COMPARISON OF LONDON AND BERLIN

Geographical study considering crime, population and police station location statistics

Abstract

The demographics of London and Berlin show that both are large and ethnically diverse metropoles.
We are aiming to compare the neighbourhoods of the two cities and determine how similar or dissimilar they are, taking crime statistics, population numbers and the geographical location of police stations into account, to consider whether to start a business in one of the two European capitals.

Crime Comparison Berlin and London

Introduction:

The Capstone analysis project focuses on the comparison of the two cities of London and Berlin. Both cities are very diverse and are the capitals of their respective countries. The demographics of London and Berlin show that both are large and ethnically diverse metropoles. London was home to 9,4 million people in 2021, with the population increasing by 1.31% compared to the previous year. While Berlin had a population of 3,56 million in 2021 up by 0.14% in contrast to the previous year.

We are aiming to compare the neighbourhoods of the two cities and determine how similar or dissimilar they are, taking crime statistics, population numbers and the geographical location of police stations into account. I will aim to explore in how far the two cities vary in the distribution of these demographic and geographical data, to determine whether one city could be considered safer than the other, or whether particular areas within these cities could be considered more safe than others.

Data Sources: The following datasets was used for this project, using a combination of methods including web scrapping, foursquare api data and reading of csv. files.

London Data:

[1]: "Crime rates in London by Borough"

[2]: "London Boroughs & Population data from Wikipedia"

[3]: "Foursquare London Location data for Police Stations"

Berlin Data:

[1]: "Crime rates in Berlin by Borough"

[2]: "Berlin Borough & Population data from Wikipedia"

[3]: "Foursquare Berlin Location data for Police Stations"

Methodology:

A variety of techniques learned in this course were used as part of the methodology to assess and analyse the available data sets. The data was cleaned and grouped to perform meaningful analysis. Foursquare location data was used to determine the spatial distribution of police stations in conjunction with crime data by Borough for the respective cities. We also further explored population density in the city boroughs of Berlin and London to see whether there was a relationship between population numbers and crime rates.

Problem Statement:

To compare the two cities of London and Berlin in terms of crime and demographic data to assess whether one city could be considered safer than the other, to inform a business which of the two capitals is better for relocation.

Discussion of results:

The statistical data for Berlin was acquired from a reliable source, the Berlin Police Department under the following link https://www.berlin.de/polizei/verschiedenes/polizeiliche-kriminalstatistik/.

As this is a reliable source in key terms for statistic labelling in German are translated into English, to make the charts accessible. The following listed data table 1. shows the total crime occurrences (Straftaten) in Berlin by borough (Bezirksregion).

Berlin Data table 1. Crime statistics by Borough for 2019 over 12 months

t[41]:

	LOR-Schlüssel (Bezirksregion)	Bezeichnung (Bezirksregion)	Straftaten \n-insgesamt-	Raub	Straßenraub,\nHandtaschen-raub	Körper-verletzungen \n-insgesamt-	Gefährl. und schwere Körper-verletzung	Freiheits-beraubung, Nötigung,\nBedrohung, Nachstellung	Diebstahl \n-insgesamt-	Diebstahl von Kraftwagen	Diebstahl \nan/aus Kfz	Fahrrad-\ndiebstahl	Wohnraum-\neinbruch	Branddelikte \n-insgesamt-	Brand-\nstiftung	Sach-beschädigung -insgesamt-	Sach-beschädigung durch Graffiti	Rauschgift-delikte	Kieztaten
0	10000	Mitte	84357	707	407	7595	1951	2157	35601	401	3330	3817	845	291	104	6142	1601	4233	15967
1	10111	Tiergarten Süd	5009	60	35	365	92	128	2271	15	198	296	55	13	6	347	77	231	980
2	10112	Regierungsviertel	7950	42	20	554	136	152	3692	13	172	352	22	19	4	497	162	170	1057
3	10113	Alexanderplatz	22974	173	102	1966	500	420	11233	63	587	940	137	43	12	1307	381	1133	3813
4	10114	Brunnenstraße Süd	4252	40	29	268	64	79	1859	39	182	361	64	18	7	424	172	86	902

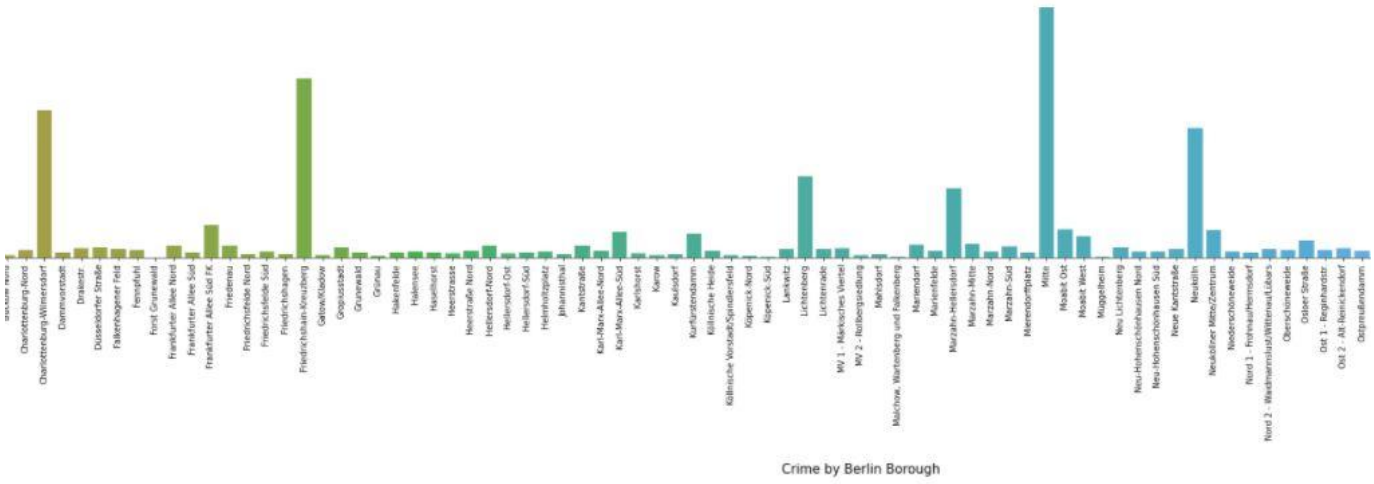
In order, to get a better view of the occurrences of crimes per month and borough, monthly averages were calculated to evaluate the crime occurrences over 12 months.

Berlin Data table 2. ordered by monthly average of crime occurrences per borough (Bezirk)

	Bezeichnung (Bezirksregion)	LOR-Schlüssel (Bezirksregion)	MonthlyAverage	Raub	Straßenraub,\nHandtaschen-raub	Körper-verletzungen \n-insgesamt-	...und schwere Körper-verletzung	Freiheits-beraubung, Nötigung,\nBedrohung, Nachstellung	Diebstahl \n-insgesamt-	Diebstahl von Kraftwagen	Diebstahl \nan/aus Kfz	Fahrrad-\ndiebstahl	Wohnraum-\neinbruch	Branddelikte \n-insgesamt-	Brand-\nstiftung	Sach-beschädigung -insgesamt-	Sach-beschädigung durch Graffiti	Rauschgift-delikte	Kieztate
0	Adlershof	90207	180.583333	14	9	171	36	78	983	48	140	240	23	9	6	240	52	28	39
1	Albrechtstr.	60102	284.833333	35	23	274	49	128	1486	79	308	271	142	31	12	426	90	59	84
2	Alexanderplatz	10113	1914.500000	173	102	1966	500	420	11233	63	587	940	137	43	12	1307	381	1133	381
3	Allende-Viertel	90414	64.000000	1	0	75	17	20	355	20	53	87	5	8	5	94	15	18	17
4	Alt-Hohenschönhausen Nord	110204	156.916667	17	9	197	41	64	891	35	136	62	35	7	6	173	13	26	46

When visualising the data in a bar chart we can see that there are significant differences per area with Mitte, Friedrichshain and Charlottenburg, being the areas with most crime averages recorded per month. It is noticeable that all boroughs are relatively central places. The high numbers might also be due to how the boundaries of these areas are defined in the police records and the larger numbers of population per km2. These factors should be further explored and taken into consideration in the future analysis of the dataset.

Berlin Data table 3. Illustrated in bar chart crime rates per borough



When looking at the population per Km2 in Berlin we can see that the top areas ranked in terms of crime rates are also the two most populated.

Berlin Data table 4. Population per km2 per borough

	Einwohner pro km²	NR	Bezirk	Ortsteile	Fläche in km²	Einwohner	Bezirksbürgermeister	Karte
11	14.404\n	2\n	Friedrichshain-Kreuzberg Friedrichshain-Kreuzb...	0201 Friedrichshain0202 Kreuzberg\n	020,16\n	290.386\n	Monika Herrmann (Grüne)\n	Lage52.513.433333333333\n
10	09.773\n	1\n	Mitte Mitte\n	0101 Mitte0102 Moabit0103 Hansaviertel0104 Tie...	039,47\n	385.748\n	Stephan von Dassel (Grüne)\n	Lage52.51666666666713.366666666667\n
9	07.343\n	8\n	Neukölln Neukölln\n	0801 Neukölln0802 Britz0803 Buckow0804 Rudow08...	044,93\n	329.917\n	Martin Hikel (SPD)\n	Lage52.48333333333313.45\n
8	06.611\n	7\n	Tempelhof-Schöneberg Tempelhof-Schöneberg\n	0701 Schöneberg0702 Friedenau0703 Tempelhof070...	053,09\n	350.984\n	Angelika Schöttler (SPD)\n	Lage52.46666666666713.383333333333\n
7	05.626\n	11\n	Lichtenberg Lichtenberg\n	1101 Friedrichsfelde1102 Karlshorst1103 Lichte...	052,29\n	294.201\n	Michael Grunst (Die Linke)\n	Lage52.53333333333313.5\n

As we can see in this table the three most populated areas also came out on top for monthly crime rates, indicating that there might be a link between population density and crime rates, possibly connected to crowded urban living. These relationships should be further explored via a heatmap to consider whether there are linked relationships between the two factors.

Foursquare API for Berlin Police Stations:

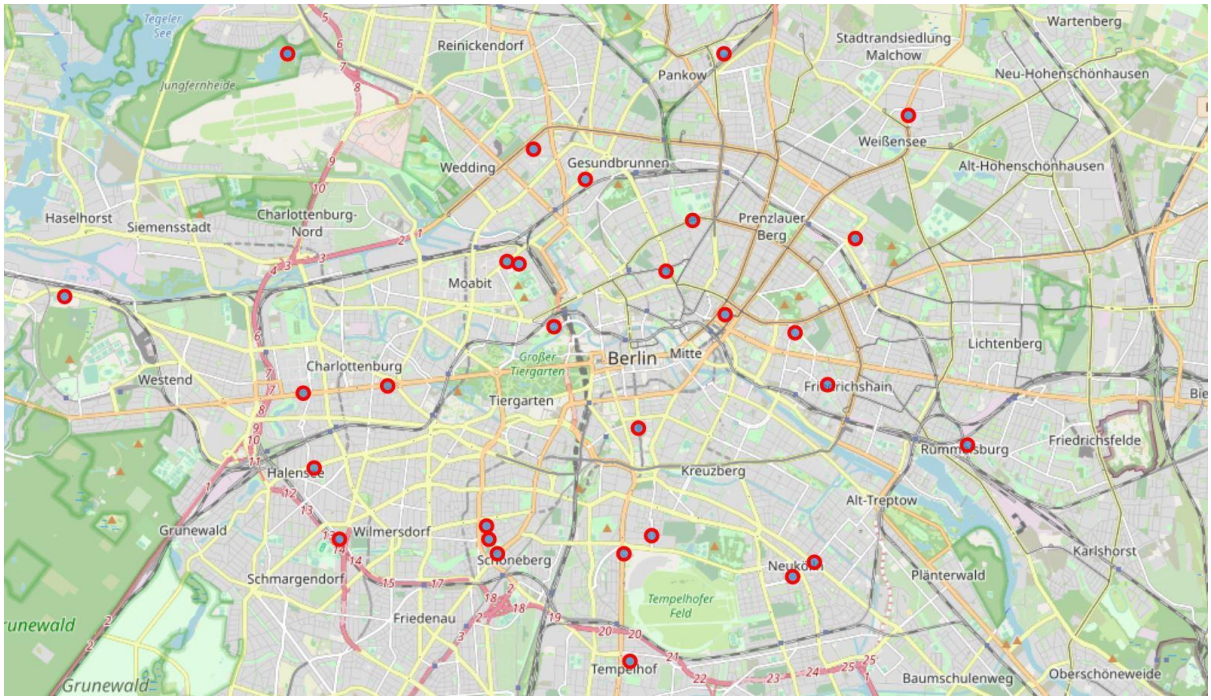
To research whether the two cities are comparable in terms of safety and police presence the foursquare API was used to establish the number and location of central police stations in both cities. In Berlin data table 3 we can see the results of using the generated token and api connection, giving us the Berlin police stations (Polizei) and location data for Berlin.

Berlin data table 5. List of berlin police stations and location acquired from the foursquare Api

	id	name	categories	referralId	hasPerk	location.address	location.lat	location.lng	location.labeledLatLngs	location.distance	location.postalCode	location.cc	location.city
0	4e14de40d164535f06168e09	Polizei Abschnitt 33	[{'id': '4bf58dd8d48988d12e941735', 'name': 'P...	v-1612885136	False	Perleberger Str. 61A	52.533865	13.353253	[{'label': 'display', 'lat': 52.53386480694763...	3053	10559	DE	Berlin
1	4be112870365c9b66fe2b238	Polizei Abschnitt 13	[{'id': '4bf58dd8d48988d12e941735', 'name': 'P...	v-1612885136	False	Hadlichstr. 37-41	52.569650	13.415130	[{'label': 'display', 'lat': 52.56964975511158...	6120	13187	DE	Berlin
2	4e7b0bbcfa76e8ee773f282b	Polizei BGSt 212, Kfz-Sicherstellung	[{'id': '4bf58dd8d48988d12e941735', 'name': 'P...	v-1612885136	False	Belziger Str. 52	52.485658	13.348148	[{'label': 'display', 'lat': 52.48565775087338...	4451	10823	DE	Berlin
3	53622cbd498ef3775e628f9b	Polizei Abschnitt 64	[{'id': '4bf58dd8d48988d12e941735', 'name': 'P...	v-1612885136	False	Nöldnerstr. 35	52.502077	13.484323	[{'label': 'display', 'lat': 52.50207700898370...	6678	10317	DE	Berlin
4	55045bc0498ef432ad294ac6	Polizei Abschnitt 53	[{'id': '4bf58dd8d48988d12e941735', 'name': 'P...	v-1612885136	False	Friedrichstraße 219	52.504960	13.390747	[{'label': 'display', 'lat': 52.50495971398627...	1350	10969	DE	Berlin

When plotting the police stations in central Berlin we can see a well distributed urban presence of police in crime spots such as Mitte and Friedriechshain. However, also areas with low crime rates such as Schoeneberg have a good distribution of police location presence.

Visual Map of police locations in Berlin



London Data:

For comparison with Berlin the geolocation data of London was taken into account. The following listed data table 1. shows the crime occurrences in London by borough. The data was acquired via the UK government London Datastore site (https://data.london.gov.uk/). The available data set provided a larger date range view than Berlin, over the last year as well as 2019.

London Data Table1 Crime statistics by Borough for 2019/2020 over 24 months

```
# Read crime records data
crime = pd.read_csv("https://data.london.gov.uk/download/recorded_crime_summary/d2e9ccfc-a054-41e3-89fb-53c2bc3ed87a/MPS%20Borough%20Level%20Crime%20%28most%20recent%2024%20months%29.csv")
crime.head()
```

	MajorText	MinorText	LookUp_BoroughName	201901	201902	201903	201904	201905	201906	201907	...	202003	202004	202005	202006	202007	202008	202009	202010	202011	202012
0	Arson and Criminal Damage	Arson	Barking and Dagenham	5	2	5	5	11	3	5	...	6	2	2	4	4	6	2	7	4	2
1	Arson and Criminal Damage	Criminal Damage	Barking and Dagenham	97	127	138	130	140	113	134	...	107	80	86	121	122	114	116	119	100	108
2	Burglary	Burglary - Business and Community	Barking and Dagenham	45	24	29	27	21	27	31	...	28	29	16	16	28	24	32	21	19	24
3	Burglary	Burglary - Residential	Barking and Dagenham	114	107	99	96	114	96	71	...	97	57	42	63	72	63	54	68	90	91
4	Drug Offences	Drug Trafficking	Barking and Dagenham	6	2	6	5	9	6	11	...	6	15	15	12	21	9	11	14	17	14

5 rows × 27 columns

In order, to be able to compare the data set with the smaller 12 months data set available for Berlin, monthly averages were calculated dividing by 24 months.

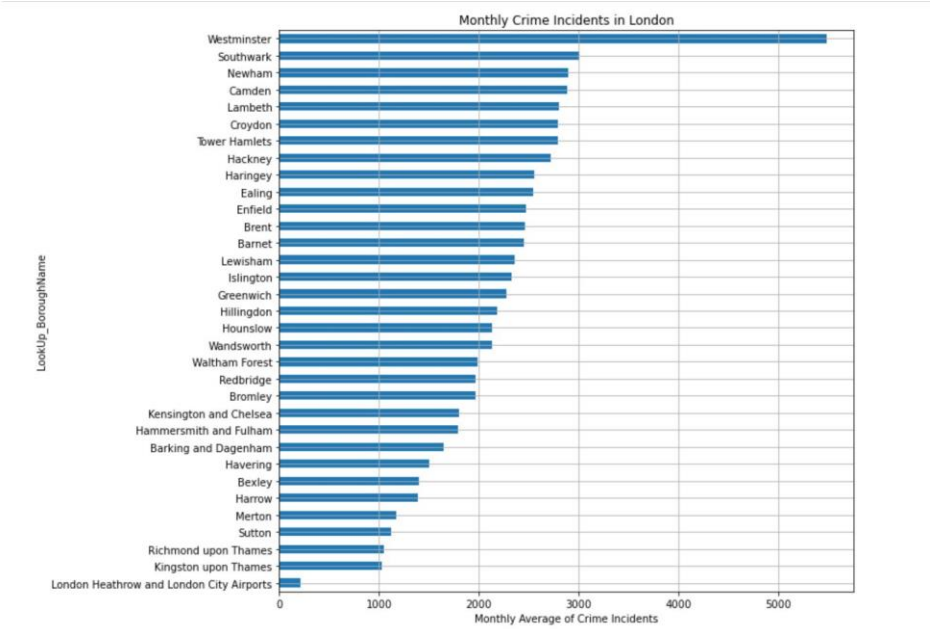
London Data Table 2 Crime statistics by Borough for 2019/2020 ranked by Monthly average per borough

```
crime.sort_values(by='MonthlyAverage', ascending=False).head()
```

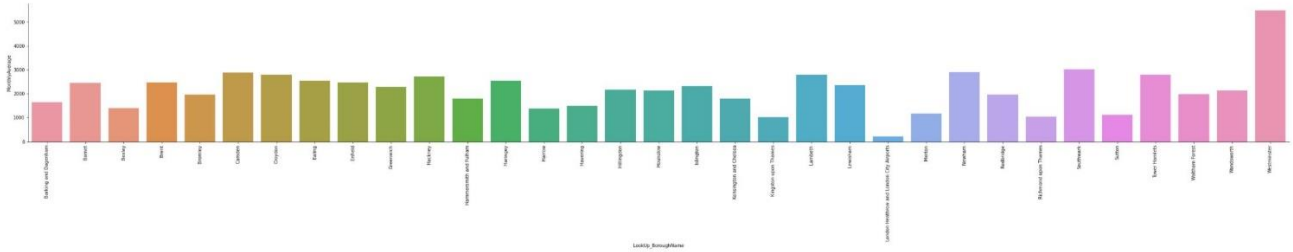
	LookUp_BoroughName	MonthlyAverage	201901	201902	201903	201904	201905	201906	201907	201908	...	202003	202004	202005	202006	202007	202008	202009	202010	202011	202012
32	Westminster	5485.500000	6280	6509	7220	7225	6852	6925	6991	7085	...	4695	1706	2261	2989	4100	4486	4469	4332	3007	3509
27	Southwark	3010.333333	3064	3013	3516	3259	3439	3342	3534	3358	...	2447	1842	2229	2544	2989	3154	2993	2977	2607	2378
24	Newham	2897.708333	2831	2890	3036	2914	3102	3159	3342	2984	...	2593	2071	2329	2692	2984	2981	2893	3092	2785	2877
5	Camden	2890.166667	3061	3174	3459	3235	3293	3163	3458	3210	...	2717	1636	1863	2082	2434	2578	2602	2717	2258	1970
20	Lambeth	2799.458333	2782	2771	3180	2929	3075	3080	3280	3134	...	2512	1889	2267	2493	2890	2961	2875	2689	2535	2192

5 rows × 26 columns

London Data table 3. Illustrated in bar chart crime rates per borough



London Data table 4. Illustrated in bar chart crime rates per borough



London Data table 5. Population density per borough

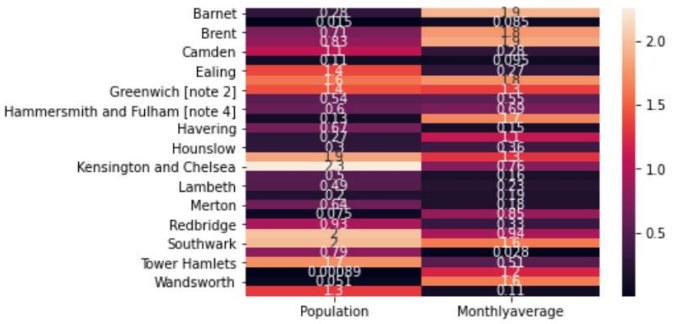
We then looked at the top most populated boroughs in London. A difference in the dataset available is that in this instance we have the total population data rather than per square meter numbers, which would give a more accurate comparison and indication of the density of population in a certain area.

:

	BoroughName	Population	Latitude	Longitude	MonthlyAverage	201901	201902	201903	201904	201905	...	202003	202004	202005	202006
5	Croydon	386710	51.3714	-0.0977	2797.208333	2642.0	2572.0	2834.0	2620.0	2800.0	...	2991.0	2248.0	2505.0	2726.0
21	Newham	353134	51.5077	0.0469	2897.708333	2831.0	2890.0	3036.0	2914.0	3102.0	...	2593.0	2071.0	2329.0	2692.0
6	Ealing	341806	51.5130	-0.3089	2547.750000	2550.0	2464.0	2915.0	2508.0	2611.0	...	2366.0	1991.0	2281.0	2324.0
7	Enfield	333794	51.6538	-0.0799	2471.916667	2509.0	2346.0	2604.0	2413.0	2454.0	...	2367.0	1860.0	2206.0	2345.0
3	Bromley	332336	51.4039	0.0198	1965.708333	2000.0	1956.0	2109.0	1968.0	2115.0	...	1767.0	1443.0	1645.0	1694.0

When using an indicative heatmap, to determine whether population density and average monthly crime rates could be linked, no clear connection could be seen. This could be further explored with other available datasets covering population numbers per square metre for instance,

London Data Visualisation Indicative heatmap of population density and Monthly crime average

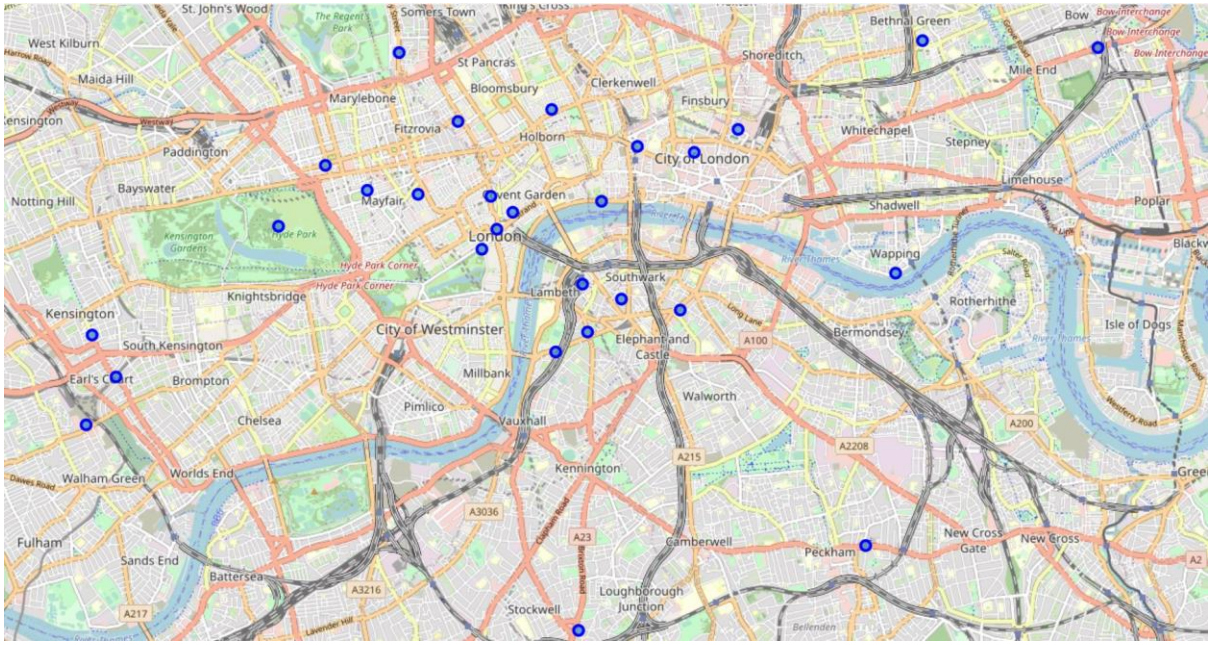


Foursquare API for London Police Stations:

The foursquare API was used to establish the number and location of central police stations in London.

	id	name	categories	referralId	hasPerk	location.address	location.lat	location.lng	location.labeledLatLngs	location.distance	location.postalCode	location.cc	location.city	location.state
0	4d344e67306160fcfe2e6b88	Trafalgar Square Police Box	[{'id': '4bf58dd8d48988d12e941735', 'name': 'P...	v-1612955003	False	Trafalgar Sq	51.507966	-0.127340	[{'label': 'display', 'lat': 51.50796613815064...	74	WC2N 5DN	GB	London	Greater London
1	52e92a52498e8f6d4ed62bfc	Metropolitan Police Central Communications Com...	[{'id': '4bf58dd8d48988d130941735', 'name': 'B...	v-1612955003	False	109 Lambeth Rd	51.494799	-0.117198	[{'label': 'display', 'lat': 51.494799, 'lng':...	1570	SE1 7JP	GB	London	Greater London
2	4b8592e9f964a5204e6631e3	Charing Cross Police Station	[{'id': '4bf58dd8d48988d12e941735', 'name': 'P...	v-1612955003	False	Agar St	51.509859	-0.124550	[{'label': 'display', 'lat': 51.50985902748898...	354	WC2N 4HN	GB	London	Greater London
3	4cc5c23d01fb236add87b3ba	Snow Hill Police Station	[{'id': '4bf58dd8d48988d12e941735', 'name': 'P...	v-1612955003	False	5 Snow Hill	51.516835	-0.103168	[{'label': 'display', 'lat': 51.51683496238804...	1999	EC1A 2DP	GB	London	Greater London
4	4c49da239e3e2d7fe77c5749	Marylebone Police Station	[{'id': '4bf58dd8d48988d12e941735', 'name': 'P...	v-1612955003	False	1-9 Seymour St.	51.514874	-0.156971	[{'label': 'display', 'lat': 51.51487350290588...	2198	W1H 7BA	GB	City of Westminster	Greater London

As we can see the distribution of police locations is comparable to Berlin, although locations seem to be concentrated more in the city centre.



Discussion Section:

The data analysis section shows that there are interesting patterns emerging in the data set, such as the central location of higher crime rate numbers in Berlin as well as in London. Further exploration to more closely link the crime data to the respective urban locations might be needed, such as taking real-time crime location data by type of crime and time of day into account.

The top crime areas tended to be in relatively central urban locations in both Berlin and London. While this might be related to population density no clear link could be found using the heatmap analysis for London. Further analysis should be conducted to see whether this factor plays a role as the top three crime areas in Berlin indicate, as they are also in the top 5 of most populated areas.

While suburban boroughs seem to be a safer bet in both cities with lower monthly average crime rates, these might not be the prime spots to open a business in terms of a trading perspective, and central location might out way the safety argument.

Conclusion section:

Both cities are relatively equal in terms of comparing crime rates per population and location data such as the distribution of police stations or higher crime rates in urban areas than suburbs. There are significant differences between the two cities in terms of the overall size as well as cultural customs, which a business planning to relocate might also consider, and might out way negative effects of high crime rates.

While there are interesting patterns emerging in the relationship of crime data and urban locations. There is further scope to take multiple additional factors into account, when relocating a business. Considering crime, population and police data might not be relevant enough for a specific business. However, in the interest of time and the scope of the course data sets were limited to a manageable amount. Further research is needed to determine all specific factors that are relevant to a business to make it thrive in the location of their choice in a European metropole.