The Analysis of the NEIGHBOURHOODS of LONDON

London is much more than just a list of 32 boroughs, or the squiggly colored lines of the tube map. In this project, I try to analyse, understand, and explore neighbourhoods. To get the most common venue categories in each area, and then use this feature to group the neighbourhoods into clusters.

I use:

- The Foursquare API to get the relevant data for each neighbourhood and to explore each of the areas.

- Folium library helps to visualize the neighbourhoods in London City and their emerging clusters.

- The k-means clustering algorithm is used to grouping the data points into distinct subgroups.

This project may be useful for people coming to the City of London to help them to uncover the places that London has to offer, to help to understand how similar and diverse neighbourhoods in London are.

Data

public data from Wikipedia and Foursquare

https://en.wikipedia.org/wiki/List_of_London_boroughs

Loading the data:

In [192]:
```
table = pd.read_html("https://en.wikipedia.org/wiki/List_of_London_boroughs",encoding="utf-8-sig")
df = pd.DataFrame(table[0])
df.head()
```

Out[192]:

	0	1	2	3	4	5	6	7	8	9
0	Borough	Inner	Status	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 est)[1]	Co-ordinates	Nr. in map
1	Barking and Dagenham [note 1]	NaN	NaN	Barking and Dagenham London Borough Council	Labour	Town Hall, 1 Town Square	13.93	194352	51°33′39″N 0°09′21″E / 51.5607°N 0.1557°E	25
2	Barnet	NaN	NaN	Barnet London Borough Council	Conservative	Barnet House, 2 Bristol Avenue, Colindale	33.49	369088	51°37′31″N 0°09′06″W / 51.6252°N 0.1517°W	31
3	Bexley	NaN	NaN	Bexley London Borough Council	Conservative	Civic Offices, 2 Watling Street	23.38	236687	51°27′18″N 0°09′02″E / 51.4549°N 0.1505°E	23
4	Brent	NaN	NaN	Brent London Borough Council	Labour	Brent Civic Centre, Engineers Way	16.70	317264	51°33′32″N 0°16′54″W / 51.5588°N 0.2817°W	12

In [193]:
```
df.shape
```

Out[193]: (33, 10)

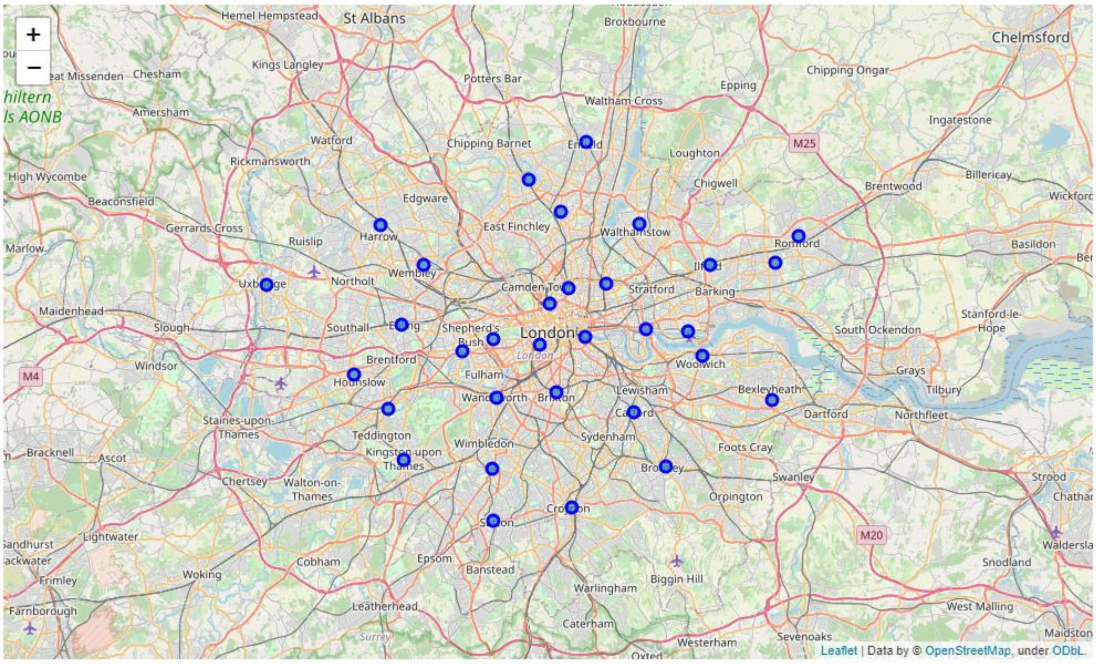
Cleaned dataset after some simple manipulations:

Out[180]:

	Borough	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 est)[1]	Co-ordinates	Nr. in map	Co-ordinates1	Latitude	Longitude
1	Barking and Dagenham [note 1]	Barking and Dagenham London Borough Council	Labour	Town Hall, 1 Town Square	13.93	194352	51°33′39″N 0°09′21″E / 51.5607°N 0.1557°E	25	51.5607°N 0.1557°E	51.5607	0.1557
2	Barnet	Barnet London Borough Council	Conservative	Barnet House, 2 Bristol Avenue, Colindale	33.49	369088	51°37′31″N 0°09′06″W / 51.6252°N 0.1517°W	31	51.6252°N 0.1517°W	51.6252	-0.1517
3	Bexley	Bexley London Borough Council	Conservative	Civic Offices, 2 Watling Street	23.38	236687	51°27′18″N 0°09′02″E / 51.4549°N 0.1505°E	23	51.4549°N 0.1505°E	51.4549	0.1505
4	Brent	Brent London Borough Council	Labour	Brent Civic Centre, Engineers Way	16.70	317264	51°33′32″N 0°16′54″W / 51.5588°N 0.2817°W	12	51.5588°N 0.2817°W	51.5588	-0.2817
5	Bromley	Bromley London Borough Council	Conservative	Civic Centre, Stockwell Close	57.97	317899	51°24′14″N 0°01′11″E / 51.4039°N 0.0198°E	20	51.4039°N 0.0198°E	51.4039	0.0198

Visualisation of the areas using Folium

Out[179]:



Adding venues to the dataset (using Foursquare API):

In [263]:
```
london_venues.head()
```

Out[263]:

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Barking and Dagenham [note 1]	51.5607	0.1557	Central Park	51.559560	0.161981	Park
1	Barking and Dagenham [note 1]	51.5607	0.1557	Lara Grill	51.562445	0.147178	Turkish Restaurant
2	Barking and Dagenham [note 1]	51.5607	0.1557	Hoo Hing	51.567561	0.135999	Grocery Store
3	Barking and Dagenham [note 1]	51.5607	0.1557	Asda	51.565751	0.143392	Supermarket
4	Barking and Dagenham [note 1]	51.5607	0.1557	Iceland	51.560578	0.147685	Grocery Store

In [264]:
```
london_venues.shape
```

Out[264]: (2781, 7)

Let’s try to visualize the data:

2 .Data Exploration

```
In [19]: print('There are {} uniques categories.'.format(len(london_venues['Venue Category'].unique())))
london_venue_unique_count = london_venues['Venue Category'].value_counts().to_frame(name='Count')
print ('Categories by count: ')
london_venue_unique_count
```

```
There are 276 uniques categories.
Categories by count:
```

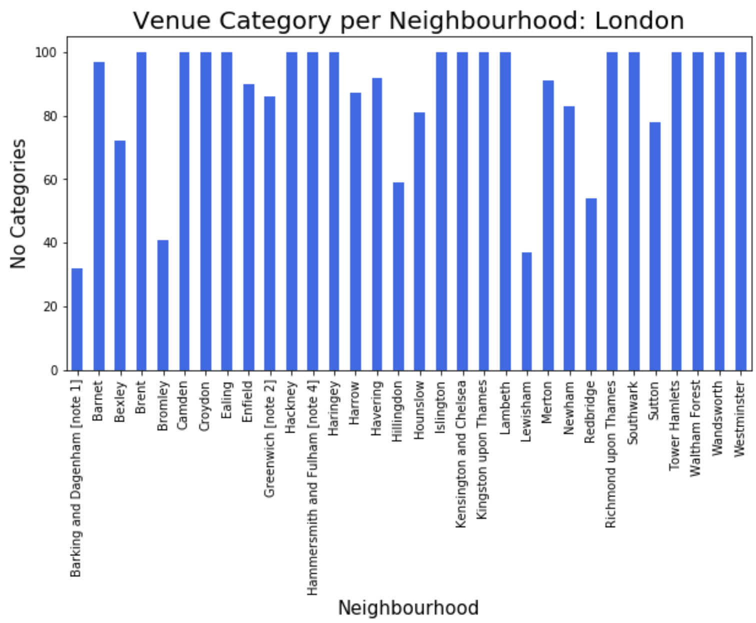
Out[19]:

	Count
Pub	223
Coffee Shop	193
Café	116
Park	109
Hotel	102
Grocery Store	97
Supermarket	79
Pizza Place	74
Italian Restaurant	70

We can see that most popular category are Pub and Coffee Shop. Lets check additional details.

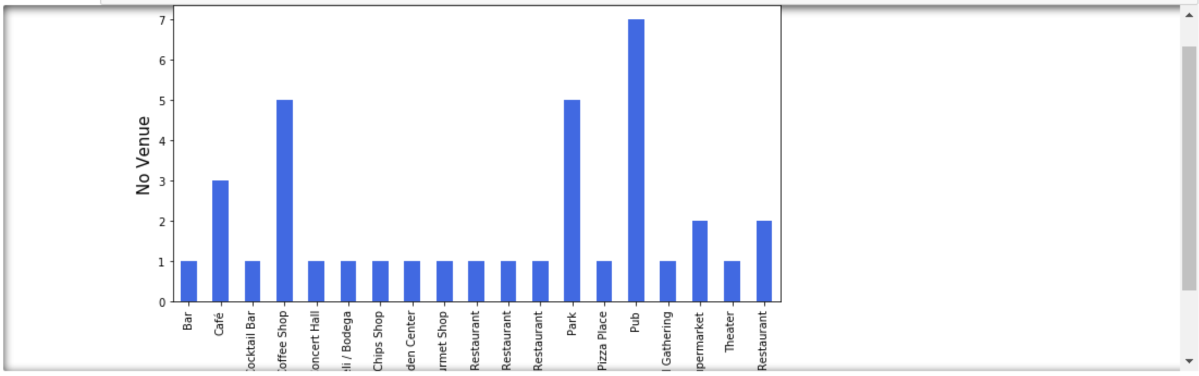
In the next chart we can see number of different categories per neighbourhood.

```
In [20]: clr = "royalblue"
london_venues.groupby('Neighbourhood')['Venue Category'].count().plot.bar(figsize=(10,5), color=clr)
plt.title('Venue Category per Neighbourhood: London', fontsize = 20)
plt.xlabel('Neighbourhood', fontsize = 15)
plt.ylabel('No Categories',fontsize = 15)
plt.xticks(rotation = 'vertical')
plt.show()
```



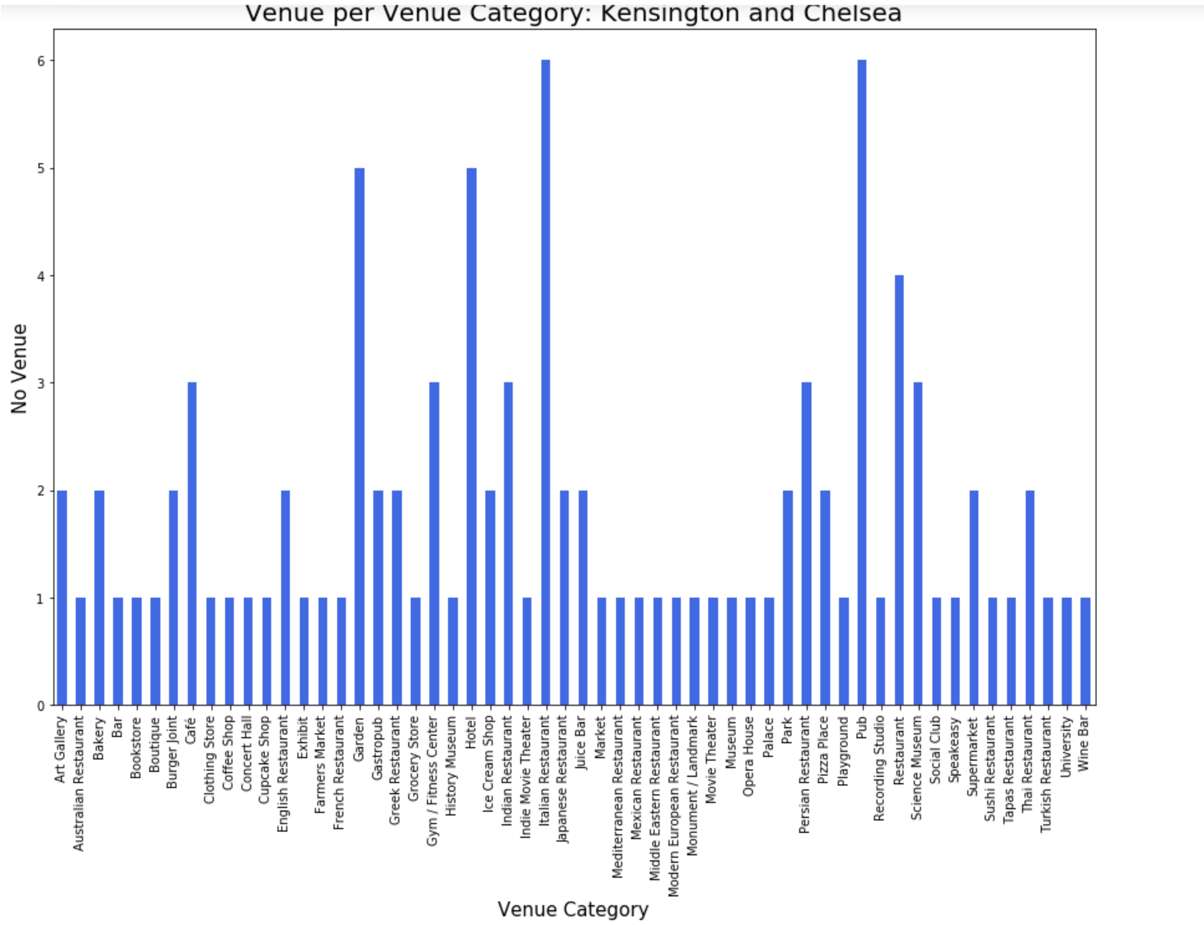
We can see that Barking and Dagenham or Lewisham have a relatively small number of venue categories, we can drill down to explore more details:

In [24]:
```python
df_Lewisham = london_venues[london_venues['Neighbourhood'] == 'Lewisham']
df_Lewisham.groupby('Venue Category')['Venue'].count().plot.bar(figsize=(10,5), color=clr)
plt.title('Venue per Venue Category: Lewisham', fontsize = 20)
plt.xlabel('Venue Category', fontsize = 15)
plt.ylabel('No Venue',fontsize = 15)
plt.xticks(rotation = 'vertical')
plt.show()
```



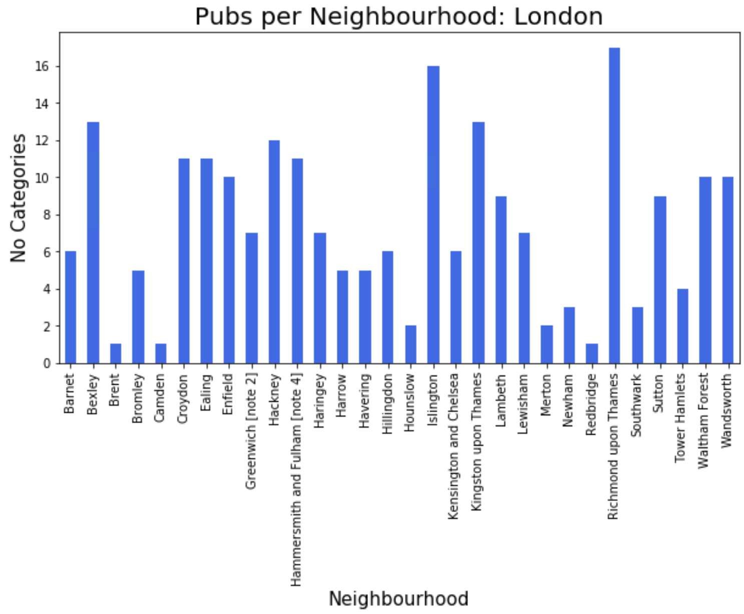
Now let's explore details of borough with bigger number if categories:

In [22]:
```python
df_KandC = london_venues[london_venues['Neighbourhood'] == 'Kensington and Chelsea']
df_KandC.groupby('Venue Category')['Venue'].count().plot.bar(figsize=(15,10), color=clr)
plt.title('Venue per Venue Category: Kensington and Chelsea', fontsize = 20)
plt.xlabel('Venue Category', fontsize = 15)
plt.ylabel('No Venue',fontsize = 15)
plt.xticks(rotation = 'vertical')
plt.show()
```



We can see that in both cases the most common category is Pub:

```
In [21]: df_pub = london_venues[london_venues['Venue Category'] == 'Pub']
         df_pub.groupby('Neighbourhood')['Venue Category'].count().plot.bar(figsize=(10,5), color=clr)
         plt.title('Pubs per Neighbourhood: London', fontsize = 20)
         plt.xlabel('Neighbourhood', fontsize = 15)
         plt.ylabel('No Categories',fontsize = 15)
         plt.xticks(rotation = 'vertical')
         plt.show()
```



Now, base on this data I'll try to build 5 clusters to compare/analyse the neighbourhoods.

Using One Hot encoding on the above data frame, create another data frame and then group all the neighbourhoods based on the frequency of occurrence of each venue category.

```
In [26]: ln_onehot = pd.get_dummies(london_venues[['Venue Category']], prefix = "", prefix_sep = "")
         # add neighborhood column back to dataframe
         ln_onehot['Neighbourhood'] = london_venues['Neighbourhood']
```

```
In [27]: # move neighborhood column to the first column
         fixed_columns = [ln_onehot.columns[-1]] + list(ln_onehot.columns[:-1])
         ln_onehot = ln_onehot[fixed_columns]
         ln_onehot.head()
```

Out[27]:

	Neighbourhood	Afghan Restaurant	African Restaurant	Airport	Airport Lounge	Airport Service	American Restaurant	Arepa Restaurant	Argentinian Restaurant	Art Gallery	Art Museum	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	Au: Res
0	Barking and Dagenham [note 1]	0	0	0	0	0	0	0	0	0	0	0	0	0	
1	Barking and Dagenham [note 1]	0	0	0	0	0	0	0	0	0	0	0	0	0	
2	Barking and Dagenham [note 1]	0	0	0	0	0	0	0	0	0	0	0	0	0	
3	Barking and Dagenham [note 1]	0	0	0	0	0	0	0	0	0	0	0	0	0	
4	Barking and Dagenham [note 1]	0	0	0	0	0	0	0	0	0	0	0	0	0	

Now let’s create the new data frame with most common venues

Out[50]:

	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	Barking and Dagenham [note 1]	Grocery Store	Supermarket	Platform	Gas Station	Park	Pool	Racetrack	Bus Stop	Bus Station	Café
1	Barnet	Coffee Shop	Café	Pub	Grocery Store	Park	Pharmacy	Italian Restaurant	Pizza Place	Supermarket	Fast Food Restaurant
2	Bexley	Pub	Supermarket	Clothing Store	Fast Food Restaurant	Coffee Shop	American Restaurant	Hotel	Grocery Store	Italian Restaurant	Pharmacy
3	Brent	Coffee Shop	Indian Restaurant	Hotel	Clothing Store	Grocery Store	Sandwich Place	Pizza Place	Sporting Goods Shop	Bar	Ice Cream Shop
4	Bromley	Pub	Pizza Place	Gym / Fitness Center	Coffee Shop	Clothing Store	Indian Restaurant	Park	Portuguese Restaurant	Donut Shop	Bar

And create the clusters:

In [51]:
```
ln_grouped_clustering = ln_grouped.drop('Neighbourhood', 1)
# set number of clusters
kclusters = 5
# run k-means clustering
kmeans = KMeans(n_clusters = kclusters, random_state=0).fit(ln_grouped_clustering)
# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]
```

Out[51]:
```
array([4, 1, 1, 2, 3, 0, 1, 3, 1, 1])
```

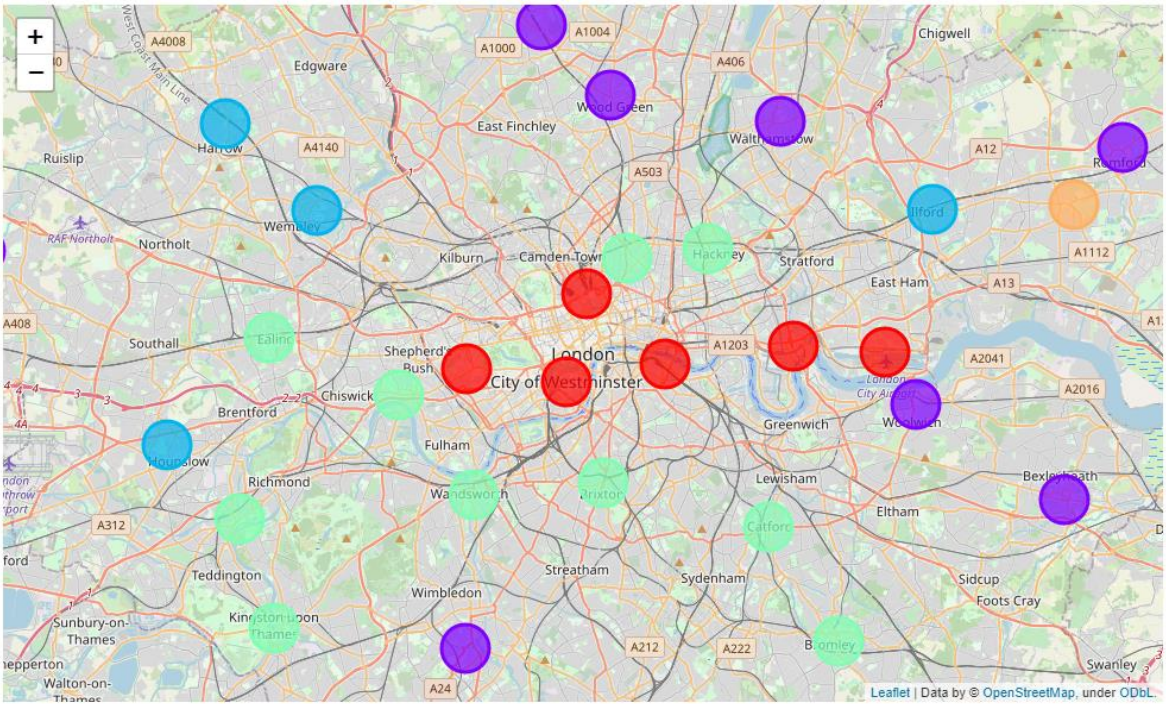
In [52]:
```
# add clustering labels
neighbourhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)
```

In [53]:
```
merged = df
# match/merge SE London data with latitude/longitude for each neighborhood
merged_latlong = merged.join(neighbourhoods_venues_sorted.set_index('Neighbourhood'), on = 'Borough')
merged_latlong.head()
```

Out[53]:

	Borough	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 est) [1]	Co-ordinates	Nr. in map	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th M Comm Ve
1	Barking and Dagenham [note 1]	Barking and Dagenham London Borough Council	Labour	Town Hall, 1 Town Square	13.93	194352	51.5607°N 0.1557°E	25	51.5607	0.1557	4	Grocery Store	Supermarket	Platform	Sta
2	Barnet	Barnet London Borough Council	Conservative	Barnet House, 2 Bristol Avenue, Colindale	33.49	369088	51.6252°N 0.1517°W	31	51.6252	-0.1517	1	Coffee Shop	Café	Pub	Gro S
3	Bexley	Bexley London Borough Council	Conservative	Civic Offices, 2 Watling Street	23.38	236687	51.4549°N 0.1505°E	23	51.4549	0.1505	1	Pub	Supermarket	Clothing Store	Fast F Restau
4	Brent	Brent London Borough Council	Labour	Brent Civic Centre, Engineers Way	16.70	317264	51.5588°N 0.2817°W	12	51.5588	-0.2817	2	Coffee Shop	Indian Restaurant	Hotel	Clot S

Visualizing the Resulting Clusters



Cluster 1:

```
In [55]: # Cluster 1
merged_latlong.loc[merged_latlong['Cluster Labels'] == 0, merged_latlong.columns[[1] + list(range(5, merged_latlong.shape[1]))]]
```

Out[55]:

	Local authority	Population (2013 est) [1]	Co-ordinates	Nr. in map	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 I Com Ve
6	Camden London Borough Council	229719	51.5290°N 0.1255°W	11	51.5290	-0.1255	0	Pizza Place	Coffee Shop	Hotel	Cocktail Bar	Bookstore	Sushi Restaurant	Café	Bee
19	Kensington and Chelsea London Borough Council	155594	51.5020°N 0.1947°W	3	51.5020	-0.1947	0	Pub	Italian Restaurant	Garden	Hotel	Restaurant	Science Museum	Persian Restaurant	Ir Restai
24	Newham London Borough Council	318227	51.5077°N 0.0469°E	27	51.5077	0.0469	0	Hotel	Coffee Shop	Café	Gym / Fitness Center	Grocery Store	Sandwich Place	Pub	Hote
27	Southwark London Borough Council	298464	51.5035°N 0.0804°W	7	51.5035	-0.0804	0	Coffee Shop	Hotel	Italian Restaurant	Cocktail Bar	Scenic Lookout	Gym / Fitness Center	Park	Mus
29	Tower Hamlets London Borough Council	272890	51.5099°N 0.0059°W	8	51.5099	-0.0059	0	Coffee Shop	Hotel	Burger Joint	Pub	Bar	Gym / Fitness Center	Italian Restaurant	F
32	Westminster City Council	226841	51.4973°N 0.1372°W	2	51.4973	-0.1372	0	Hotel	Cocktail Bar	Plaza	Park	Garden	Café	Coffee Shop	Mus

Cluster 2:

In [56]:
```python
# Cluster 2
merged_latlong.loc[merged_latlong['Cluster Labels'] == 1, merged_latlong.columns[[1] + list(range(5, merged_latlong.shape[1]))]]
```

Out[56]:

	Local authority	Population (2013 est) [1]	Co-ordinates	Nr. in map	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 M Comn Ve
2	Barnet London Borough Council	369088	51.6252°N 0.1517°W	31	51.6252	-0.1517	1	Coffee Shop	Café	Pub	Grocery Store	Park	Pharmacy	Ita Restau
3	Bexley London Borough Council	236687	51.4549°N 0.1505°E	23	51.4549	0.1505	1	Pub	Supermarket	Clothing Store	Fast Food Restaurant	Coffee Shop	American Restaurant	H
7	Croydon London Borough Council	372752	51.3714°N 0.0977°W	19	51.3714	-0.0977	1	Pub	Coffee Shop	Clothing Store	Supermarket	Park	Mediterranean Restaurant	Groc S
9	Enfield London Borough Council	320524	51.6538°N 0.0799°W	30	51.6538	-0.0799	1	Pub	Supermarket	Coffee Shop	Pizza Place	Grocery Store	Clothing Store	Pharm
10	Greenwich London Borough Council	264008	51.4892°N 0.0648°E	22	51.4892	0.0648	1	Grocery Store	Pub	Park	Coffee Shop	Plaza	Supermarket	C
13	Haringey London Borough Council	263386	51.6000°N 0.1119°W	29	51.6000	-0.1119	1	Café	Grocery Store	Pub	Park	Mediterranean Restaurant	Turkish Restaurant	Bal
15	Havering London Borough Council	242080	51.5812°N 0.1837°E	24	51.5812	0.1837	1	Coffee Shop	Grocery Store	Fast Food Restaurant	Pub	Supermarket	Clothing Store	Shopp I

Cluster 3:

In [57]:
```python
# Cluster 3
merged_latlong.loc[merged_latlong['Cluster Labels'] == 2, merged_latlong.columns[[1] + list(range(5, merged_latlong.shape[1]))]]
```

Out[57]:

	Local authority	Population (2013 est) [1]	Co-ordinates	Nr. in map	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 Com V
4	Brent London Borough Council	317264	51.5588°N 0.2817°W	12	51.5588	-0.2817	2	Coffee Shop	Indian Restaurant	Hotel	Clothing Store	Grocery Store	Sandwich Place	Pizza Place	Spo G
14	Harrow London Borough Council	243372	51.5898°N 0.3346°W	32	51.5898	-0.3346	2	Coffee Shop	Indian Restaurant	Pub	Grocery Store	Fast Food Restaurant	Park	Café	Sanc I
17	Hounslow London Borough Council	262407	51.4746°N 0.3680°W	14	51.4746	-0.3680	2	Indian Restaurant	Coffee Shop	Bus Stop	Clothing Store	Metro Station	Grocery Store	Sandwich Place	Fast Resta
25	Redbridge London Borough Council	288272	51.5590°N 0.0741°E	26	51.5590	0.0741	2	Grocery Store	Supermarket	Coffee Shop	Hotel	Clothing Store	Fast Food Restaurant	Department Store	Sanc I

Cluster 4:

```
In [58]: # Cluster 4
merged_latlong.loc[merged_latlong['Cluster Labels'] == 3, merged_latlong.columns[[1] + list(range(5, merged_latlong.shape[1]))]]
```

Out[58]:

Co-nates	Nr. in map	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	9th Most Common Venue	10th Most Common Venue
39°N 198°E	20	51.4039	0.0198	3	Pub	Pizza Place	Gym / Fitness Center	Coffee Shop	Clothing Store	Indian Restaurant	Park	Portuguese Restaurant	Donut Shop	Bar
30°N 89°W	13	51.5130	-0.3089	3	Pub	Coffee Shop	Park	Hotel	Italian Restaurant	Pizza Place	Café	Burger Joint	Vietnamese Restaurant	Greek Restaurant
50°N 53°W	9	51.5450	-0.0553	3	Pub	Coffee Shop	Café	Bakery	Cocktail Bar	Park	Wine Shop	Brewery	Supermarket	Garden
27°N 39°W	4	51.4927	-0.2339	3	Pub	Café	Coffee Shop	Indian Restaurant	Park	Gastropub	Japanese Restaurant	Pizza Place	French Restaurant	Thai Restaurant
16°N 22°W	10	51.5416	-0.1022	3	Pub	Café	Park	Gastropub	Theater	Canal	Trail	Coffee Shop	Gym / Fitness Center	Mediterranean Restaurant
85°N 64°W	16	51.4085	-0.3064	3	Pub	Café	Coffee Shop	Thai Restaurant	Burger Joint	Gastropub	Italian Restaurant	Department Store	Park	Japanese Restaurant
07°N 63°W	6	51.4607	-0.1163	3	Coffee Shop	Pub	Cocktail Bar	Market	Café	Park	Pizza Place	Restaurant	Brewery	Caribbean Restaurant
52°N 09°W	21	51.4452	-0.0209	3	Pub	Coffee Shop	Park	Café	Turkish Restaurant	Supermarket	Concert Hall	Gourmet Shop	Deli / Bodega	Cocktail Bar

Cluster 5:

```
In [59]: # Cluster 5
merged_latlong.loc[merged_latlong['Cluster Labels'] == 4, merged_latlong.columns[[1] + list(range(5, merged_latlong.shape[1]))]]
```

Out[59]:

	Local authority	Population (2013 est) [1]	Co-ordinates	Nr. in map	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
1	Barking and Dagenham London Borough Council	194352	51.5607°N 0.1557°E	25	51.5607	0.1557	4	Grocery Store	Supermarket	Platform	Gas Station	Park	Pool	Racetrack	Bus Stop

Results

Results The following are the highlights of the 5 clusters above:

Pubs, Cafe, Coffee Shops are popular in the South East London.

As for restaurants, for example, the Indian Restaurants are very popular in Brent and Harrow areas.

Although, the Clusters have variations, a very visible presence is the predominance of pubs. (That we also could see on data analysis section)