

# Clustering the Neighborhoods of two Cities

# A Tale of Two cities - Clustering the Neighbourhoods of London and Paris

## 1. Introduction

A Tale of Two cities, a novel written by Charles Dickens was set in London and Paris which takes place during the French Revolution. These cities were both happening then and now. A lot has changed over the years and we now take a look at how the cities have grown.

London and Paris are quite the popular tourist and vacation destinations for people all around the world. They are diverse and multicultural and offer a wide variety of experiences that is widely sought after. We try to group the neighbourhoods of London and Paris respectively and draw insights to what they look like now.

## 2. Business Problem

The aim is to help tourists choose their destinations depending on the experiences that the neighbourhoods have to offer and what they would want to have. This also helps people make decisions if they are thinking about migrating to London or Paris or even if they want to relocate neighbourhoods within the city. Our findings will help stakeholders make informed decisions and address any concerns they have including the different kinds of cuisines, provision stores and what the city has to offer.

## 3. Data Description

We require geographical location data for both London and Paris. Postal codes in each city serve as a starting point. Using Postal codes we use can find out the neighborhoods, boroughs, venues and their most popular venue categories.

### 3.1 London

To derive our solution, We scrape our data from [https://en.wikipedia.org/wiki/List\\_of\\_areas\\_of\\_London](https://en.wikipedia.org/wiki/List_of_areas_of_London)

This wikipedia page has information about all the neighbourhoods, we limit it London.

1. *borough* : Name of Neighbourhood
2. *town* : Name of borough
3. *post\_code* : Postal codes for London.

This wikipedia page lacks information about the geographical locations. To solve this problem we use ArcGIS API

### 3.2 ArcGIS API

ArcGIS Online enables you to connect people, locations, and data using interactive maps. Work with smart, data-driven styles and intuitive analysis tools that deliver location intelligence. Share your insights with the world or specific groups.

More specifically, we use ArcGIS to get the geo locations of the neighbourhoods of London. The following columns are added to our initial dataset which prepares our data.

4. *latitude* : Latitude for Neighbourhood
5. *longitude* : Longitude for Neighbourhood

### 3.3 Paris

To derive our solution, We leverage JSON data available at <https://www.data.gouv.fr/fr/datasets/r/e88c6fda-1d09-42a0-a069-606d3259114e>

The JSON file has data about all the neighbourhoods in France, we limit it to Paris.

1. *postal\_code* : Postal codes for France
2. *nom\_comm* : Name of Neighbourhoods in France
3. *nom\_dept* : Name of the boroughs, equivalent to towns in France
4. *geo\_point\_2d* : Tuple containing the latitude and longitude of the Neighbourhoods.

### 3.4 Foursquare API Data

We will need data about different venues in different neighbourhoods of that specific borough. In order to gain that information we will use "Foursquare" locational information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API.

After finding the list of neighbourhoods, we then connect to the Foursquare API to gather information about venues inside each and every neighbourhood. For each neighbourhood, we have chosen the radius to be 500 meters.

The data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

1. *Neighbourhood* : Name of the Neighbourhood
2. *Neighbourhood Latitude* : Latitude of the Neighbourhood
3. *Neighbourhood Longitude* : Longitude of the Neighbourhood

## 4. Methodology

We will be creating our model with the help of Python so we start off by importing all the required packages.

```
import pandas as pd
import requests
import numpy as np
import matplotlib.cm as cm
import matplotlib.colors as colors
import folium
from sklearn.cluster import KMeans
```

Package breakdown:

- *Pandas* : To collect and manipulate data in JSON and HTML and then data analysis
- *requests* : Handle http requests
- *matplotlib* : Detailing the generated maps
- *folium* : Generating maps of London and Paris
- *sklearn* : To import Kmeans which is the machine learning model that we are using.

The approach taken here is to explore each of the cities individually, plot the map to show the neighbourhoods being considered and then build our model by clustering all of the similar neighbourhoods together and finally plot the new map with the clustered neighbourhoods. We draw insights and then compare and discuss our findings.

# Wiki London Data

|     | Location        | London borough                    | Post town      | Postcode district | Dial code | OS grid ref |
|-----|-----------------|-----------------------------------|----------------|-------------------|-----------|-------------|
| 0   | Abbey Wood      | Bexley, Greenwich [7]             | LONDON         | SE2               | 020       | TQ465785    |
| 1   | Acton           | Ealing, Hammersmith and Fulham[8] | LONDON         | W3, W4            | 020       | TQ205805    |
| 2   | Addington       | Croydon[8]                        | CROYDON        | CR0               | 020       | TQ375645    |
| 3   | Addiscombe      | Croydon[8]                        | CROYDON        | CR0               | 020       | TQ345665    |
| 4   | Albany Park     | Bexley                            | BEXLEY, SIDCUP | DA5, DA14         | 020       | TQ478728    |
| ... | ...             | ...                               | ...            | ...               | ...       | ...         |
| 528 | Woolwich        | Greenwich                         | LONDON         | SE18              | 020       | TQ435795    |
| 529 | Worcester Park  | Sutton, Kingston upon Thames      | WORCESTER PARK | KT4               | 020       | TQ225655    |
| 530 | Wormwood Scrubs | Hammersmith and Fulham            | LONDON         | W12               | 020       | TQ225815    |
| 531 | Yeading         | Hillingdon                        | HAYES          | UB4               | 020       | TQ115825    |
| 532 | Yiewsley        | Hillingdon                        | WEST DRAYTON   | UB7               | 020       | TQ063804    |

# Paris Raw Data

|   | datasetid                              | recordid                                 | fields   | geometry  | record_timestamp                  |
|---|--|--|--|---|-----------------------------------|
| 0 | correspondances-code-insee-code-postal | 21e809b1d4480333c8b6fe7addd8f3b06f343e2c | {'code_comm': '003',<br>'nom_dept': 'VAL-DE-MARNE... | {'type': 'Point',<br>'coordinates':<br>[2.3335102498... | 2016-09-<br>21T00:29:06.175+02:00 |
| 1 | correspondances-code-insee-code-postal | c38925e974a8875071da3eb1391a6935d9c97e07 | {'code_comm': '430',<br>'nom_dept': 'SEINE-ET-MAR... | {'type': 'Point',<br>'coordinates':<br>[2.7879422114... | 2016-09-<br>21T00:29:06.175+02:00 |
| 2 | correspondances-code-insee-code-postal | 7c0aa8ba7a7b4320a9cf5abf12288eb76e3eead8 | {'code_comm': '412',<br>'nom_dept': 'SEINE-ET-MAR... | {'type': 'Point',<br>'coordinates':<br>[2.5107818983... | 2016-09-<br>21T00:29:06.175+02:00 |
| 3 | correspondances-code-insee-code-postal | b123405b4d069c33725418aab20ca0b741f8a5d8 | {'code_comm': '598',<br>'nom_dept': 'VAL-D'OISE',... | {'type': 'Point',<br>'coordinates':<br>[2.3004997834... | 2016-09-<br>21T00:29:06.175+02:00 |
| 4 | correspondances-code-insee-code-postal | 33dea89ab43606076200134a51f2b9d2d7d62256 | {'code_comm': '040',<br>'nom_dept': 'SEINE-ET-MAR... | {'type': 'Point',<br>'coordinates':<br>[2.5699190953... | 2016-09-<br>21T00:29:06.175+02:00 |

# London Feature Engineered

|            | <b>borough</b>                 | <b>town</b>            | <b>post_code</b> | <b>latitude</b> | <b>longitude</b> |
|------------|--------------------------------|------------------------|------------------|-----------------|------------------|
| <b>0</b>   | Bexley, Greenwich              | LONDON                 | SE2              | 51.49245        | 0.12127          |
| <b>1</b>   | Ealing, Hammersmith and Fulham | LONDON                 | W3, W4           | 51.51324        | -0.26746         |
| <b>6</b>   | City                           | LONDON                 | EC3              | 51.51200        | -0.08058         |
| <b>7</b>   | Westminster                    | LONDON                 | WC2              | 51.51651        | -0.11968         |
| <b>9</b>   | Bromley                        | LONDON                 | SE20             | 51.41009        | -0.05683         |
| ...        | ...                            | ...                    | ...              | ...             | ...              |
| <b>523</b> | Redbridge                      | LONDON                 | IG8, E18         | 51.58977        | 0.03052          |
| <b>524</b> | Redbridge, Waltham Forest      | LONDON, WOODFORD GREEN | IG8              | 51.50642        | -0.12721         |
| <b>527</b> | Barnet                         | LONDON                 | N12              | 51.61592        | -0.17674         |
| <b>528</b> | Greenwich                      | LONDON                 | SE18             | 51.48207        | 0.07143          |
| <b>530</b> | Hammersmith and Fulham         | LONDON                 | W12              | 51.50645        | -0.23691         |



# Paris Feature Engineered

|    | postal_code | nom_comm                 | nom_dept | Latitude  | Longitude |
|----|-------------|--------------------------|----------|-----------|-----------|
| 0  | 75010       | PARIS-10E-ARRONDISSEMENT | PARIS    | 48.876029 | 2.361113  |
| 1  | 75016       | PARIS-16E-ARRONDISSEMENT | PARIS    | 48.860399 | 2.262100  |
| 2  | 75009       | PARIS-9E-ARRONDISSEMENT  | PARIS    | 48.876896 | 2.337460  |
| 3  | 75015       | PARIS-15E-ARRONDISSEMENT | PARIS    | 48.840155 | 2.293559  |
| 4  | 75002       | PARIS-2E-ARRONDISSEMENT  | PARIS    | 48.867903 | 2.344107  |
| 5  | 75011       | PARIS-11E-ARRONDISSEMENT | PARIS    | 48.859415 | 2.378741  |
| 6  | 75005       | PARIS-5E-ARRONDISSEMENT  | PARIS    | 48.844509 | 2.349859  |
| 7  | 75019       | PARIS-19E-ARRONDISSEMENT | PARIS    | 48.886869 | 2.384694  |
| 8  | 75020       | PARIS-20E-ARRONDISSEMENT | PARIS    | 48.863187 | 2.400820  |
| 9  | 75003       | PARIS-3E-ARRONDISSEMENT  | PARIS    | 48.863054 | 2.359361  |
| 10 | 75006       | PARIS-6E-ARRONDISSEMENT  | PARIS    | 48.848968 | 2.332671  |
| 11 | 75018       | PARIS-18E-ARRONDISSEMENT | PARIS    | 48.892735 | 2.348712  |
| 12 | 75008       | PARIS-8E-ARRONDISSEMENT  | PARIS    | 48.872527 | 2.312583  |
| 13 | 75013       | PARIS-13E-ARRONDISSEMENT | PARIS    | 48.828718 | 2.362468  |
| 14 | 75012       | PARIS-12E-ARRONDISSEMENT | PARIS    | 48.835156 | 2.419807  |
| 15 | 75007       | PARIS-7E-ARRONDISSEMENT  | PARIS    | 48.856083 | 2.312439  |

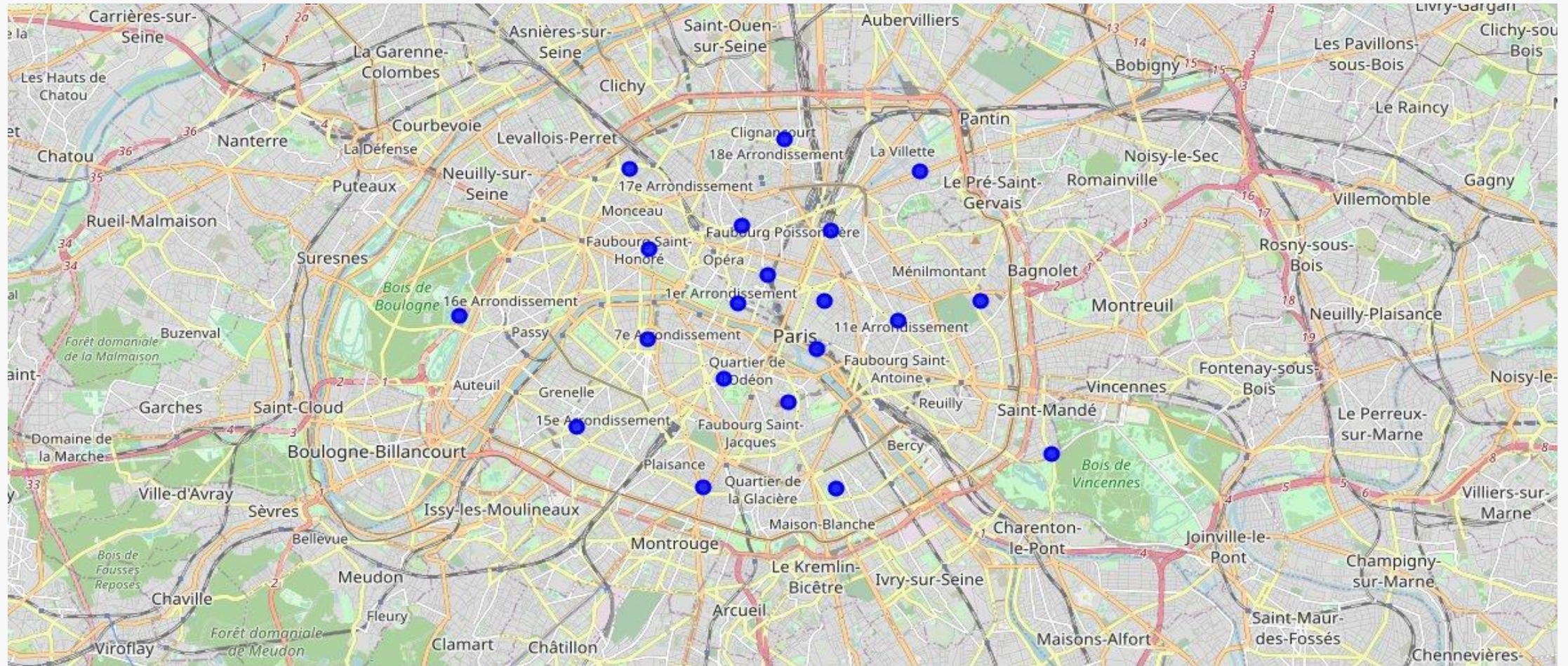


# London





# Paris



# Venues

|   | Neighbourhood     | Neighbourhood Latitude | Neighbourhood Longitude | Venue                            | Venue Category |
|---|-------------------|------------------------|-------------------------|----------------------------------|----------------|
| 0 | Bexley, Greenwich | 51.49245               | 0.12127                 | Sainsbury's                      | Supermarket    |
| 1 | Bexley, Greenwich | 51.49245               | 0.12127                 | Lesnes Abbey                     | Historic Site  |
| 2 | Bexley, Greenwich | 51.49245               | 0.12127                 | Lidl                             | Supermarket    |
| 3 | Bexley, Greenwich | 51.49245               | 0.12127                 | Abbey Wood Railway Station (ABW) | Train Station  |
| 4 | Bexley, Greenwich | 51.49245               | 0.12127                 | Bean @ Work                      | Coffee Shop    |



# One Hot Encoding

|   | Neighbourhood         | Accessories Store | Adult Boutique | African Restaurant | American Restaurant | Antique Shop | Arcade | Arepa Restaurant | Argentinian Restaurant | Art Gallery | Art Museum | Arts & Crafts Store | Asian Restaurant | Athletics & Sports |
|---|-----------------------|-------------------|----------------|--------------------|---------------------|--------------|--------|------------------|------------------------|-------------|------------|---------------------|------------------|--------------------|
| 0 | Barnet                | 0.0               | 0.0            | 0.0                | 0.001887            | 0.0          | 0.0    | 0.0              | 0.007547               | 0.0         | 0.0        | 0.0                 | 0.020755         | 0.0                |
| 1 | Barnet, Brent, Camden | 0.0               | 0.0            | 0.0                | 0.000000            | 0.0          | 0.0    | 0.0              | 0.000000               | 0.0         | 0.0        | 0.0                 | 0.000000         | 0.0                |
| 2 | Bexley                | 0.0               | 0.0            | 0.0                | 0.000000            | 0.0          | 0.0    | 0.0              | 0.000000               | 0.0         | 0.0        | 0.0                 | 0.000000         | 0.0                |
| 3 | Bexley, Greenwich     | 0.0               | 0.0            | 0.0                | 0.000000            | 0.0          | 0.0    | 0.0              | 0.000000               | 0.0         | 0.0        | 0.0                 | 0.000000         | 0.0                |
| 4 | Bexley, Greenwich     | 0.0               | 0.0            | 0.0                | 0.000000            | 0.0          | 0.0    | 0.0              | 0.000000               | 0.0         | 0.0        | 0.0                 | 0.000000         | 0.0                |

# Top 10 Categories

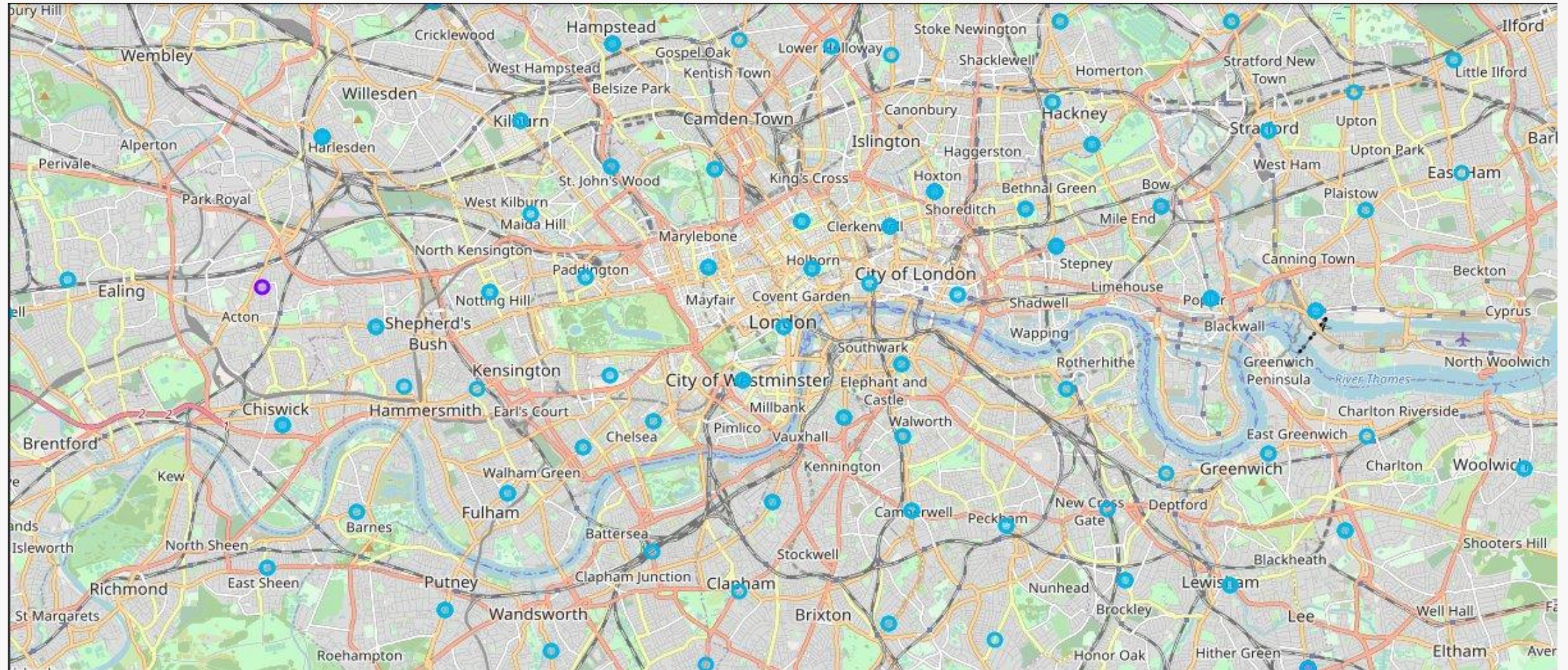
|   | Neighbourhood         | 1st Most Common Venue | 2nd Most Common Venue      | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue | 9th Most Common Venue      | 10th Most Common Venue |
|---|-----------------------|-----------------------|----------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|----------------------------|------------------------|
| 0 | Barnet                | Coffee Shop           | Café                       | Grocery Store         | Pub                   | Italian Restaurant    | Supermarket           | Pharmacy              | Chinese Restaurant    | Turkish Restaurant         | Pizza Place            |
| 1 | Barnet, Brent, Camden | Gym / Fitness Center  | Music Venue                | Clothing Store        | Supermarket           | Zoo Exhibit           | Film Studio           | Event Space           | Exhibit               | Falafel Restaurant         | Farmers Market         |
| 2 | Bexley                | Supermarket           | Historic Site              | Train Station         | Platform              | Convenience Store     | Coffee Shop           | Bus Stop              | Golf Course           | Construction & Landscaping | Park                   |
| 3 | Bexley, Greenwich     | Park                  | Construction & Landscaping | Sports Club           | Bus Stop              | Golf Course           | Historic Site         | Food Service          | Convenience Store     | Department Store           | Cycle Studio           |
| 4 | Bexley, Greenwich     | Supermarket           | Platform                   | Convenience Store     | Historic Site         | Train Station         | Coffee Shop           | Zoo Exhibit           | Film Studio           | Event Space                | Exhibit                |

# K-Means Labelled

|   | borough                        | town   | post_code | latitude | longitude | Cluster Labels | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | 6th Most Common Venue | 7th Most Common Venue | 8th Most Common Venue |
|---|--------------------------------|--------|-----------|----------|-----------|----------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 0 | Bexley, Greenwich              | LONDON | SE2       | 51.49245 | 0.12127   | 4              | Supermarket           | Platform              | Convenience Store     | Historic Site         | Train Station         | Coffee Shop           | Zoo Exhibit           | Film Studio           |
| 1 | Ealing, Hammersmith and Fulham | LONDON | W3, W4    | 51.51324 | -0.26746  | 1              | Grocery Store         | Train Station         | Breakfast Spot        | Park                  | Indian Restaurant     | Deli / Bodega         | Fish Market           | Exhibit               |
| 6 | City                           | LONDON | EC3       | 51.51200 | -0.08058  | 2              | Coffee Shop           | Italian Restaurant    | Hotel                 | Pub                   | Gym / Fitness Center  | Food Truck            | Sandwich Place        | Beer                  |
| 7 | Westminster                    | LONDON | WC2       | 51.51651 | -0.11968  | 2              | Hotel                 | Coffee Shop           | Pub                   | Sandwich Place        | Café                  | Italian Restaurant    | Restaurant            | Theatre               |
| 9 | Bromley                        | LONDON | SE20      | 51.41009 | -0.05683  | 2              | Supermarket           | Grocery Store         | Convenience Store     | Hotel                 | Fast Food Restaurant  | Park                  | Italian Restaurant    | Gym / Fitness Center  |

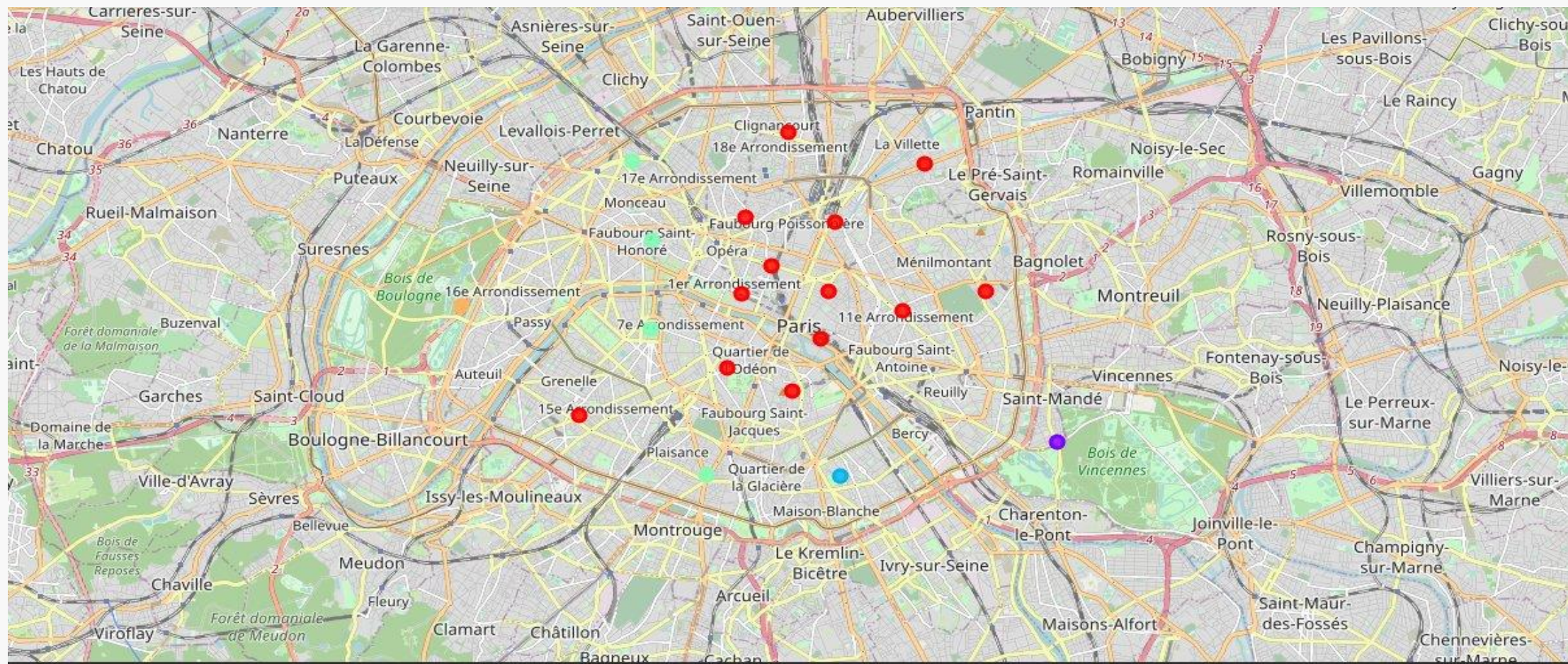


# London Clustered





# Paris Clustered





## 5. Results and Discussion

The neighbourhoods of London are very multicultural. There are a lot of different cuisines including Indian, Italian, Turkish and Chinese. London seems to take a step further in this direction by having a lot of Restaurants, bars, juice bars, coffee shops, Fish and Chips shop and Breakfast spots. It has a lot of shopping options too with that of the Flea markets, flower shops, fish markets, Fishing stores, clothing stores. The main modes of transport seem to be Buses and trains. For leisure, the neighbourhoods are set up to have lots of parks, golf courses, zoo, gyms and Historic sites. Overall, the city of London offers a multicultural, diverse and certainly an entertaining experience.

Paris is relatively small in size geographically. It has a wide variety of cuisines and eateries including French, Thai, Cambodian, Asian, Chinese etc. There are a lot of hangout spots including many Restaurants and Bars. Paris has a lot of Bistro's. Different means of public transport in Paris which includes buses, bikes, boats or ferries. For leisure and sight seeing, there are a lot of Plazas, Trails, Parks, Historic sites, clothing shops, Art galleries and Museums. Overall, Paris seems like the relaxing vacation spot with a mix of lakes, historic spots and a wide variety of cuisines to try out.

## 6. Conclusion

The purpose of this project was to explore the cities of London and Paris and see how attractive it is to potential tourists and migrants. We explored both the cities based on their postal codes and then extrapolated the common venues present in each of the neighbourhoods finally concluding with clustering similar neighbourhoods together.

We could see that each of the neighbourhoods in both the cities have a wide variety of experiences to offer which is unique in its own way. The cultural diversity is quite evident which also gives the feeling of a sense of inclusion.

Both Paris and London seem to offer a vacation stay or a romantic getaway with a lot of places to explore, beautiful landscapes, amazing food and a wide variety of culture. Overall, it's up to the stakeholders to decide which experience they would prefer more and which would more to their liking.