

Data sources

1. Ethnic groups in London - https://en.wikipedia.org/wiki/Ethnic_groups_in_London
2. Population in London Boroughs - <https://data.london.gov.uk/download/2011-census-demography/62f62c4d-eb60-4846-9efd-1e1373641452/london-unrounded-data.xls>
3. 24 month crime data by London Borough - 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
4. Nearest most popular venues data for the boroughs from FourSquare API

Data will be sourced from 4 different websites. Please note that it is very important to understand the data, its completeness, gaps, caveats and any errors to arrive at robust insights.

1. Ethnic groups in London – all the ethnic group data will be scraped, cleaned and converted into relevant data form (Object to Float for numbers etc.) from this website (https://en.wikipedia.org/wiki/Ethnic_groups_in_London)

	Borough	Indian	Pakistani	Bangladeshi	Chinese	Other_Asian	Total_Asian
0	Newham	42484	30307	37262	3930	19912	133895
1	Redbridge	45660	31051	16011	3000	20781	116503
2	Brent	58017	14381	1749	3250	28589	105986
3	Tower Hamlets	6787	2442	81377	8109	5786	104501
4	Harrow	63051	7797	1378	2629	26953	101808
5	Ealing	48240	14711	1786	4132	31570	100439
6	Hounslow	48161	13676	2189	2405	20826	87257
7	Hillingdon	36795	9200	2639	2889	17730	69253
8	Barnet	27920	5344	2215	8259	22180	65918
9	Croydon	24660	10865	2570	3925	17607	59627
10	Waltham Forest	9134	26347	4632	2579	11697	54389

Absolute numbers may not be relevant as the proportion of a particular ethnic group in that population (or concentration) will paint a better picture for each borough, so we also need total population by each borough.

2. Population in London Boroughs – We will download and use the following excel file from official London Data website and use 'Persons' sheet (<https://data.london.gov.uk/download/2011-census-demography/62f62c4d-eb60-4846-9efd-1e1373641452/london-unrounded-data.xls>).

	Borough	Total_Population
1	City of London	7375.0
2	Barking and Dagenham	185911.0
3	Barnet	356386.0
4	Bexley	231997.0
5	Brent	311215.0
6	Bromley	309392.0
7	Camden	220338.0
8	Croydon	363378.0
9	Ealing	338449.0
10	Enfield	312466.0

Now we have population by ethnicity and total population by borough, so we can calculate Indian concentration (% Indian) population.

	Borough	IndianPercent
4	Harrow	26.37
6	Hounslow	18.96
2	Brent	18.64
1	Redbridge	16.37
5	Ealing	14.25
0	Newham	13.79
7	Hillingdon	13.43
8	Barnet	7.83
9	Croydon	6.79
11	Merton	4.06
17	Barking and Dagenham	4.00
13	Enfield	3.73
10	Waltham Forest	3.54
15	Westminster	3.29
16	Greenwich	3.08

3. Crime data of London Boroughs (24 months) – from official London website
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

It has data by crime type and subtype for the past 24 months (1 column for each month)

	MajorText	MinorText	LookUp_BoroughName	201712	201801	201802	201803	201804	201805	201806	...	201902
1570	Vehicle Offences	Theft from a Motor Vehicle	Westminster	296	258	212	225	258	207	267	...	287
1571	Vehicle Offences	Theft or Taking of a Motor Vehicle	Westminster	50	79	63	58	57	50	58	...	48
1572	Violence Against the Person	Homicide	Westminster	0	0	1	0	0	0	0	...	0
1573	Violence Against the Person	Violence with Injury	Westminster	336	246	230	278	268	315	302	...	288

We will summarize the data over past 24 months and by all crime type for each borough.

	TotalCrime
Borough	
Barking and Dagenham	38231
Barnet	59112
Brent	60983
Camden	74864
Croydon	64392
Ealing	59413
Enfield	57762
Greenwich	54167
Harrow	31820
Hillingdon	52096
Hounslow	51863
Merton	28389
Newham	72057

4. Now will use Four Square API to get 200 venue listings for each borough within the radius of 2.5 kms and find the top 10 most popular venue types.

	Borough	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
0	Barking and Dagenham	Grocery Store	Supermarket	Gas Station	Park	Pizza Place	Soccer Stadium	Pub	Racetrack	Metro Station
1	Barnet	Coffee Shop	Pub	Italian Restaurant	Grocery Store	Café	Turkish Restaurant	Park	Fish & Chips Shop	Pizza Place
2	Ealing	Pub	Coffee Shop	Park	Hotel	Pizza Place	Italian Restaurant	Café	Persian Restaurant	Sandwich Place
3	Harrow	Coffee Shop	Indian Restaurant	Park	Sandwich Place	Pub	Fast Food Restaurant	Grocery Store	Supermarket	Gym / Fitness Center
4	Wandsworth	Coffee Shop	Pub	Park	Café	Pizza Place	Bakery	French Restaurant	Thai Restaurant	Supermarket

Finally, to map them out we will also generate latitude and longitudes of all the boroughs

	Borough	Latitude	Longitude
0	Newham	51.530000	0.029318
1	Redbridge	51.576320	0.045410
2	Brent	30.471943	-87.246916
3	Tower Hamlets	51.128863	1.298669
4	Harrow	51.596769	-0.337275
5	Ealing	51.512655	-0.305195
6	Hounslow	51.468613	-0.361347
7	Hillingdon	51.542519	-0.448335
8	Barnet	51.648784	-0.172913
9	Croydon	51.371305	-0.101957
10	Waltham Forest	51.556999	-0.005835
11	Merton	51.410803	-0.188099