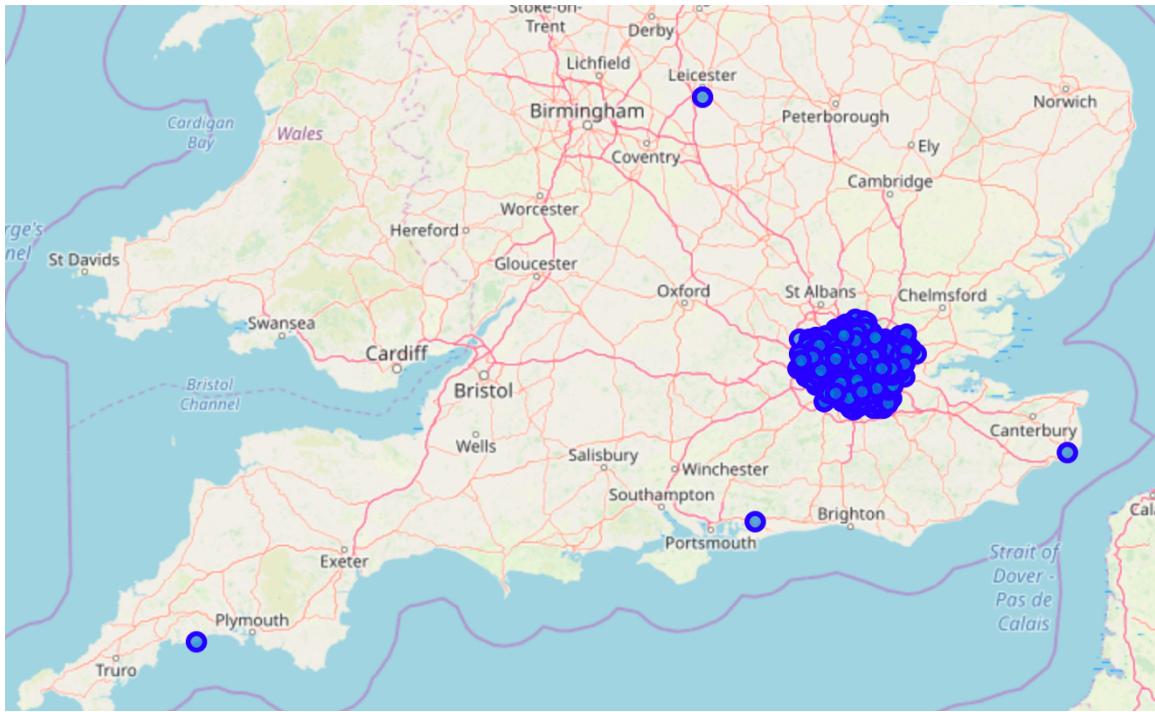


All the blue markers on the map above are for neighborhoods. It is evident that the city is more congested at the center and widespread in the outskirts.

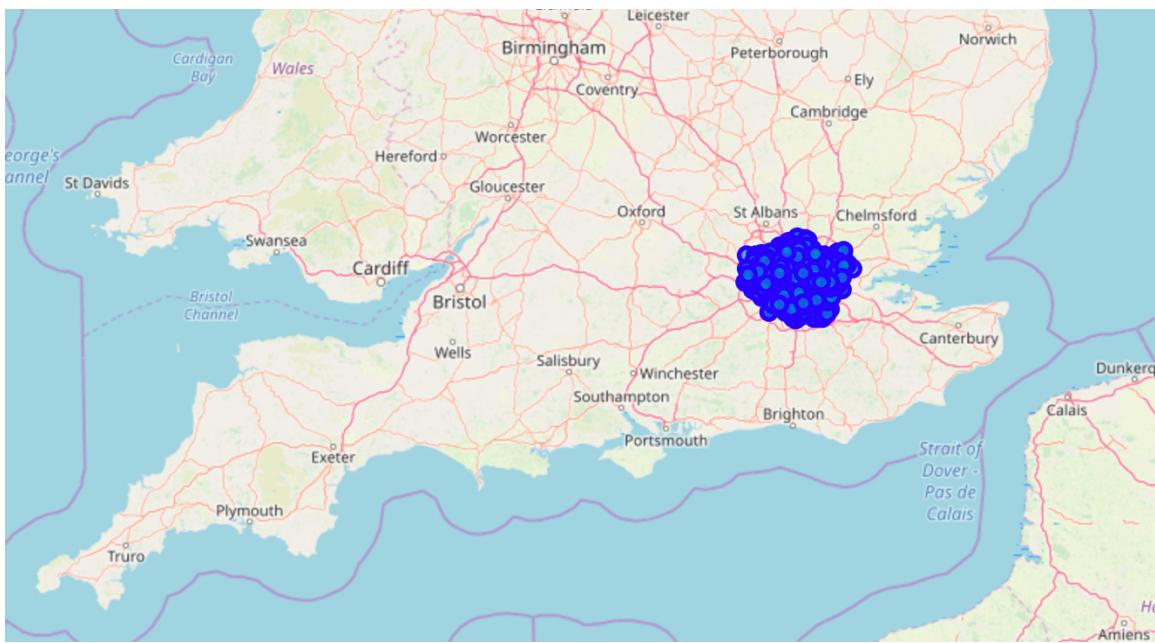
B. PROBLEM APPROACH AND USING K-MEANS CLUSTERING

The process starts with scrapping the Wikipedia web page containing information about London boroughs and neighborhoods present in each borough using library BeautifulSoup. During the process Pandas data frame is constructed with two rows having boroughs and corresponding neighborhood information.

The neighborhoods in the resulting data frame has to be plotted on a map using folium library. In order to do so, latitude and longitude value for each neighborhood is determined using geopy library. After getting latitude and longitude values, it is merged with the original data frame containing neighborhood and borough information appropriately. Finally, each neighborhood is plotted on folium map. Neighborhood with undetermined latitude and longitude values by geopy are excluded from evaluation. Also, the resulting values create four outliers and they are dealt using corresponding longitude values.



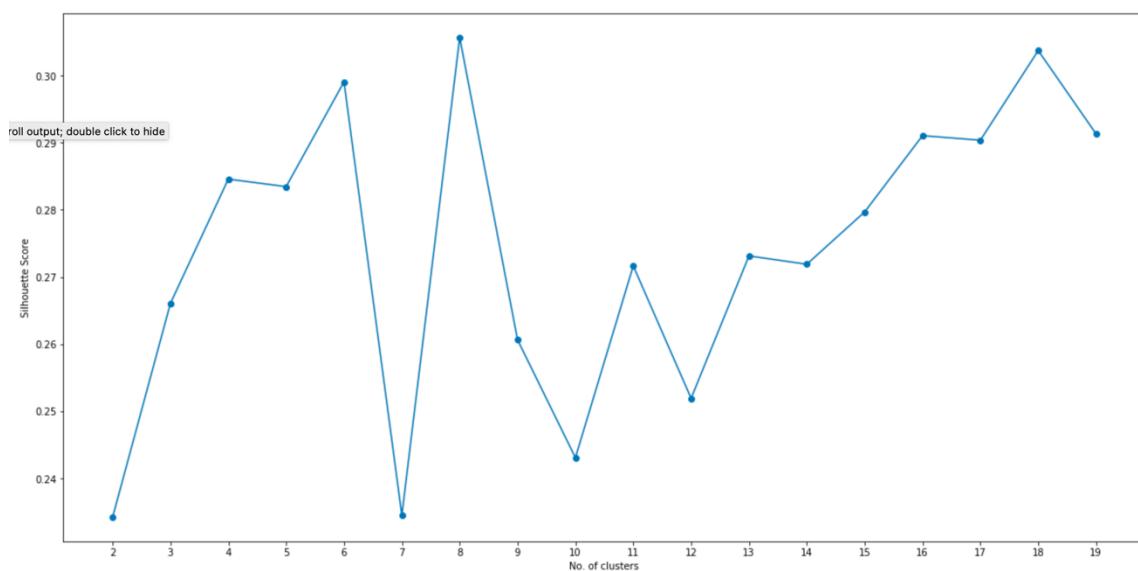
The extreme left location to be considered for study is Longford, Hillington with longitude value of '-0.490805'. Similarly, the extreme right is North Ockendon, Havering with longitude value of '0.293698'. So, any longitude value less than '-0.5' and greater than '0.3' will not be considered for evaluation.



Now Foursquare API is used to explore the neighborhoods and segment them. First, data frame with all the venues information provided by foursquare for the given latitude and longitude values. Foursquare API returns 12237 results with 423 unique categories. Now, only restaurants are extracted from venue category list. The resulting data frame has 87 unique categories or cuisines available in London. Next, one hot encoding is performed on the resulting data frame for each neighborhood. The results have 2657 unique restaurants in London with 87 different style of cuisines.

Rows are grouped by neighborhood to determine the frequency of occurrence of each restaurant. A new data frame is created with each row assigned for neighborhood and its corresponding top ten common restaurants based on cuisine.

Finally, k-means clustering is performed on the data frame to check the pattern for each neighborhood and get the information about the top ten common restaurants for each neighborhood. Before fitting the data frame best value of k for k-means clustering is determined by based on silhouette_score from sklearn.metrics. It is observed from the graph below that, $k = 8$ would yield more better results for the computation using k-means clustering.



Finally, k-means clustering is performed on the following data frame with $k = 8$, to determine the pattern of top 10 restaurants based on cuisine in every neighborhood.

| | Neighborhood | Neighborhood Latitude | Neighborhood Longitude | Venue | Venue Latitude | Venue Longitude | Venue Category | Cluster Labels | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue |
|----|--------------|-----------------------|------------------------|----------------------------|----------------|-----------------|----------------------|----------------|-----------------------|-----------------------|------------------------|
| 11 | Acton | 51.508140 | -0.273261 | Amigo's Peri Peri | 51.508396 | -0.274561 | Fast Food Restaurant | 2 | Japanese Restaurant | Thai Restaurant | Chinese Restaurant |
| 15 | Acton | 51.508140 | -0.273261 | Ting Tong Thai | 51.508363 | -0.277755 | Thai Restaurant | 2 | Japanese Restaurant | Thai Restaurant | Chinese Restaurant |
| 18 | Acton | 51.508140 | -0.273261 | North China Restaurant | 51.508251 | -0.277435 | Chinese Restaurant | 2 | Japanese Restaurant | Thai Restaurant | Chinese Restaurant |
| 24 | Acton | 51.508140 | -0.273261 | Hasu | 51.508167 | -0.269494 | Japanese Restaurant | 2 | Japanese Restaurant | Thai Restaurant | Chinese Restaurant |
| 30 | Addington | 51.358637 | -0.031635 | The Cricketers (Harvester) | 51.357833 | -0.032844 | English Restaurant | 1 | English Restaurant | Xinjiang Restaurant | Gluten-free Restaurant |

III. RESULTS

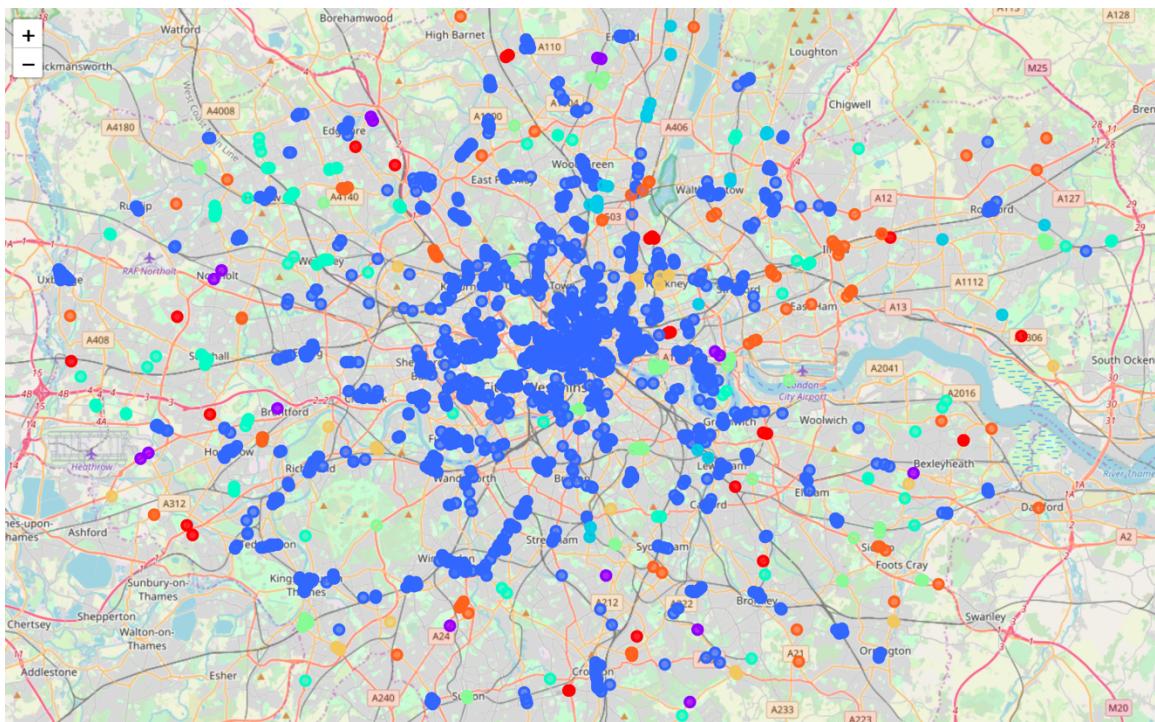
Neighborhood K-Means clustering based on mean occurrence of venue category:

All 8 clusters follow unique pattern for top ten common restaurants for a particular neighborhood. The detail shows the number of neighborhoods assigned to each cluster. Cluster 2 has the highest neighborhoods of 2306 and cluster 1 has the least with 16.

| | |
|---|------|
| 2 | 2306 |
| 4 | 97 |
| 7 | 77 |
| 5 | 64 |
| 3 | 36 |
| 6 | 33 |
| 0 | 28 |
| 1 | 16 |

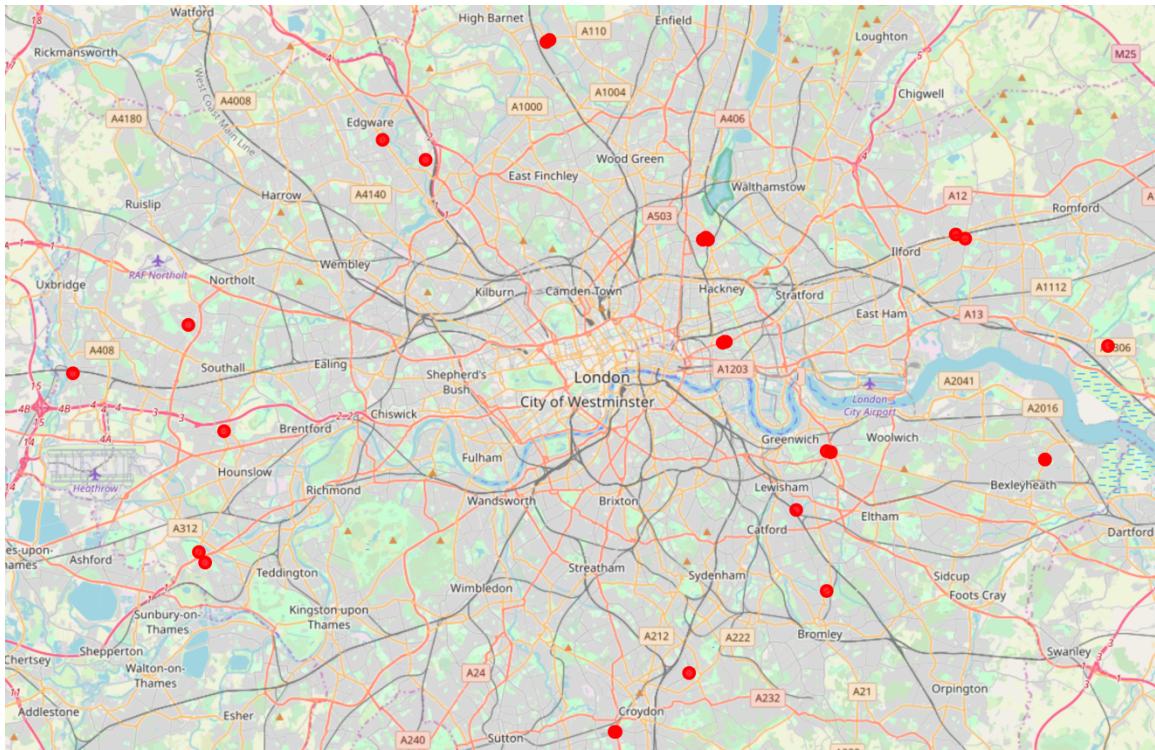
Name: Cluster Labels, dtype: int64

The clustered map is shown below:



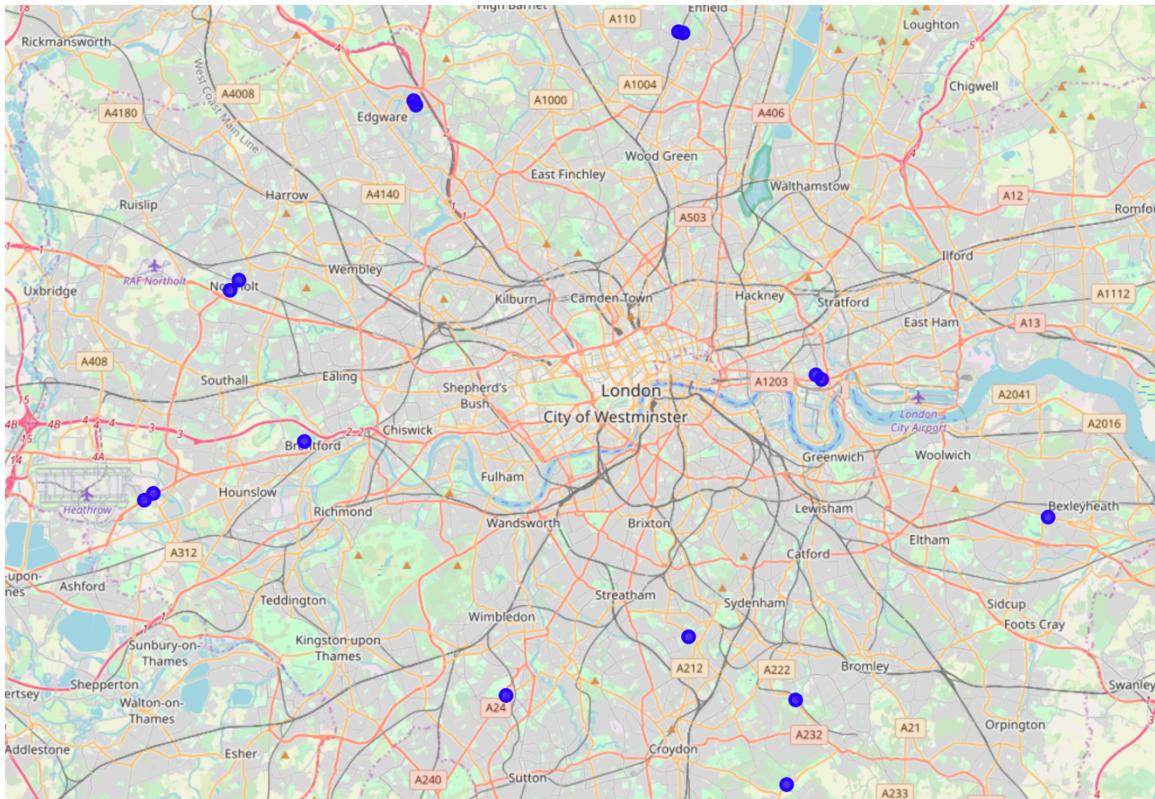
The resulting data frames for each cluster are shown below.

- Cluster 1 (Most Common: Chinese & Xinjiang Restaurant)



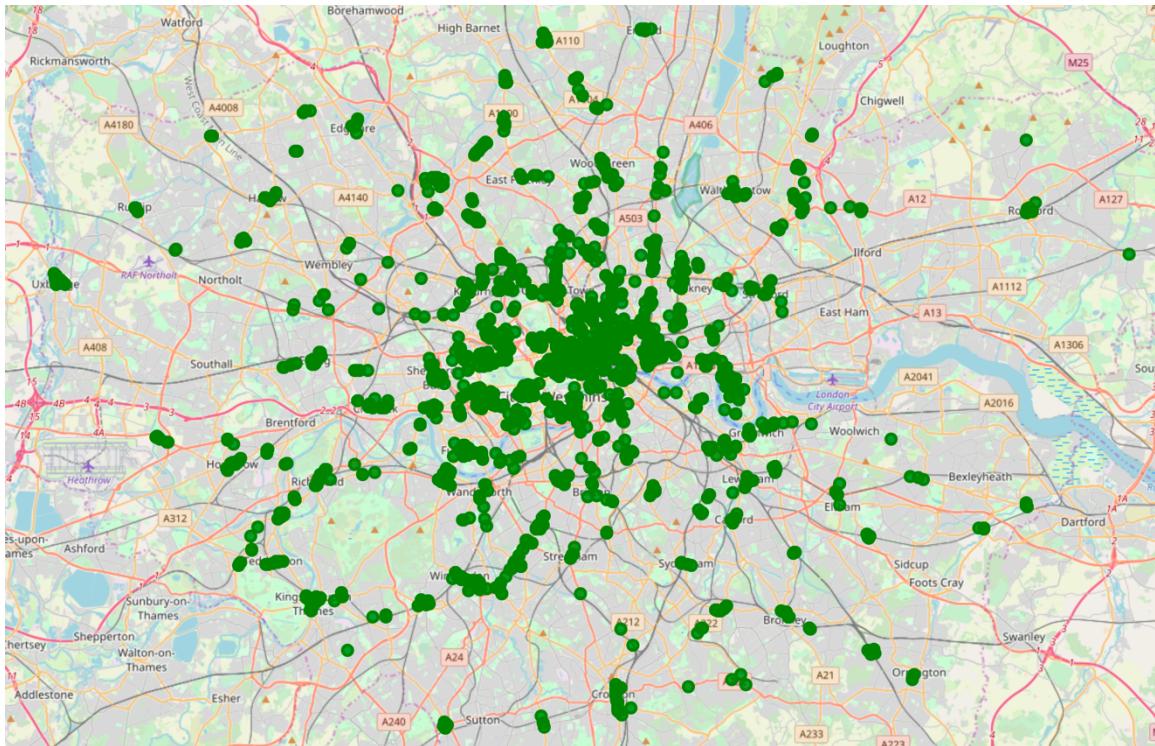
| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | Category |
|------|---------------------------|-----------------------|-----------------------|-----------------------|-----------------------|----------|
| 1285 | Blackheath Royal Standard | Chinese Restaurant | Fast Food Restaurant | Xinjiang Restaurant | Dumpling Restaurant | Food |
| 1686 | Burnt Oak | Chinese Restaurant | Xinjiang Restaurant | Dumpling Restaurant | Empanada Restaurant | Food |
| 2757 | Colindale | Chinese Restaurant | Xinjiang Restaurant | Dumpling Restaurant | Empanada Restaurant | Food |
| 3525 | East Barnet | Chinese Restaurant | Greek Restaurant | Xinjiang Restaurant | German Restaurant | Food |
| 4471 | Goodmayes | Chinese Restaurant | Fast Food Restaurant | Xinjiang Restaurant | Dumpling Restaurant | Food |
| 5007 | Hanworth | Restaurant | Chinese Restaurant | Xinjiang Restaurant | German Restaurant | Food |
| 5268 | Heston | Chinese Restaurant | Xinjiang Restaurant | Dumpling Restaurant | Empanada Restaurant | Food |

- Cluster 2 (Most Common: English Restaurant)



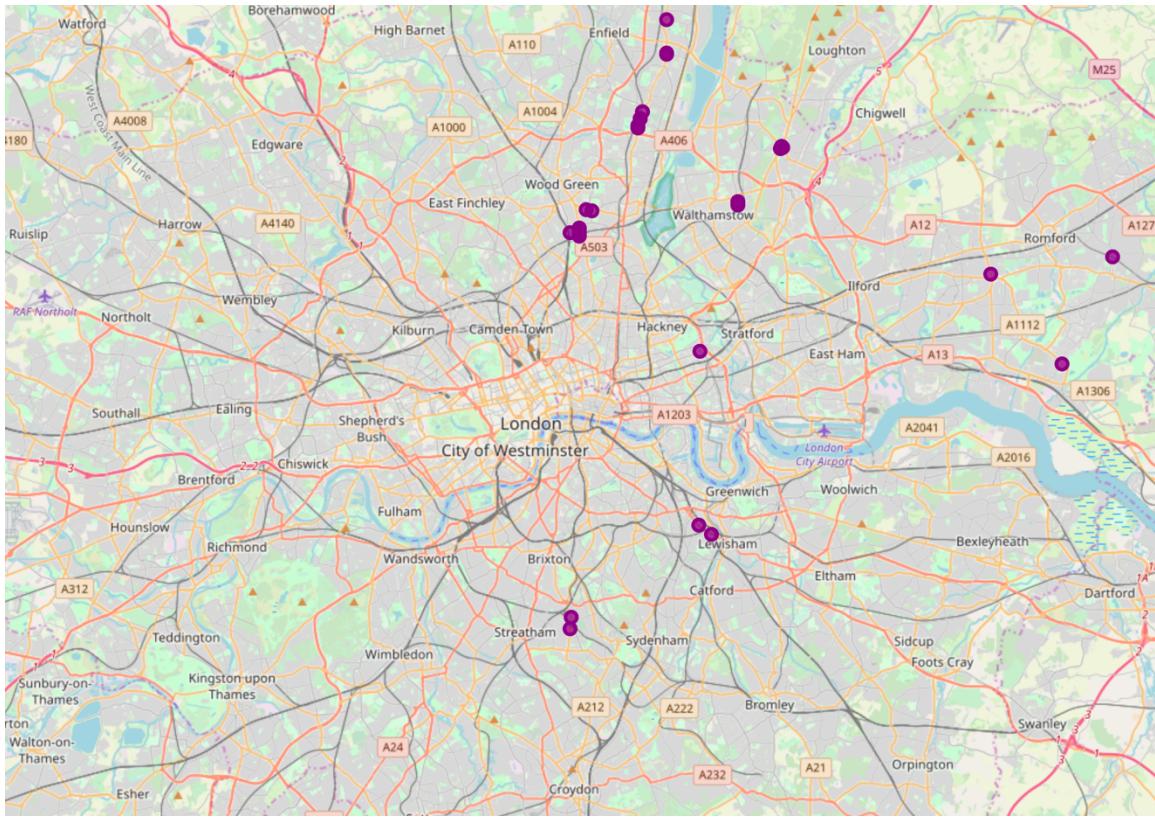
| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | Co |
|------|--------------|-----------------------|--------------------------|------------------------|------------------------|----|
| 30 | Addington | English Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant | |
| 1468 | Brentford | English Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant | |
| 3056 | Crook Log | English Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant | |
| 3659 | Eden Park | English Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant | |
| 4495 | Grange Park | English Restaurant | Indian Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | |
| 4814 | The Hale | English Restaurant | Mediterranean Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | |
| 5183 | Hatton | English Restaurant | Fast Food Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | |

- Cluster 3 (Most Common: Depends upon neighborhood)



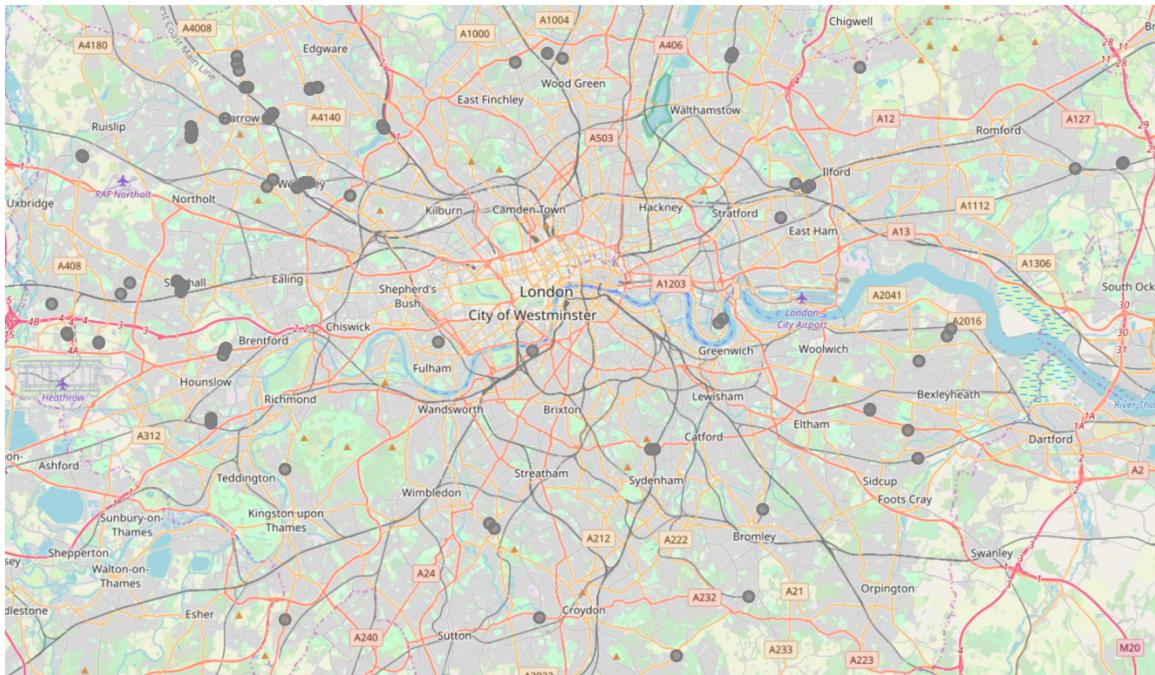
| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue |
|-----|--------------|---------------------------|-----------------------|-------------------------------|---------------------------|
| 11 | Acton | Japanese Restaurant | Thai Restaurant | Chinese Restaurant | Fast Food Restaurant |
| 58 | Aldgate | Indian Restaurant | Thai Restaurant | Asian Restaurant | Middle Eastern Restaurant |
| 152 | Aldwych | Restaurant | Seafood Restaurant | Sushi Restaurant | Italian Restaurant |
| 254 | Alperton | Middle Eastern Restaurant | Indian Restaurant | Asian Restaurant | Fast Food Restaurant |
| 282 | Angel | Sushi Restaurant | Vietnamese Restaurant | Mediterranean Restaurant | Indian Restaurant |
| 356 | Archway | Italian Restaurant | Japanese Restaurant | Vegetarian / Vegan Restaurant | Kebab Restaurant |
| 398 | Arnos Grove | French Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant |

- Cluster 4 (Most Common: Turkish Restaurant)



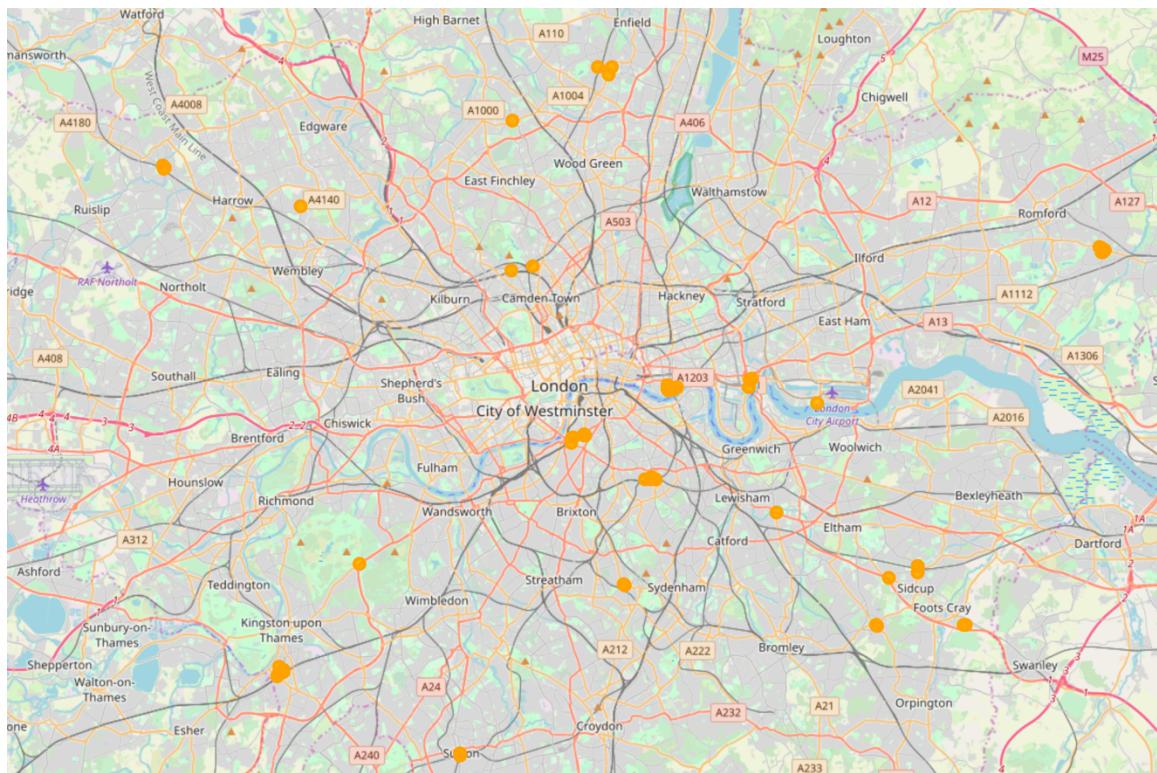
| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | C |
|------|-----------------|-----------------------|--------------------------|------------------------|------------------------|---|
| 898 | Becontree Heath | Turkish Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant | |
| 3691 | Edmonton | Turkish Restaurant | Kebab Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | |
| 3819 | Emerson Park | Turkish Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant | |
| 3827 | Enfield Highway | Turkish Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant | |
| 4282 | Freezywater | Turkish Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant | |
| 5050 | Harringay | Turkish Restaurant | Mediterranean Restaurant | Tapas Restaurant | Fast Food Restaurant | |
| 7756 | Old Ford | Turkish Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant | |

- Cluster 5 (Most Common: Indian Restaurant)



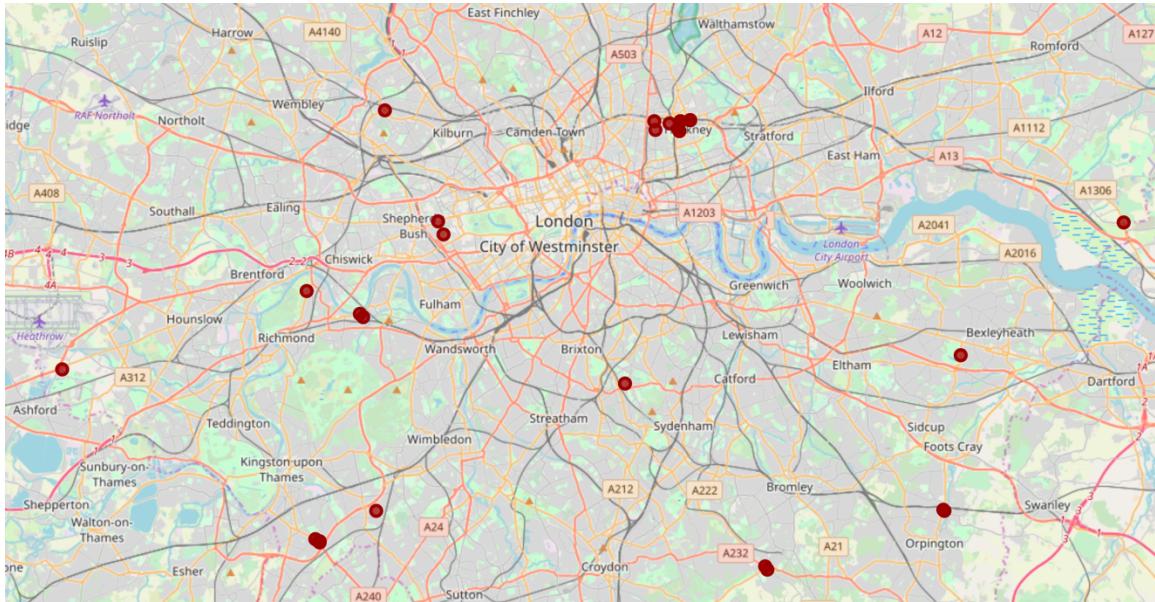
| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue |
|------|--------------|-----------------------------|-----------------------|--------------------------|-----------------------|
| 47 | Albany Park | Indian Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant |
| 908 | Beddington | Indian Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant |
| 1052 | Belvedere | Eastern European Restaurant | Indian Restaurant | Gluten-free Restaurant | Empanada Restaurant |
| 1312 | Blendon | Indian Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant |
| 1426 | Bounds Green | Indian Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant |
| 1449 | Bowes Park | Greek Restaurant | Indian Restaurant | Gluten-free Restaurant | Empanada Restaurant |
| 1522 | Brent Park | Indian Restaurant | Fast Food Restaurant | Mediterranean Restaurant | Xinjiang Restaurant |

- Cluster 6 (Most Common: Italian Restaurant)



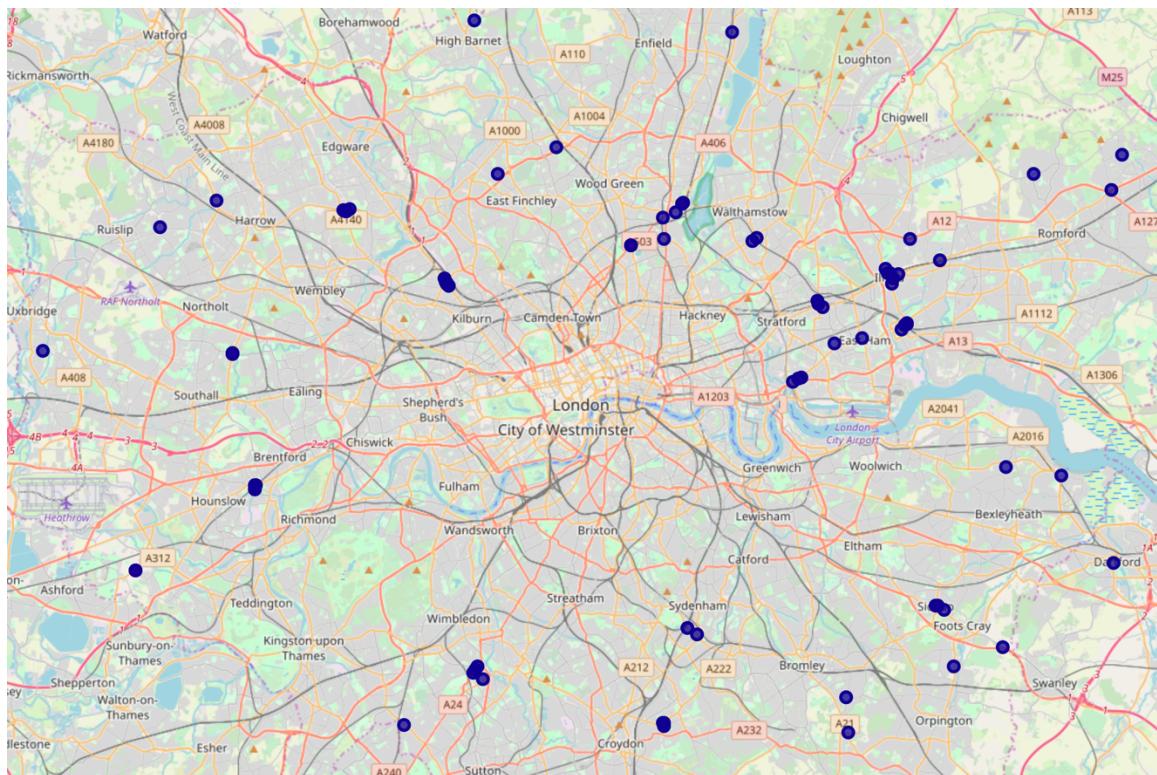
| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue |
|------|---------------|-----------------------|-----------------------|------------------------|------------------------|
| 1294 | Blackwall | Italian Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant |
| 2453 | Chislehurst | Italian Restaurant | Indian Restaurant | Xinjiang Restaurant | German Restaurant |
| 4285 | Friern Barnet | Italian Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant |
| 4427 | Gipsy Hill | Italian Restaurant | Indian Restaurant | Xinjiang Restaurant | German Restaurant |
| 4479 | Gospel Oak | Italian Restaurant | French Restaurant | Xinjiang Restaurant | Gluten-free Restaurant |
| 5604 | Hornchurch | Italian Restaurant | American Restaurant | Thai Restaurant | Portuguese Restaurant |
| 6240 | Kingston Vale | Italian Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant |

- Cluster 7 (Most Common: Multi-cuisine Restaurant)



| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue |
|-------------|---------------------|------------------------------|------------------------------|------------------------------|------------------------------|
| 2805 | Coney Hall | Restaurant | Dumpling Restaurant | Eastern European Restaurant | Empanada Restaurant |
| 3213 | Dalston | Restaurant | Modern European Restaurant | Israeli Restaurant | Indonesian Restaurant |
| 3323 | Dollis Hill | Restaurant | Dumpling Restaurant | Eastern European Restaurant | Empanada Restaurant |
| 3340 | Dulwich | Restaurant | Dumpling Restaurant | Eastern European Restaurant | Empanada Restaurant |
| 3533 | East Bedfont | Restaurant | Dumpling Restaurant | Eastern European Restaurant | Empanada Restaurant |
| 4630 | Hackney | Vietnamese Restaurant | Restaurant | Italian Restaurant | Israeli Restaurant |
| 5214 | Hazelwood | Restaurant | Dumpling Restaurant | Eastern European Restaurant | Empanada Restaurant |

- Cluster 8 (Most Common: Fast Food Restaurant)



| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue |
|-------------|---------------------|------------------------------|------------------------------|------------------------------|------------------------------|
| 37 | Addiscombe | Fast Food Restaurant | Chinese Restaurant | Xinjiang Restaurant | Dumpling Restaurant |
| 627 | Barking | Fast Food Restaurant | Portuguese Restaurant | Chinese Restaurant | Xinjiang Restaurant |
| 1533 | Brimsdown | Fast Food Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant |
| 1636 | Bromley Common | Fast Food Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant |
| 1963 | Canning Town | Fast Food Restaurant | Turkish Restaurant | Xinjiang Restaurant | Gluten-free Restaurant |
| 2764 | Collier Row | Fast Food Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant |
| 2942 | Cowley | Fast Food Restaurant | Xinjiang Restaurant | Gluten-free Restaurant | Empanada Restaurant |

IV. DISCUSSION

The results can be approached in two ways;

1. *If STARBUCKS company want to open a store in preferred location and irrespective of cuisine, refer to that neighborhood in specific cluster and chose cuisine with the least common restaurant for better profits*
2. *If STARBUCKS company wants to open a store with a preferred cuisine and irrespective of location, refer to the cluster with the least number of restaurants with that specific cuisine and select one among the neighborhoods based on company's preference.*

V. CONCLUSION

London has so many restaurants, yet certain neighborhood or borough doesn't have a specific cuisine restaurant available.

Depending on the culinary and focus different clusters in the city appears. If we want to test stores close to fast food restaurants, Italian, Indian, French, American, English, multifood may be a useful beta testa to replicate the investment in the clusters

As per the neighborhood or restaurant type mentioned like Indian Restaurant analysis can be checked. A venue customized to the Indian lovers can be placed close to these venues.