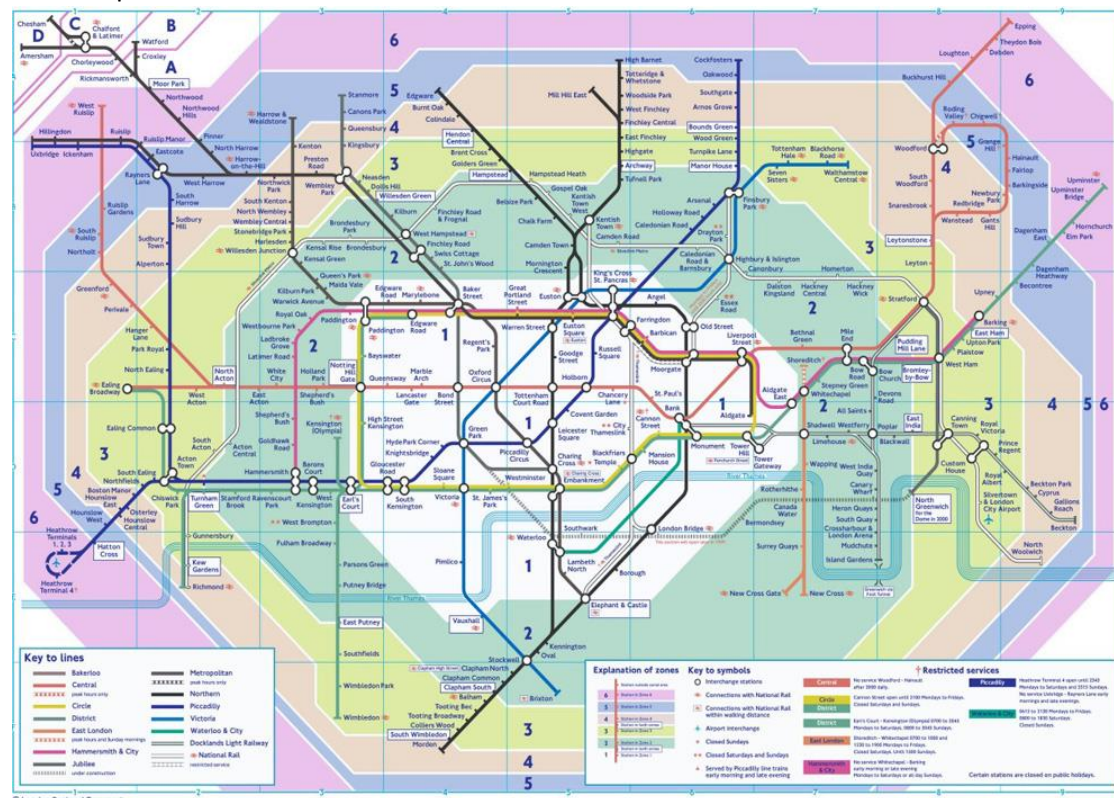


Comparing Neighbourhoods in London

Project description:

I recently made a big decision in my life - Moving to one of the biggest cities in the world - LONDON! London has proven to be all I dreamed of and more, but comparing it to the small place I used to call home, there is lots more to learn.

To give you an idea of the overwhelming feeling I had, below you will see an image of the London Tube maps and zones.



The zones work as follows: Central London is Zone 1 (the white area in the image), further out is Zone 2 (green), etc.

The more central you are, the busier London is. Tourists everywhere, many things to do and LOTS of shopping!

In my journey so far, I have discovered 4 areas that I really like:

- Putney
- Chancery Lane
- Mayfair
- Hammersmith

Chancery Lane and Mayfair are in the centre of Zone 1, Hammersmith is more West (zone 2) and Putney is more South West (border zone 3).

The objective of this project is to see if these 4 neighbourhoods are similar to each other and what venues I would be able to find in each of these areas. This is where Foursquare data will be used. I will be finding the 100 closest venues to each of the 4 London locations:

Name	Longitude	Latitude
PUTNEY	-0.21547	51.46456
CHANCERY LANE	-0.11214	51.51802
MAYFAIR	-0.14271	51.50745
HAMMERSMITH	-0.22814	51.49244

I will do a k-means test to see if some of the neighbourhoods are similar to one another. The overall objective is just to explore these areas and perhaps decide which one is my favourite!

Data description:

The objective of this project is to see if these 4 neighbourhoods are similar to each other and what venues I would be able to find in each of these areas. Foursquare data is used to find the 100 closest venues to each of the 4 London locations.

After extracting Foursquare data for the 4 areas, the neighbourhood data is combined and a list of 400 points is formed from the data.

Below is an extract of what the Foursquare data looks like:

Unnamed: 0		name	categories	lat	lng
0	0	Catalyst Cafe	Coffee Shop	51.519705	-0.112052
1	1	Scarves Bar	Hotel Bar	51.517813	-0.118184
2	2	Rosewood London	Hotel	51.517468	-0.117810
3	3	Apex Temple Court Hotel	Hotel	51.513738	-0.108878
4	4	The Hoxton Holborn	Hotel	51.517145	-0.122030

For each point, the area is added to the list of 400 entries as per the following example:

Unnamed: 0		name	categories	lat	lng	area
95	95	The Ginger Pig	Butcher	51.454584	-0.139801	Putney
96	96	BVLGARI Hotel London	Hotel	51.501467	-0.163548	Putney
97	97	Shirley Sherwood Gallery	Art Gallery	51.474788	-0.292470	Putney
98	98	L'ETO Caffè	Café	51.498651	-0.156969	Putney
99	99	Mari Vanna	Russian Restaurant	51.501793	-0.162479	Putney

The 400 points are then ranked by most common categories. In this case, Hotels are the most common venue, followed by Parks and then Pizza places. With Pizza places in the top 3, no wonder these areas are my favourite!!

Theatres, Gardens and Café's also appear high up on the list. In total, there are 94 unique categories of locations.

Unnamed: 0

categories	
Hotel	38
Park	24
Pizza Place	14
Theater	13
Garden	11
Café	11
Indian Restaurant	10
Coffee Shop	10
Cocktail Bar	10
Bookstore	9
Steakhouse	9
Gym / Fitness Center	8
Department Store	8
Ice Cream Shop	8
Art Museum	8
Japanese Restaurant	7
Movie Theater	7
Plaza	7
Art Gallery	7
Botanical Garden	6

Then the data is divided into the most common categories in each area:

	area	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	Chancery Lane	Hotel	Theater	Coffee Shop	Steakhouse	Cocktail Bar	Indian Restaurant	Sushi Restaurant	Pizza Place	Bakery	Art Museum
1	Hammersmith	Hotel	Park	Botanical Garden	Garden	Pizza Place	Japanese Restaurant	Breakfast Spot	Café	Ice Cream Shop	Science Museum
2	Mayfair	Hotel	Theater	Bookstore	Park	Steakhouse	Art Museum	Lounge	Pizza Place	Cocktail Bar	Department Store
3	Putney	Park	Café	Pizza Place	Coffee Shop	Japanese Restaurant	Garden	Hotel	Ice Cream Shop	Museum	Art Gallery

It can be seen that Chancery Lane, Hammersmith and Mayfair all have Hotels as the number 1 most common venue, but Putney only has this as the 7th most common venue.

Here we can already see that Putney is not too similar to the other neighbourhoods.

Hammersmith and Putney seem to have Parks quite high up on common venues, so this could be giving an indication of them being similar.

Further investigations will include plotting each area's points on a map, plotting the most common venues on a map, doing k-means to see if the 4 areas would fall into 2 separate categories and looking at a few box plots to further investigate similarities or differences.

Methodology

Before doing the actual k-means test, it is great to get an idea of where all of the data points are on the map.

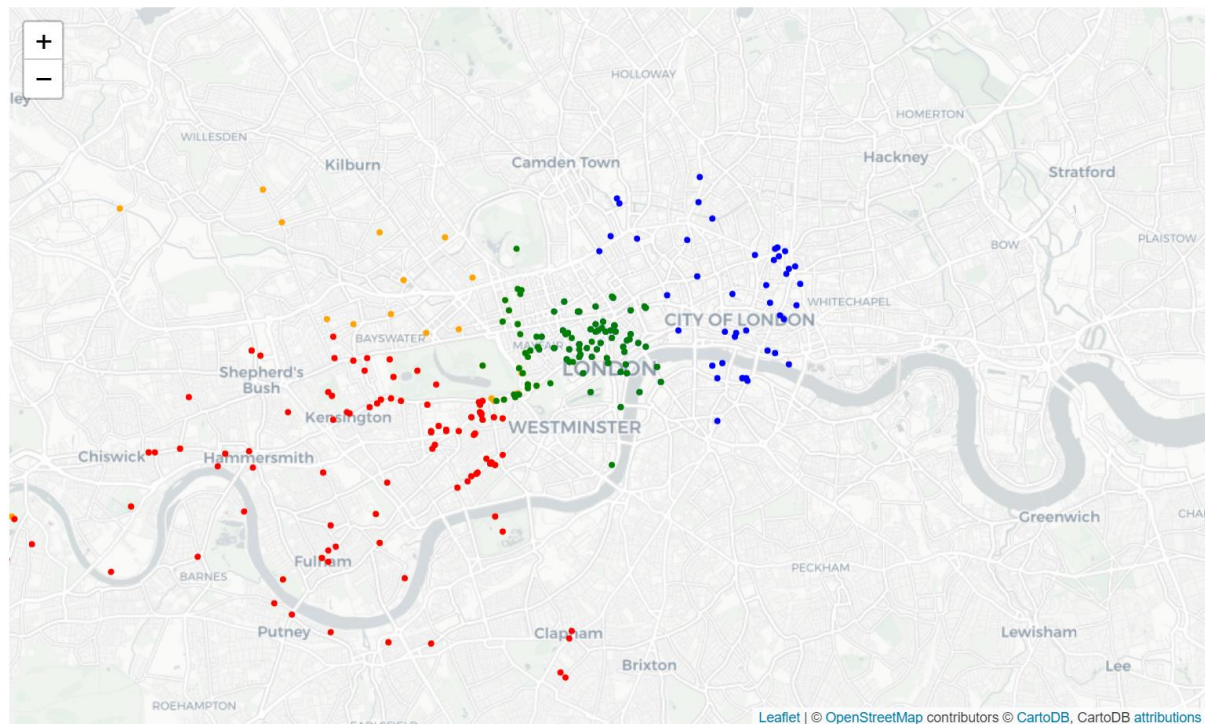
Below is a map of the 4 areas with their 100 Foursquare data points plotted in the following colours:

Chancery Lane

Mayfair

Putney

Hammersmith



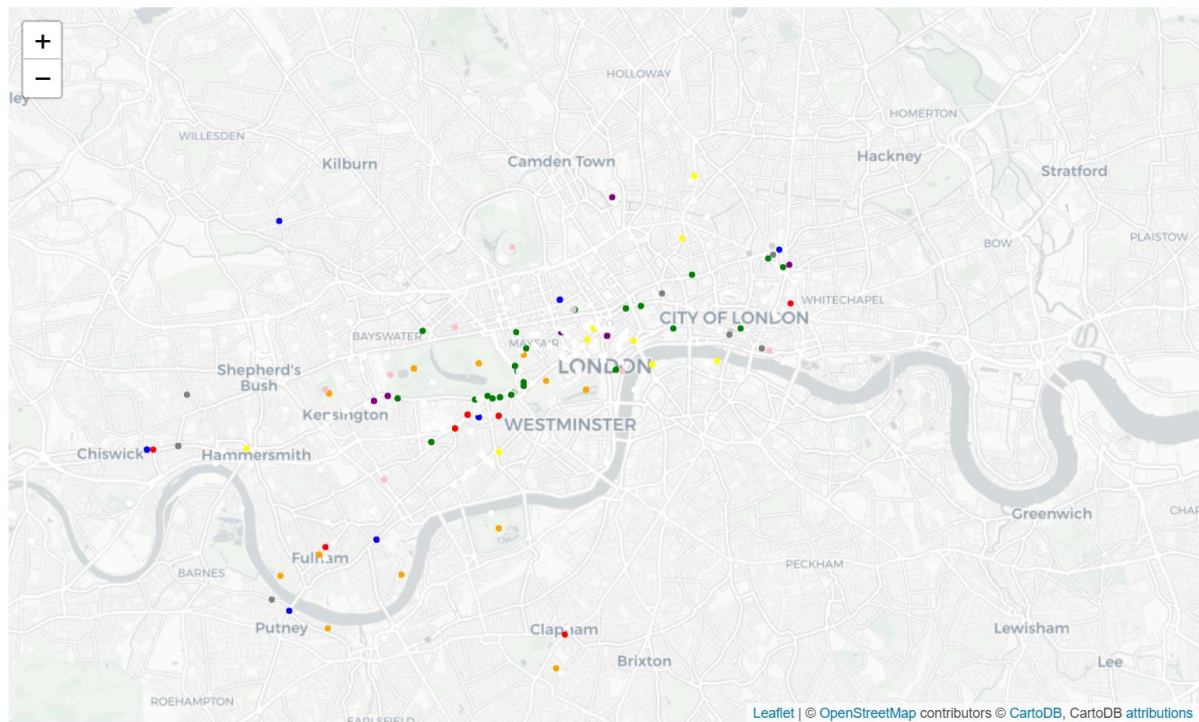
As mentioned before, Chancery Lane and Mayfair are in the centre of Zone 1, Hammersmith is more West (zone 2) and Putney is more South West (border zone 3).

From the data so far, we have already seen that Hammersmith and Putney have more parks. Whereas central London areas like Chancery Lane and Mayfair have more tourists and have more Hotels and Theatres. It is clear to see on the map that these two areas have their venues clustered more closely to one another.

Then also looking at the most common venues, which as we have seen before are as follows:

Unnamed: 0

	categories	
Hotel	Hotel	38
Park	Park	24
Pizza Place	Pizza Place	14
Theater	Theater	13
Garden	Garden	11
Café	Café	11
Indian Restaurant	Indian Restaurant	10
Coffee Shop	Coffee Shop	10
Cocktail Bar	Cocktail Bar	10



Thereafter, a k-means test was applied to the 4 neighbourhoods. It was decided to use 2 clusters for k-means to see which of the 4 neighbourhoods are more similar. The k-means test was done by dividing each of the neighbourhoods into the venues that are most common in the area. The input for the clustering is as follows as already seen in the data preparation section:

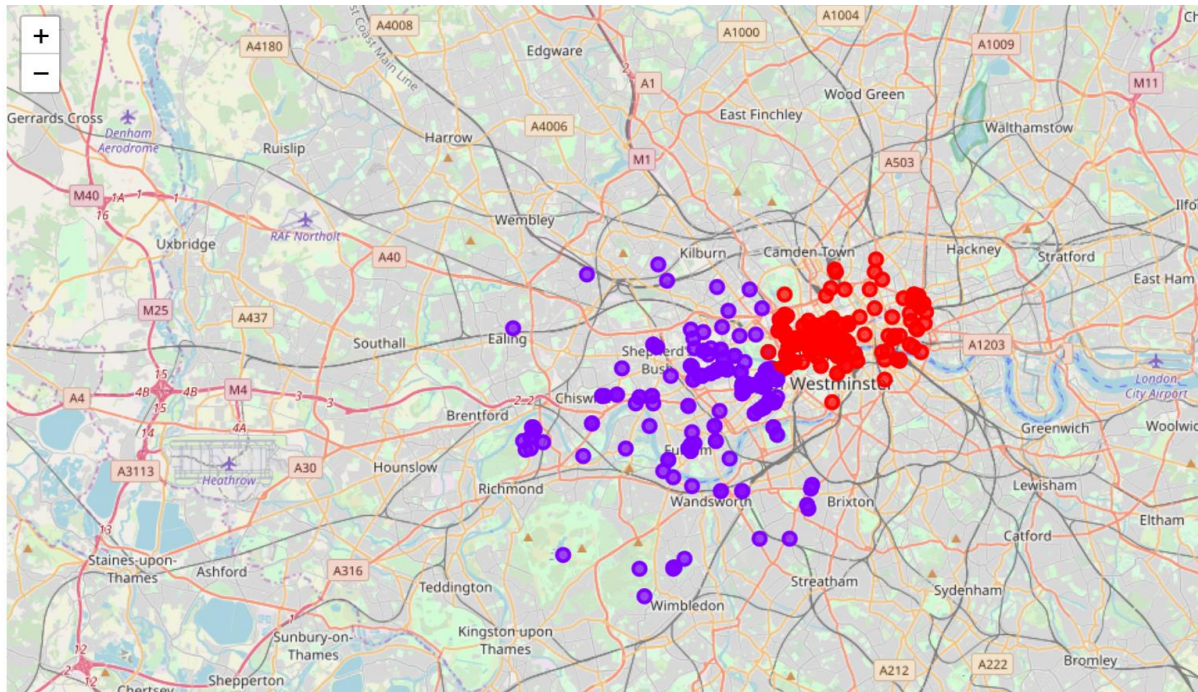
	area	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Chancery Lane	Hotel	Theater	Coffee Shop	Steakhouse	Cocktail Bar	Indian Restaurant	Sushi Restaurant	Pizza Place	Bakery	Art Museum
1	Hammersmith	Hotel	Park	Botanical Garden	Garden	Pizza Place	Japanese Restaurant	Breakfast Spot	Café	Ice Cream Shop	Science Museum
2	Mayfair	Hotel	Theater	Bookstore	Park	Steakhouse	Art Museum	Lounge	Pizza Place	Cocktail Bar	Department Store
3	Putney	Park	Café	Pizza Place	Coffee Shop	Japanese Restaurant	Garden	Hotel	Ice Cream Shop	Museum	Art Gallery

Results

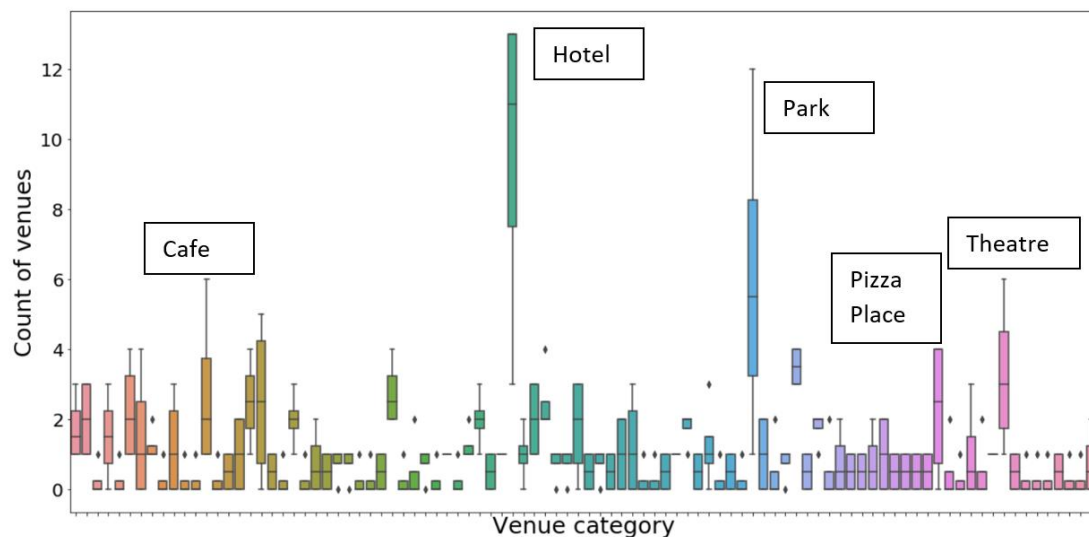
Results from the k-means test showed that Chancery Lane and Mayfair are clustered together and that Putney and Hammersmith are clustered together.

This was an early suspicion as the similar venues for each cluster as well as the physical location of the two clusters had clear differences.

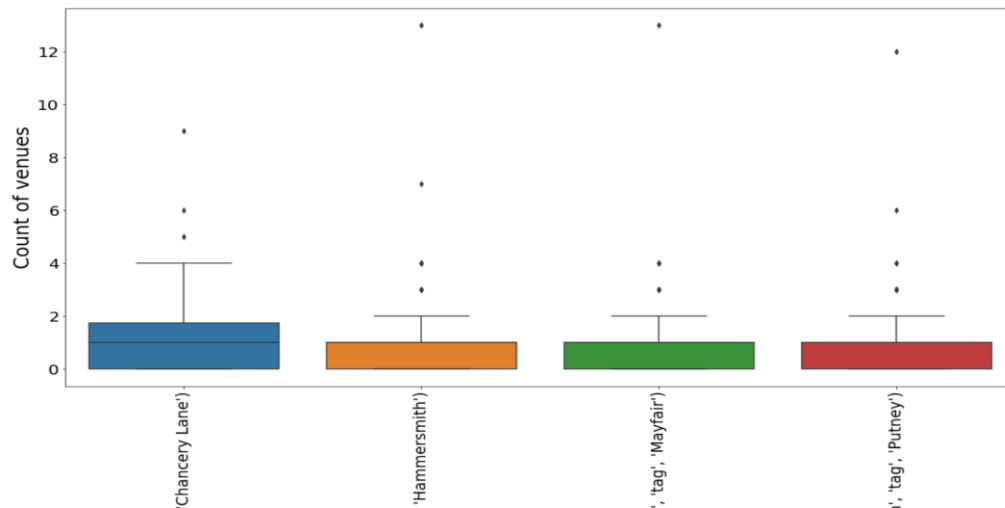
The clusters can now be visualised on a map as follows:



Thereafter some box plots had confirmed some results. Hotels, Parks, Pizza Places, Theatres and Cafes were by far the most common venues.



Though the 4 venues did indeed cluster into 2 parts, looking at overall totals number of venues in each area irrespective of which venue it is, the 4 venues were infact very similar:



Observations and Recommendations from results

MAPS:

Chancery Lane and Mayfair are zone 1 and more central.

Hammersmith and Putney are zones 2 and 3 respectively and are more west out of central London.

The maps showed that the zone 1 neighbourhoods had venues more closely situated to the points of interest. The other two neighbourhoods had venues that were more spaced out.

We also saw that Hotels, Parks and Pizza places are the most common venues over all 4 neighbourhoods.

K-MEANS:

The preparation for k-means showed us the most common venues for each of the neighbourhoods.

The two zone 1 neighbourhoods were clustered together, having similar venues.

BOX PLOTS:

We can clearly see that Hotels, Parks and Pizza places are most common venues. Though there were 2 clusters, the 4 areas had similar distributions of total types of venues.

Conclusions

As one could have suspected, the two central London neighbourhoods are more similar than the outskirts.

Looking at the two clusters and the most common venues, I would prefer the outskirts due to them having more parks and perhaps a bit of extra space to wander around the areas! Tourists would prefer central London and there are many Hotels, Theatres and Food options to choose from.

If ever you visit London, I hope you would have a wonderful time in any of these neighbourhoods :)