Shifting home in London - Choosing the best borough

Author: Praveen Ramasubramaniam
Date: October 8, 2019

Introduction

About London

London is considered to be one of the world's most important global cities and has been termed the world's most powerful, most desirable, most influential, most visited, innovative, sustainable and most popular for work city in the world according to various sources. London exerts a considerable impact upon the arts, commerce, education, entertainment, fashion, finance, healthcare, media, professional services, research and development, tourism and transportation. It is one of the largest financial centres and is the most-visited city as measured by international arrivals and has the busiest city airport system as measured by passenger traffic. It is the leading investment destination, hosting more international retailers and ultra-high net-worth individuals than any other city. London's universities form the largest concentration of higher education institutes in Europe, and is home of world-class institutions such as Imperial College London in science, technology, engineering, and mathematics and the London School of Economics in economics, finance, and business.

London has a diverse range of people and cultures, and more than 300 languages are spoken in the region. It is the most populous of any city in the European Union and accounting for 13.4% of the UK population. The population within the London commuter belt is the most populous in the EU.

Source: Wikipedia

Background

I am fortunate to work in London and reside in one of the outer Greater London boroughs which is also a commuter belt. There comes a point in time where due to the change in circumstances - socio-economic, financial, cultural or due to any other factors like change in job, expanding family, one may decide to shift his / her home and relocate to another property. There are a number of factors that one may consider to choose where to shift. Some of the factors can be: easy access to public transport, low crime rate in the area, good schools or educational institutions, healthcare facilities, entertainment avenues, make-up of the population in the area, socio-economic indicators within the area amongst others.

Problem Statement

This project is aimed at shortlisting one or more London boroughs to relocate based on a number of criteria that are specific to the author. The project will try to utilise most of the relevant data science topics learned in the lead up to the certification to answer the question. The actual criteria for selection would be reviewed, weighted and finalised as part of the final submission, however some of the initial indicators which are to be considered would be:

1. Access to public transport
2. Access to essential amenities - schools, medical care
3. Community safety - crime rate, incidents etc.,
4. Housing indicators - like average house prices, repossessions etc.,
5. Make-up of the borough in terms of entertainment avenues, most visited places etc.,
6. Make-up of the population in terms of age, socio-economic indicators etc.,
7. Environment quality - pollution, green spaces etc.,

Interest

This exercise would be of particular interest to anyone thinking of relocating to London or those already in London / Greater London area planning to relocate to another area in order to shortlist the boroughs to concentrate and focus on.

Data Acquisition and cleaning

Data sources

The advantage with London is the ready availability of data corresponding to various categories in different formats. The chief source of data for all the analysis is taken from the London Data Store which is a free and open data-sharing portal. Once the data source is identified, the main activity is to choose the specific datasets that correspond to the problem at hand. Therefore, I have considered the following datasets for the analysis grouped together as below:

Category	Dataset
Crime	Borough Level Crimes (over the last 24 months)
General Borough make-up	Qualifications of the working age population
	Population estimates by borough
	Public Transport Accessibility Levels
Educational indicators	Schools – Pupils
	GCSE results
	KS1 results
	KS2 results
	Early Year Foundation Stages results
Health	Childhood obesity
	Smoking prevalence
	Sports participation
	Walking / Cycling frequency
Mental Health	Personal well-being
Socio-economic indicators	Earnings for Working age population

Category	Dataset
	Homeless population
	Vehicles
Economic indicator	Average House Prices

Further, I have also taken location data from FourSquare API for nearby venues to each of the London Boroughs on a 350m radius and a limit of 20 venues for each borough.

Data Cleaning

With such a wide range of data, there is a need to perform exhaustive data cleaning and feature selection to make the relevant analysis for the London boroughs. The data is sourced in two major formats: comma separated value files as well as excel spreadsheets.

There are several problems with the datasets. Some of the issues are:

1. Non-London borough data which needs to be discarded for the purpose of the analysis
2. Historic data not relevant to latest analysis which needs to be discarded
3. Very granular data – with multiple features in a single dataset which needs to be aggregated / grouped for pertinent analysis for the London boroughs
4. Data spread over multiple spreadsheets which needs to be collated and the data merged to create a common data frame.

After these initial data cleansing and sorting, I check for missing data and make appropriate provisions for the same. I also check for data outliers – these are mostly concentrated around the City of London borough mainly because of its extremely small size and population.

Feature Selection

Once the data has been cleansed, the next problem to tackle is to select the features to be taken for analysis. For some of the data sets, it is already present in the data set and it is easy to focus on the feature for analysis, but for the majority of the data sets – the feature set needs to be created based on the existing data based on aggregation, average or other functions.

Given the thousands of available features, it is a major undertaking to reduce the feature set to its most appropriate characteristics for the sake of this exercise.

For e.g. for crime related data, the dataset had 27 columns with month-wise data split of the crimes in each London borough starting from 2017/10 to 2019/09. Though this data is undoubtedly useful for granular level of crime data for London boroughs, from our perspective, where crime is just one of the factors under consideration, having that level of granularity is an overkill. Hence, I had to reduce the dataset by aggregating the crimes between 2017/10 to 2018/09 as 2017-18 and similarly the crimes between 2018/10 to 2019/09 as 2018-19 for simpler and meaningful analysis when looking at trends over a longer period of time.

The following table shows the details of the features that were selected / dropped and the rationale for the same:

Dataset	Key Feature(s) for analysis	Derivation of key feature(s)	Reason for choosing the key feature(s)
Crime	Crimes in 2017-18 and 2018-19	Aggregation of month-wise crime figures for each borough	To visualize trend in crimes over the last 2 years
Qualifications of working age population	NVQ 4+ qualifications of working population for each borough over the last 4 years	Removal of non-London data	Not relevant to current analysis
		Removal of data between 2004 and 2013	Historic data – focus only on recent past
		Removal of other qualifications	Focus on highest qualifications for simplicity
Population	Total Working Population	Removal of non-London data	Not relevant to current analysis
		Removal of data older than 2013	Historic data – focus only on recent past
		Defining new feature: Total Working Population	To focus on segment of population bringing maximum economic benefits
Schools / Pupils	Pupils per school	Aggregation of all types of schools in each borough	To focus on total number of schools in each borough
		Aggregation of all types of students in each borough	To focus on total number of students in each borough
		Defining new feature: Pupils per school	To simplify the analysis for meaningful comparison
GCSE	Attainment 8 values for 2016/17 and 2015/16	Removal of non-London data	Not relevant to current analysis
		Consideration of total student achievements	Focus on all students – not gender differences
		Attainment 8 for last 2 years	Focus on the latest data for analysis
KS1	Average Expected Standard and Average Greater Depth	Removal of non-London data	Not relevant to current analysis
		Defining new feature: Average Expected Standard	Looking at wider perspective than individual subjects
		Defining new feature: Average Greater Depth	Looking at wider perspective than individual subjects

Dataset	Key Feature(s) for analysis	Derivation of key feature(s)	Reason for choosing the key feature(s)
KS2	Expected Standard All (Read, Write, Maths) Higher Standard All (Read, Write, Maths)	Focus on Expected Standard and Higher Standard for all subjects	Looking at wider perspective than individual subjects
Early Years Foundation Stage	% Exp Levels across all ELGs % Good Level of development	Focus on specific subset of data	Focus on performance – not number of students or gender differences
Income Levels	Median Income 2014-15 Median Income 2015-16 Median Income 2016-17	Remove population numbers	Focus on median income not on population numbers
		Removal of older data	Focus on most recent set of numbers
Childhood Obesity Indicators	Healthy Weight Y6% Obese Y6 % Severely Obese Y6 %	Removal of non-London data	Not relevant to current analysis
		Removal of confidence intervals	Focus on prevalence rather than confidence brackets
		Removal of Reception metrics	Too young to report any meaningful observations
Smoking	Current Smoking Rate % 2015 Current Smoking Rate % 2016 Current Smoking Rate % 2017	Removal of non-London data	Not relevant to current analysis
		Removal of older data	Focus on the most recent set of data for analysis
		Removal of number of smokers	Focus on percentage rather than numbers
Well-being	Happiness Mean: 2015-16 and 2016-17 Anxiety Mean: 2015-16 and 2016-17	Removal of non-London data	Not relevant to current analysis
		Removal of older data	Focus on the most recent set of data for analysis
		Focus on Happiness and Anxiety as the main metrics	More stronger metrics compared to Life Satisfaction and Worthwhile
Sports Participation	1 per week %: 2014-15 and 2015-16	Removal of non-London data	Not relevant to current analysis
		Removal of older data	Focus on the most recent set of data for analysis

Dataset	Key Feature(s) for analysis	Derivation of key feature(s)	Reason for choosing the key feature(s)
	3 per week %: 2014-15 and 2015-16	Focus on percentage of population	Remove focus on total numbers
		Focus on active population	Remove focus on population not participating in any sport
Homelessness	Number of priority homeless (per 1000 households) Households accommodated by authority (per 1000 households)	Focus only on the latest set of data	To get the most recent picture of the state of the boroughs
		Focus on numbers per 1000 households	Disregard ethnicity and total numbers or split in type of accommodation provided by Las
Vehicles	Total PLG	Focus on the most recent set of data on vehicles	To get the most recent picture of the state of the boroughs
		Focus on total number of vehicles	Disregard the type of vehicles – cars / others as well as ownership – company or private
Walking / Cycling	Walking 3x or 5x per week: 2015/16 and 2016/17 Cycling 3x or 5x per week: 2015/16 and 2016/17	Removal of non-London data	Not relevant to current analysis
		Removal of older data	Focus on the most recent set of data for analysis
		Focus on the most active population metrics	Removal of data corresponding to walking 1x per week, 1x per month or lesser frequency
Public Transport Accessibility Levels	Avg PTAI 2015	Focus on the Accessibility Value	For easier comparison between boroughs especially when two boroughs are in same band
Median House Prices	Median House Price: 2015, 2016 and 2017	Removal of non-London data	Not relevant to current analysis
		Removal of older data	Focus on the most recent set of data (3 years) for analysis
		Focus on median figures	More accurate representation than average / sales numbers

Data Analysis

The ultimate aim of this activity is to rank the boroughs on the different characteristics and to help identify the suitability / desirability of a particular borough based on the weightage of the individual characteristics. This exercise will not provide a final ranking of the boroughs as each individual may have a different opinion on the suitability of the features chosen, the weightage of a particular feature based on their individual preference, socio-economic criteria, stage in their life etc., However, I will try to provide detail of one such way the boroughs can be ranked overall based on my order of preference / prioritization.

The available datasets are categorized into 7 categories as mentioned in the Data Acquisition and Cleansing section as below:
Crime, General borough make-up, Educational indicators, Health, Mental health, socio-economic indicators and economic indicators.

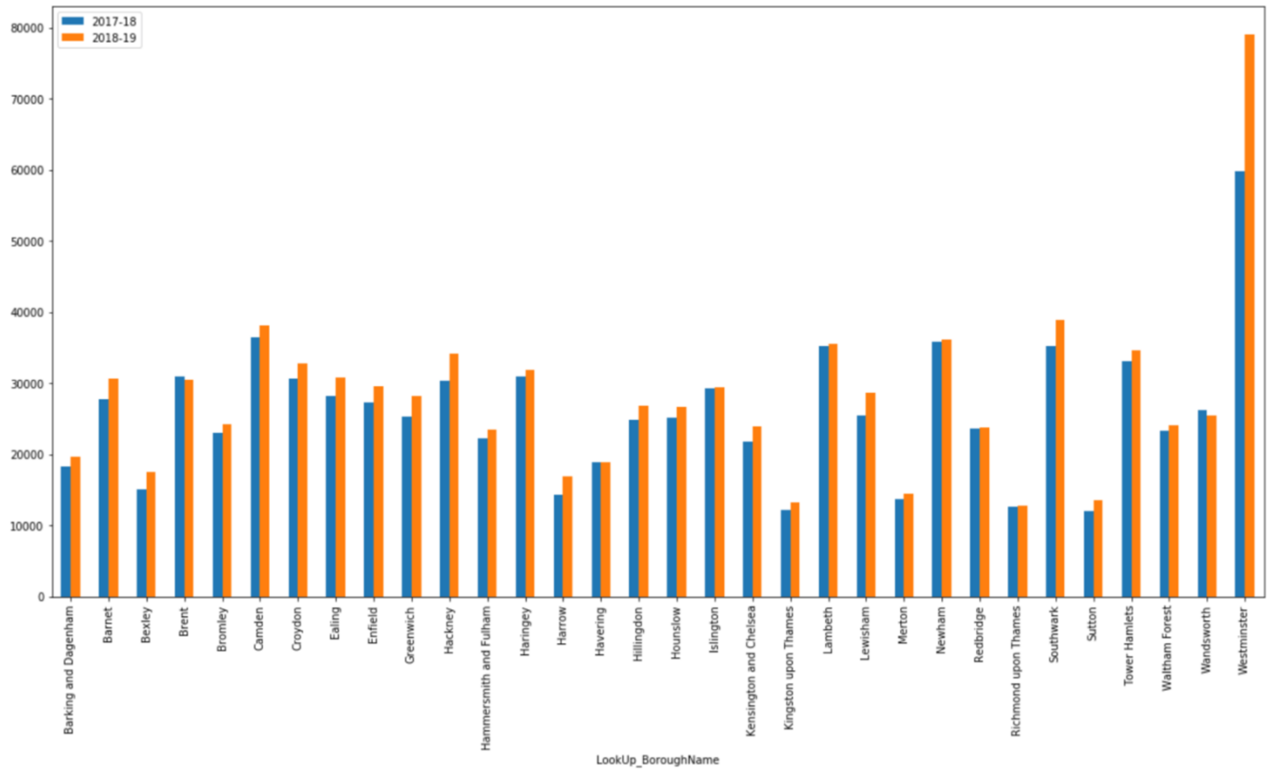
For each of the category and the individual dataset associated with the categories, I have tried to analyze the data by cleaning the data and feature selection as detailed in the earlier sections to come with 3 outputs:

1. Bar chart plotting the Key Feature against the different boroughs
2. Heat map plotting the Key Feature or subset of the key feature against the London boroughs
3. List of the 10 boroughs that are leading in the statistics for that Key Feature

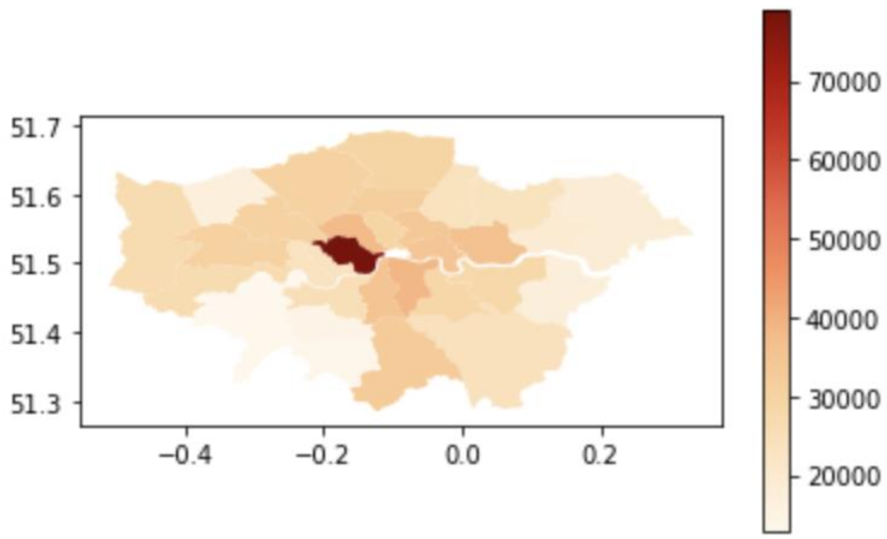
Crime

I believe safety or living in a safe neighbourhood with fewer crimes especially serious or violent crimes would be a main criterion applicable for almost everyone. Further, we need to look into the most recent set of crime data to get a clearer picture of the borough make-up. Though the dataset available is very extensive dealing with month level number of crimes split by major and minor crime categories – given that it is just one of the seven categories and 18 different datasets – this level of granularity is not required. So, I have considered an aggregate number of crimes over the last 2 years on an individual borough basis.

Bar chart - Crime



Borough Crime Heatmap



Top 10 boroughs – with least crimes

	2017-18	2018-19
LookUp_BoroughName		
Richmond upon Thames	12633	12792
Kingston upon Thames	12099	13199
Sutton	11919	13453
Merton	13645	14441
Harrow	14283	16896
Bexley	15107	17474
Havering	18802	18890
Barking and Dagenham	18224	19631
Hammersmith and Fulham	22261	23440
Redbridge	23551	23767

The bar chart shows visually not only the number of crimes for each borough but also the increase / decrease in the crime numbers compared to the last year. This is also borne out in the heatmap where Westminster is shown as deep red as the borough having the most crimes. The bar chart is validated via the table which shows the 10 boroughs with the least number of crimes irrespective of the type of crime.

General borough make-up

I have taken the following datasets to be categorized into the general borough make-up and these are:
Qualifications of the working age population
Population estimates – Total working population
Public Transport Accessibility Levels

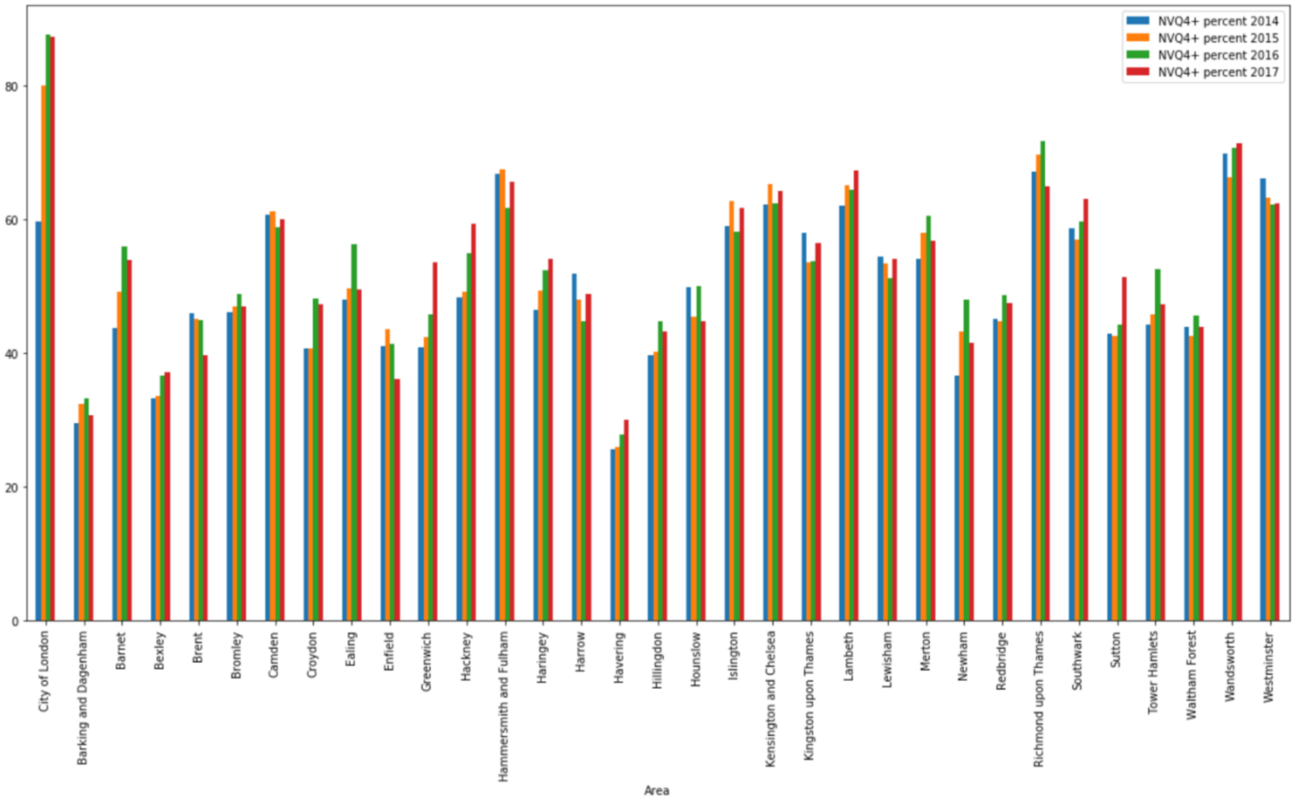
I am more interested in the statistics around the type of population in each borough as categorized by the percentage of population with the highest qualifications terms as NVQ 4+ and the Total Working Population in numbers as defined as the population between 18 and 60. Finally, the Public Transport Accessibility Levels shows the ratings of each borough in terms of density and accessibility of the public transport network for each borough based on the latest available data.

The weightage to this category and maybe even the selection of this category of the data and its datasets into the analysis may differ for each individual. Some may prefer to live amongst well-educated neighbourhood populace with high numbers of economically active population. I personally give very high preference to Public Transport Accessibility levels as commuting by private transport especially to Inner London boroughs is a very big hassle due to traffic, parking, congestion zone charges and other considerations. Easy accessibility and availability of public transport trumps the need for own vehicular transport for me.

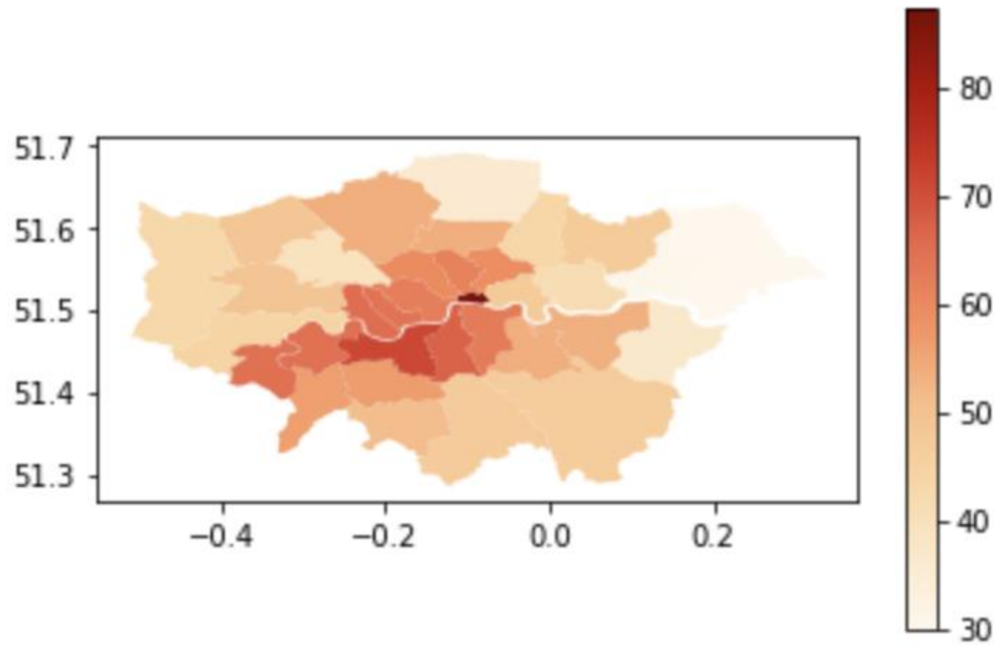
Qualifications of the working age population

The focus is on the highest qualifications termed as NVQ 4+ as percent of the population in each borough.

Bar chart



Heat Map



Top 10 boroughs - Qualifications

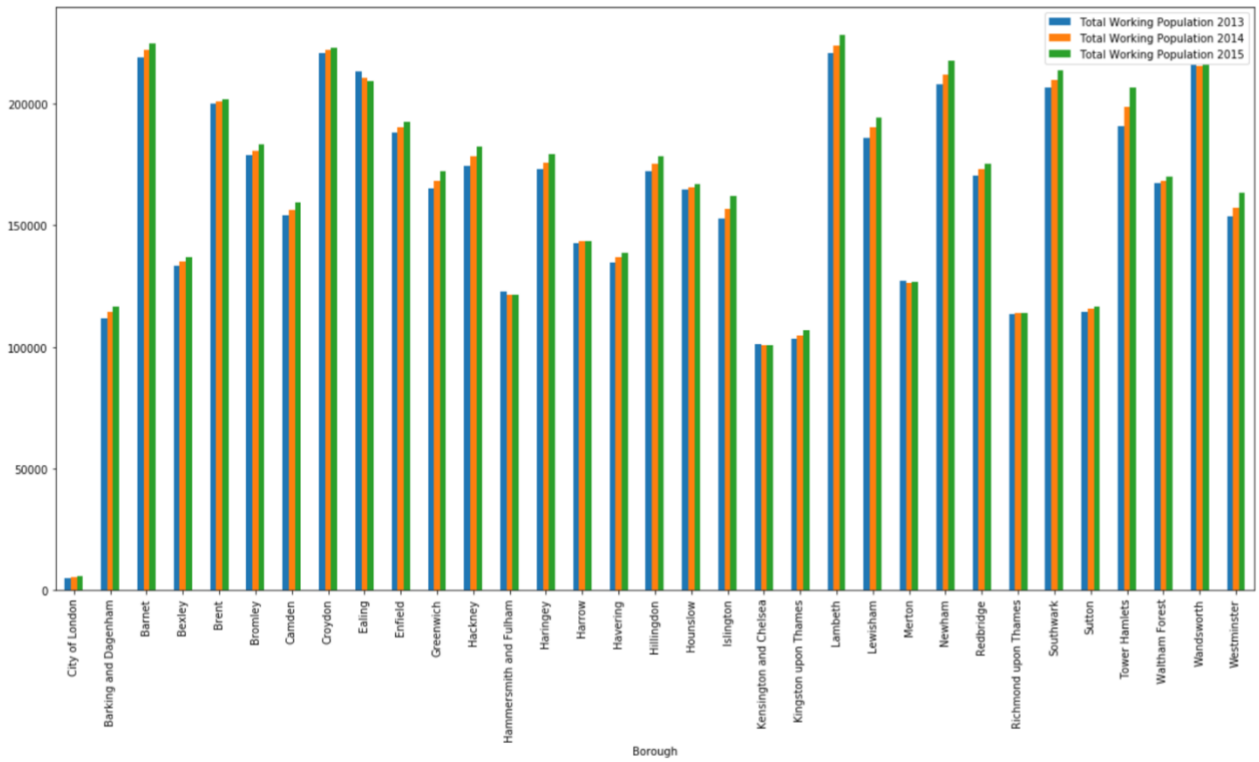
	NVQ4+ percent 2014	NVQ4+ percent 2015	NVQ4+ percent 2016	NVQ4+ percent 2017
Area				
City of London	59.6	80.0	87.6	87.2
Wandsworth	69.8	66.3	70.6	71.3
Lambeth	62.0	65.0	64.4	67.2
Hammersmith and Fulham	66.8	67.4	61.6	65.6
Richmond upon Thames	67.0	69.6	71.6	64.9
Kensington and Chelsea	62.2	65.3	62.4	64.2
Southwark	58.7	56.9	59.6	63.0
Westminster	66.1	63.2	62.1	62.3
Islington	58.9	62.7	58.1	61.7
Camden	60.6	61.1	58.8	59.9

From the charts and tables – we can see that the City of London (a very small geographical area) has the most qualified population.

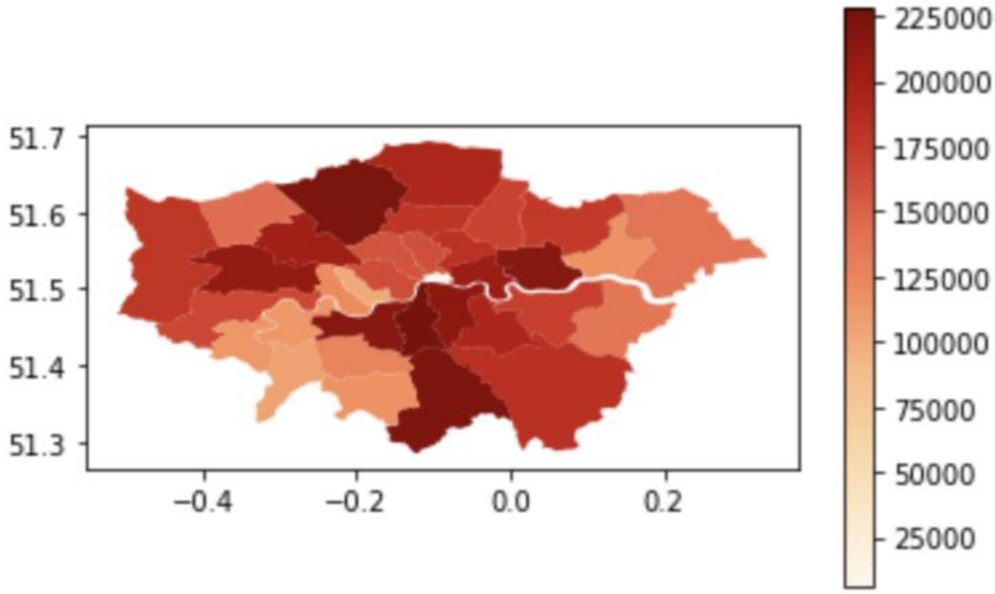
Population – Total Working Population

The focus was on the most economically active population – which I termed as the Working Population aged between 18 and 60 irrespective of the gender.

Bar chart – Total working population



Heatmap – Total working population



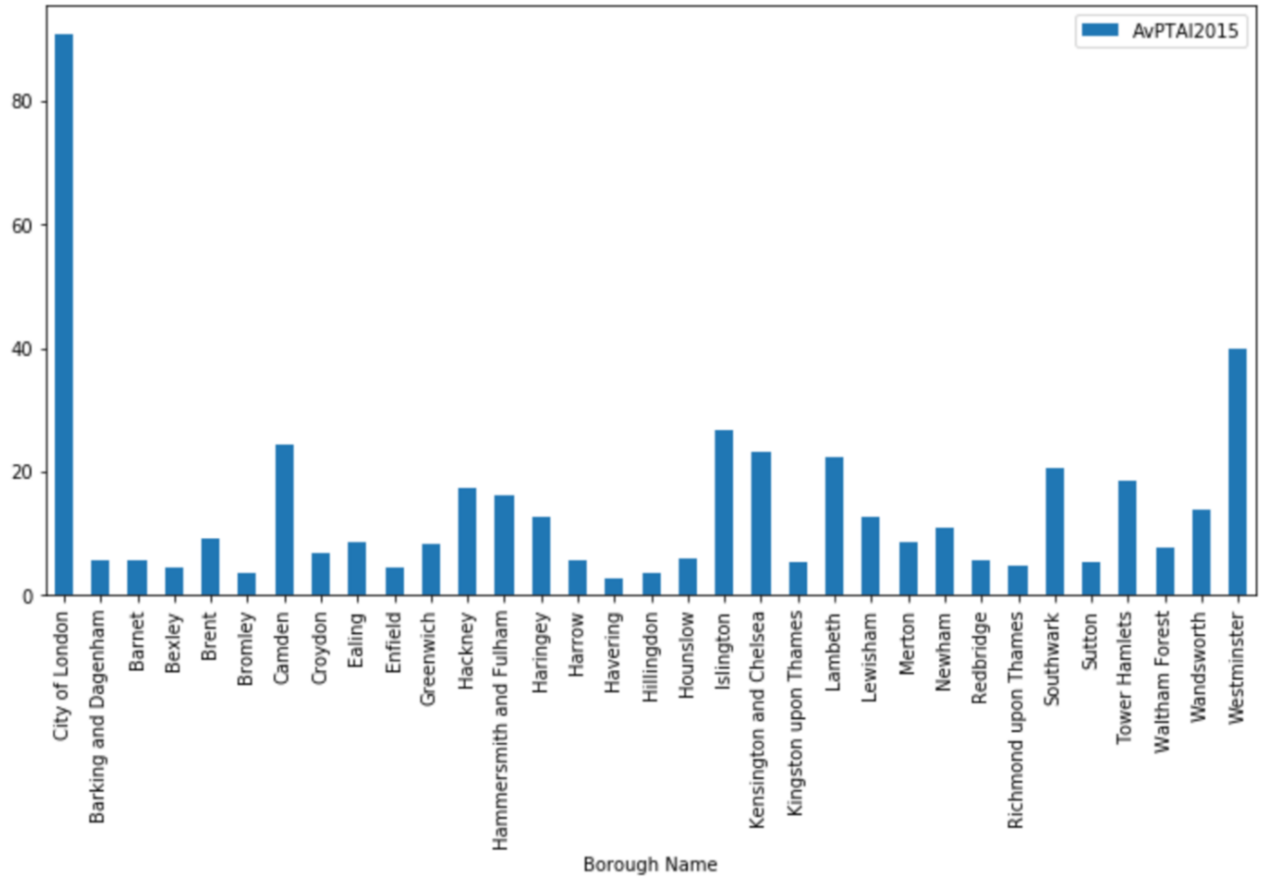
Top 10 boroughs – Total working population

	Total Working Population 2013	Total Working Population 2014	Total Working Population 2015
Borough			
Lambeth	220786	223775	228348
Barnet	218826	222108	224859
Croydon	220595	221873	222944
Newham	207984	211681	217693
Wandsworth	215907	215597	216218
Southwark	206763	209542	213829
Ealing	213099	210691	209412
Tower Hamlets	190668	198597	206577
Brent	200127	200736	201761
Lewisham	185785	190106	194141

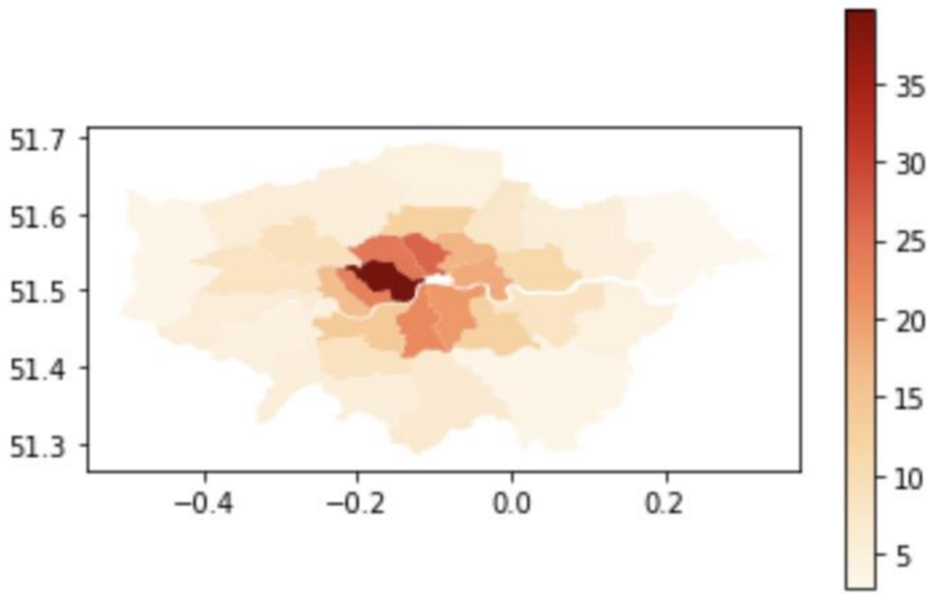
Public Transport Accessibility Levels

Public Transport Accessibility is the measure of the density and accessibility of the public transport for each borough. The focus is on the scores rather than the levels to compare better access between boroughs at the same level.

Barchart



Heatmap (removing City of London)



Top 10 boroughs – public transport accessibility

	Borough Code	AvPTAI2015	PTAL
Borough Name			
Westminster	E09000033	39.750796	6a
Islington	E09000019	26.667566	6a
Camden	E09000007	24.316782	5
Kensington and Chelsea	E09000020	23.262691	5
Lambeth	E09000022	22.239912	5
Southwark	E09000028	20.513884	5
Tower Hamlets	E09000030	18.532870	4
Hackney	E09000012	17.478762	4
Hammersmith and Fulham	E09000013	16.325973	4
Wandsworth	E09000032	13.973088	3

Educational Indicators

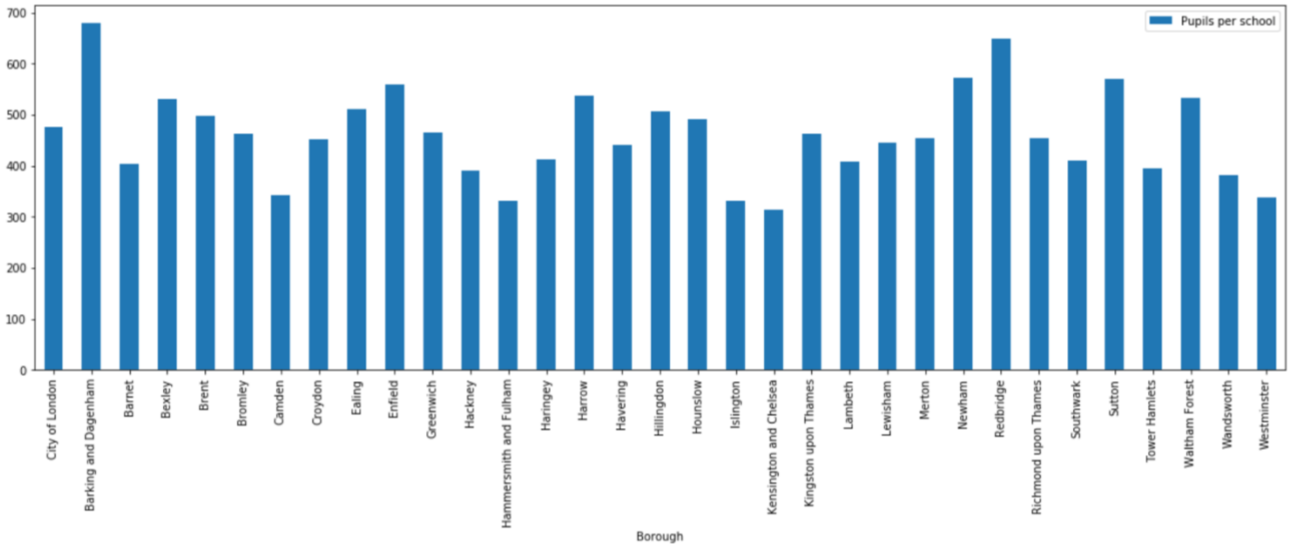
The educational indicators consist of the following datasets:
Schools-Pupils
GCSE
KS1 results
KS2 results
Early Years Foundation Stage Results

Education is one of the important categories of borough comparison and has a number of associated datasets for analysis. It covers a wide range of education spectrum right from nursery to secondary school results.

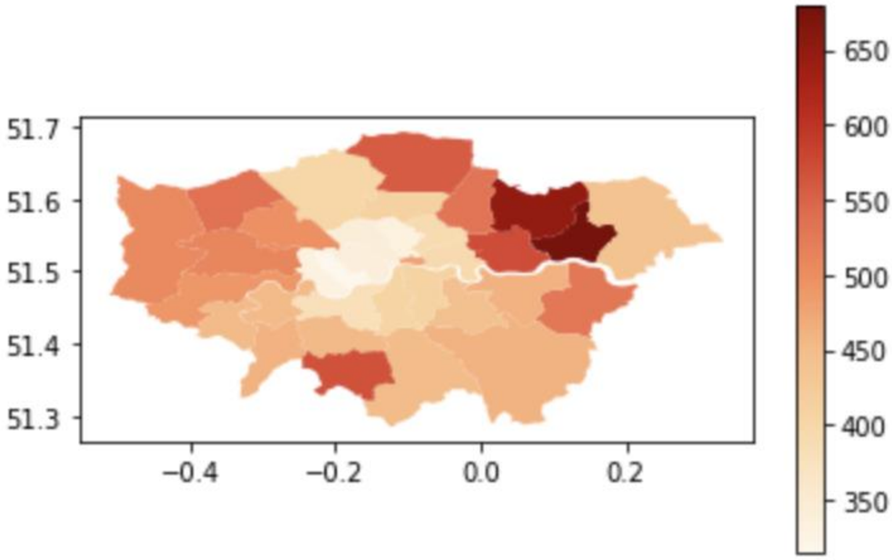
Schools-Pupils

The focus is to rank the London boroughs based on the pupils per schools across all categories of schooling – from nursery to secondary including independent schools and all types of students. The result is a single value of Pupils per school for each London borough for direct comparison.

Bar chart – Pupils per school



Heatmap – Pupils per school



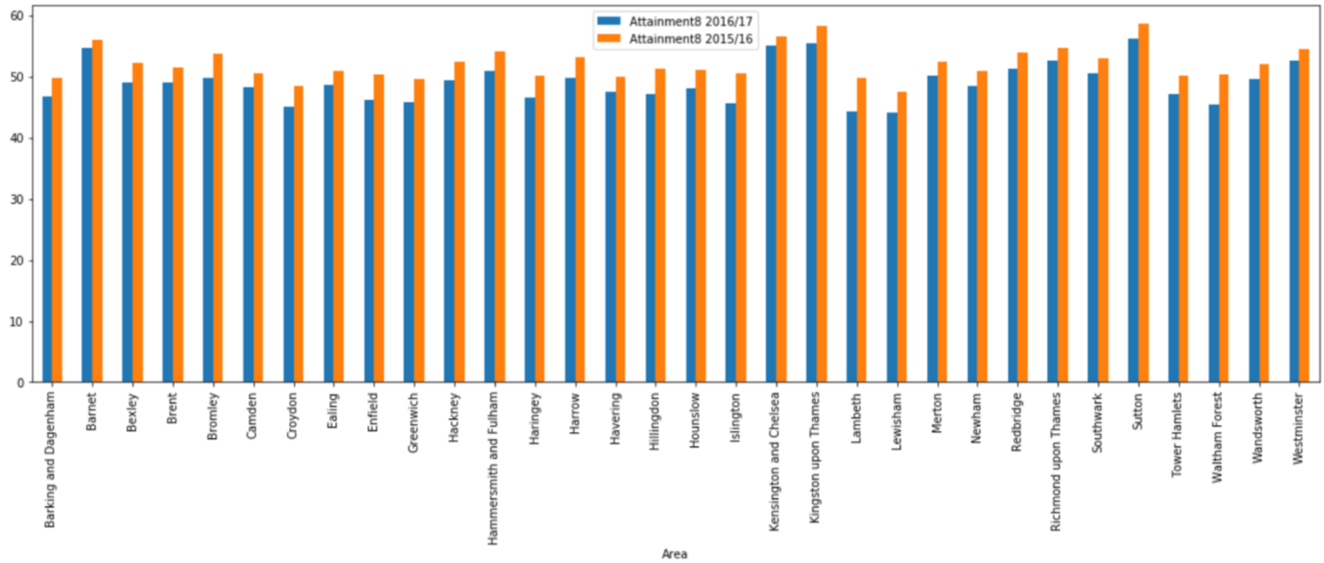
Top 10 boroughs – Pupils per school

	Pupils per school
Borough	
Kensington and Chelsea	314.444444
Islington	331.923077
Hammersmith and Fulham	332.469880
Westminster	338.617021
Camden	343.347368
Wandsworth	380.675214
Hackney	390.956140
Tower Hamlets	393.983333
Barnet	403.613497
Lambeth	408.353535

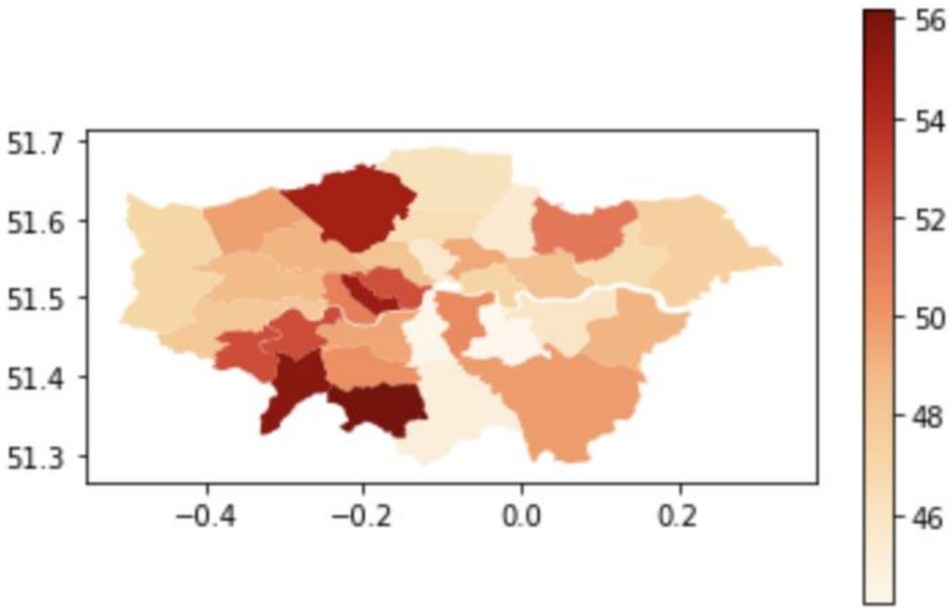
GCSE

GCSE (General Certificate of Secondary Education) is the most commonly used metric for secondary education. I have taken the Attainment 8 percentage data for analysis and ranking of London boroughs.

Bar chart



Heat Map



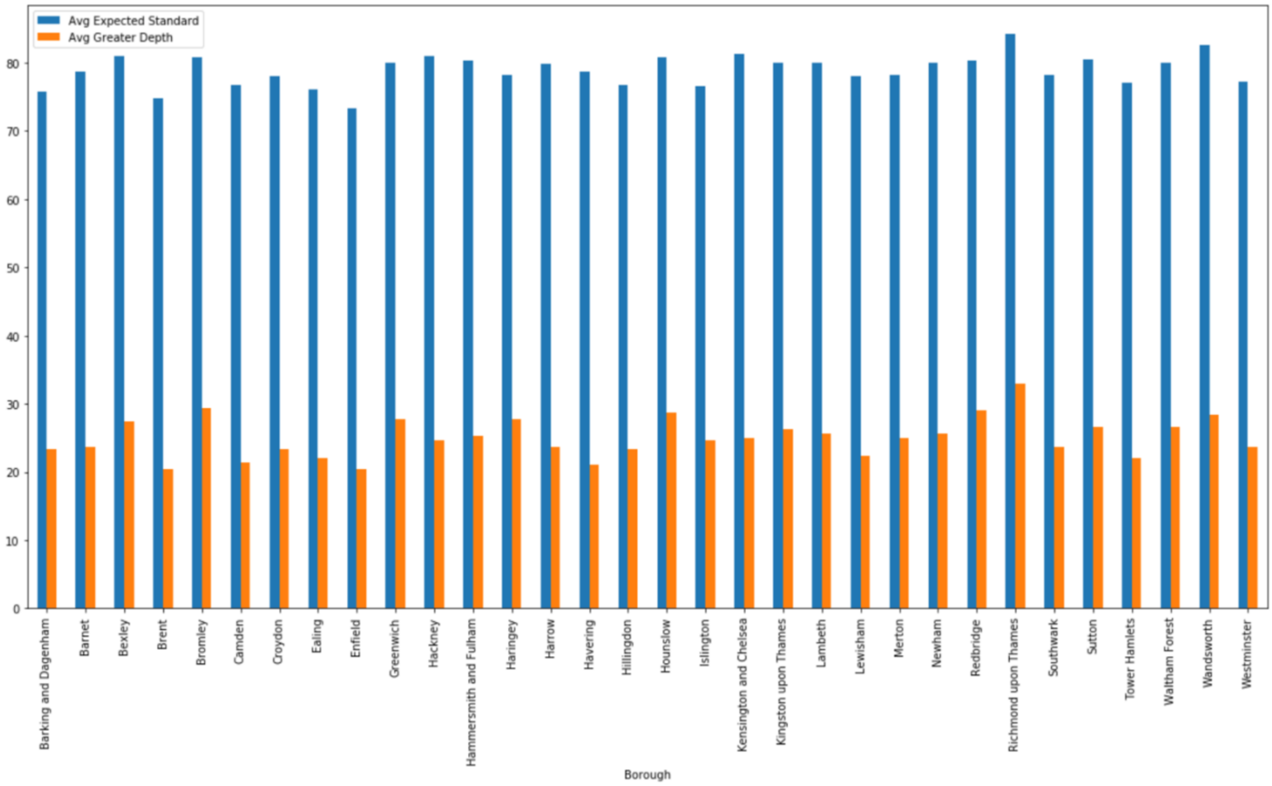
Top 10 boroughs – GCSE performance

	Attainment8 2016/17	Attainment8 2015/16
Area		
Sutton	56.2	58.7
Kingston upon Thames	55.5	58.2
Kensington and Chelsea	55.0	56.6
Barnet	54.7	56.1
Richmond upon Thames	52.7	54.6
Westminster	52.6	54.5
Redbridge	51.2	53.9
Hammersmith and Fulham	50.9	54.1
Southwark	50.5	52.9
Merton	50.2	52.4

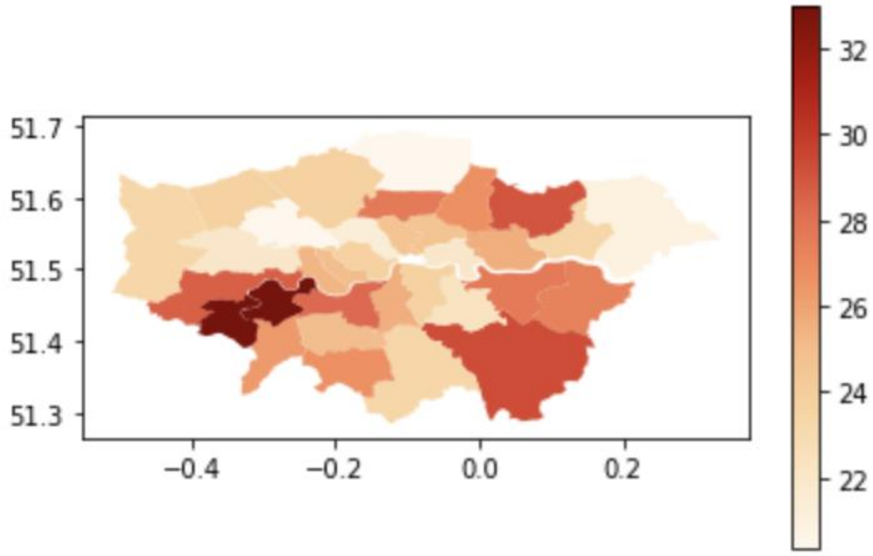
KS1

KS1 is the primary school performance metrics for pupils aged 5 to 7. The main performance metrics for comparison is Expected Standard across Read, Write and Maths as well as Greater Depth across those 3 subjects.

Bar chart – KS1



Heatmap – KS1



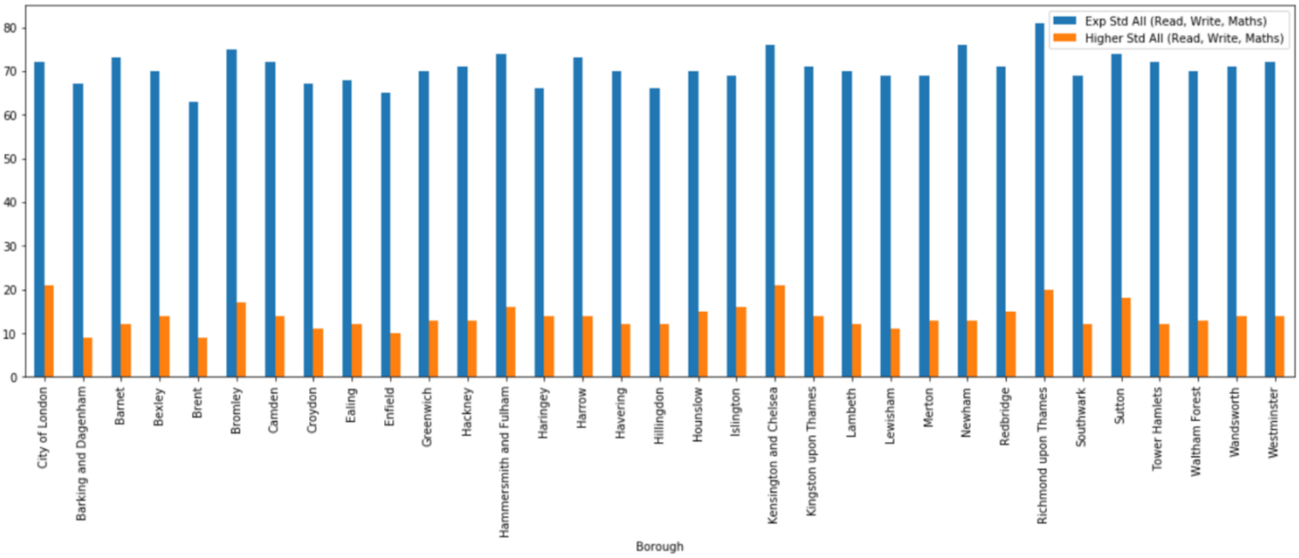
Top 10 boroughs – KS1

	Avg Expected Standard	Avg Greater Depth
Borough		
Richmond upon Thames	84.25	33.000000
Bromley	80.75	29.333333
Redbridge	80.25	29.000000
Hounslow	80.75	28.666667
Wandsworth	82.50	28.333333
Haringey	78.25	27.666667
Greenwich	80.00	27.666667
Bexley	81.00	27.333333
Sutton	80.50	26.666667
Waltham Forest	80.00	26.666667

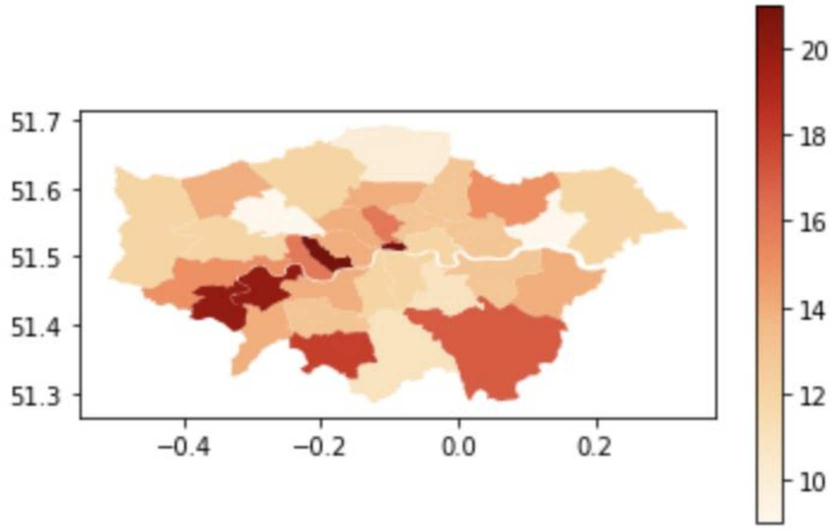
KS2

KS2 is another primary school performance metric taught in Years 3 to 6, when children are between 7 and 11 years-old. Similar to KS1, the focus is on the Expected Standard across Read, Write and Maths and Higher Standard across the same subjects.

Barchart



Heatmap



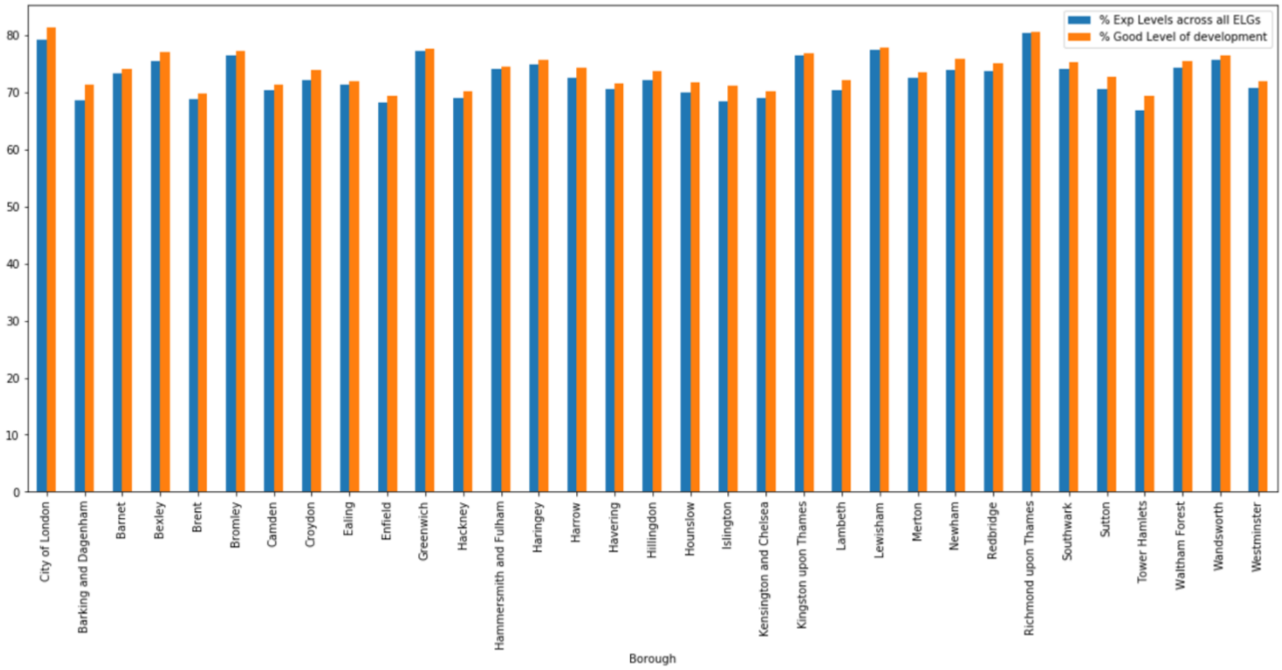
Top 10 boroughs – KS2

	Exp Std All (Read, Write, Maths)	Higher Std All (Read, Write, Maths)
Borough		
City of London	72	21
Kensington and Chelsea	76	21
Richmond upon Thames	81	20
Sutton	74	18
Bromley	75	17
Islington	69	16
Hammersmith and Fulham	74	16
Redbridge	71	15
Hounslow	70	15
Haringey	66	14

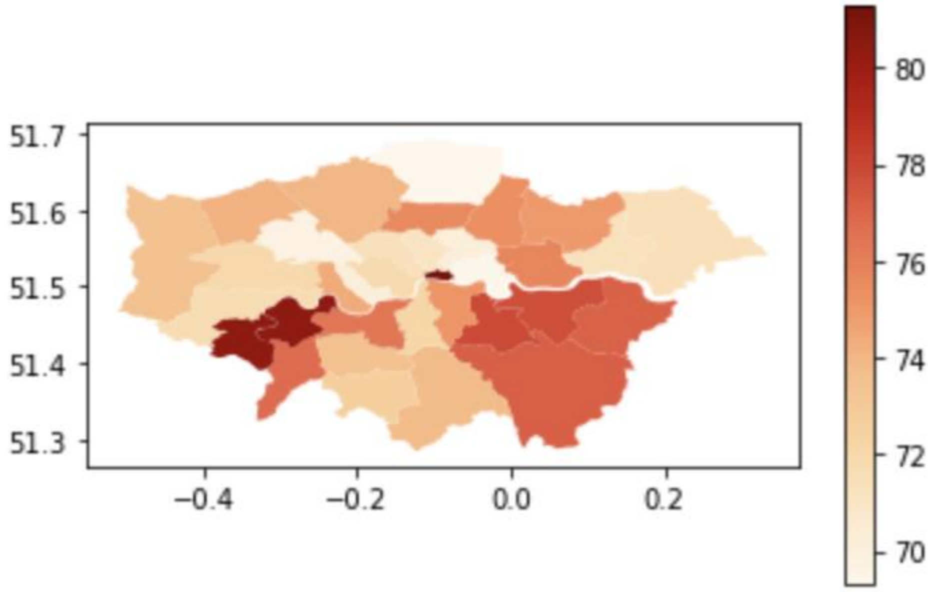
Early Year Foundation Stages (EYFS)

The early years foundation stage (EYFS) sets standards for the learning, development and care of the child from birth to 5 years old. For our analysis, we consider the ranking of the boroughs based on the Expected Levels across all Early Learning Goals (ELGs) % and Good Level of Development %.

Bar chart



Heatmap



Top 10 boroughs – EYFS

	% Exp Levels across all ELGs	% Good Level of development
Borough		
City of London	79.2	81.3
Richmond upon Thames	80.3	80.5
Lewisham	77.4	77.9
Greenwich	77.3	77.7
Bromley	76.5	77.2
Bexley	75.4	77.1
Kingston upon Thames	76.5	76.8
Wandsworth	75.7	76.4
Newham	73.9	75.8
Haringey	74.8	75.6

Health

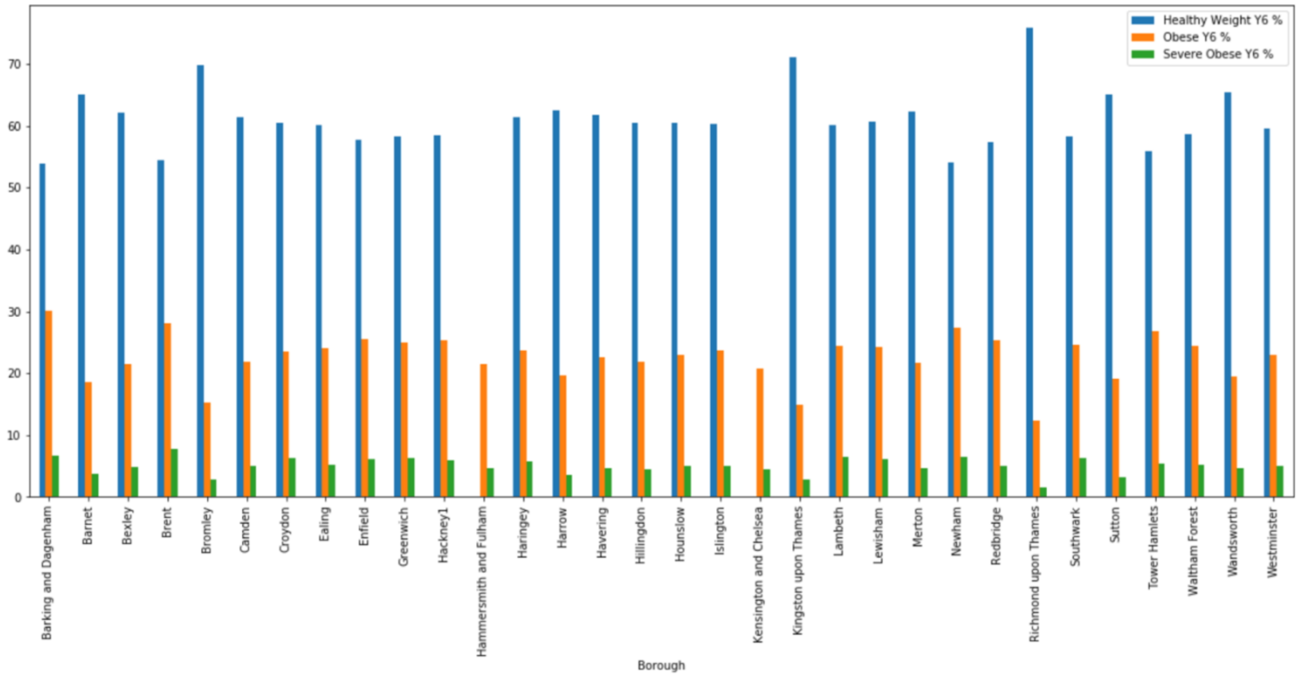
Health is another major category for analysis and consist of the following datasets:
Childhood obesity, Smoking prevalence, Sports participation, Walking / Cycling frequency

A healthier neighbourhood puts lesser pressure on the NHS thereby helping channel and funding for more serious and preventive measures.

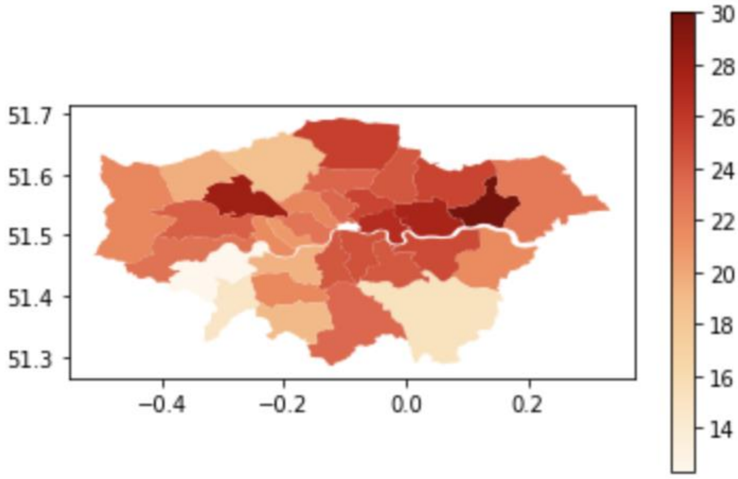
Childhood obesity

Obesity is one of the most important health concerns facing our society today and the issue of childhood obesity is a very serious issue as it leads to lifelong complications if not identified and treated appropriately. The focus of this exercise is to consider the Childhood Obesity measures at Year 6 (approx. 11 years old) to categorize the child population to healthy weight, obese and seriously obese percentage.

Bar chart



Heatmap



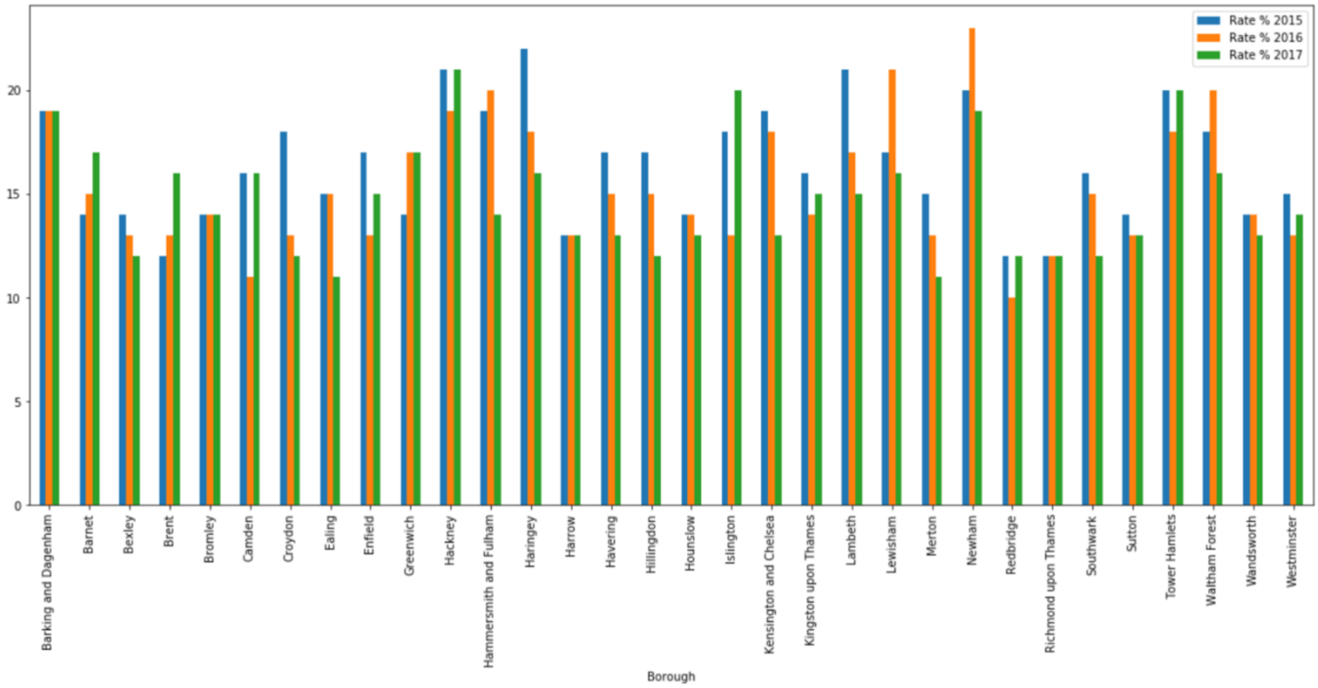
Top 10 boroughs – based on least Obesity percentage

	Healthy Weight Y6 %	Obese Y6 %	Severe Obese Y6 %
Borough			
Richmond upon Thames	75.815085	12.262774	1.508516
Kingston upon Thames	71.154930	14.816901	2.760563
Bromley	69.856322	15.201149	2.844828
Barnet	65.104301	18.593871	3.682720
Sutton	65.020576	19.112940	3.109282
Wandsworth	65.347826	19.521739	4.608696
Harrow	62.458349	19.733432	3.554239
Kensington and Chelsea	NaN	20.775027	4.413348
Hammersmith and Fulham	NaN	21.534847	4.698512
Bexley	62.222936	21.554770	4.754256

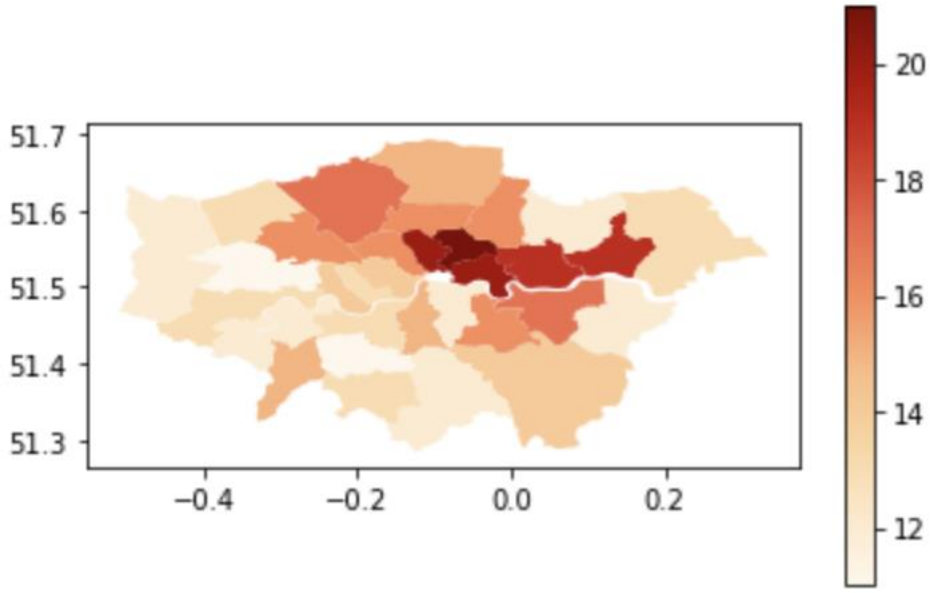
Smoking prevalence

This analysis is pretty self-explanatory – but the focus is on ranking the boroughs based on rate of population in the boroughs considered as current / active smokers.

Bar chart



Heat map



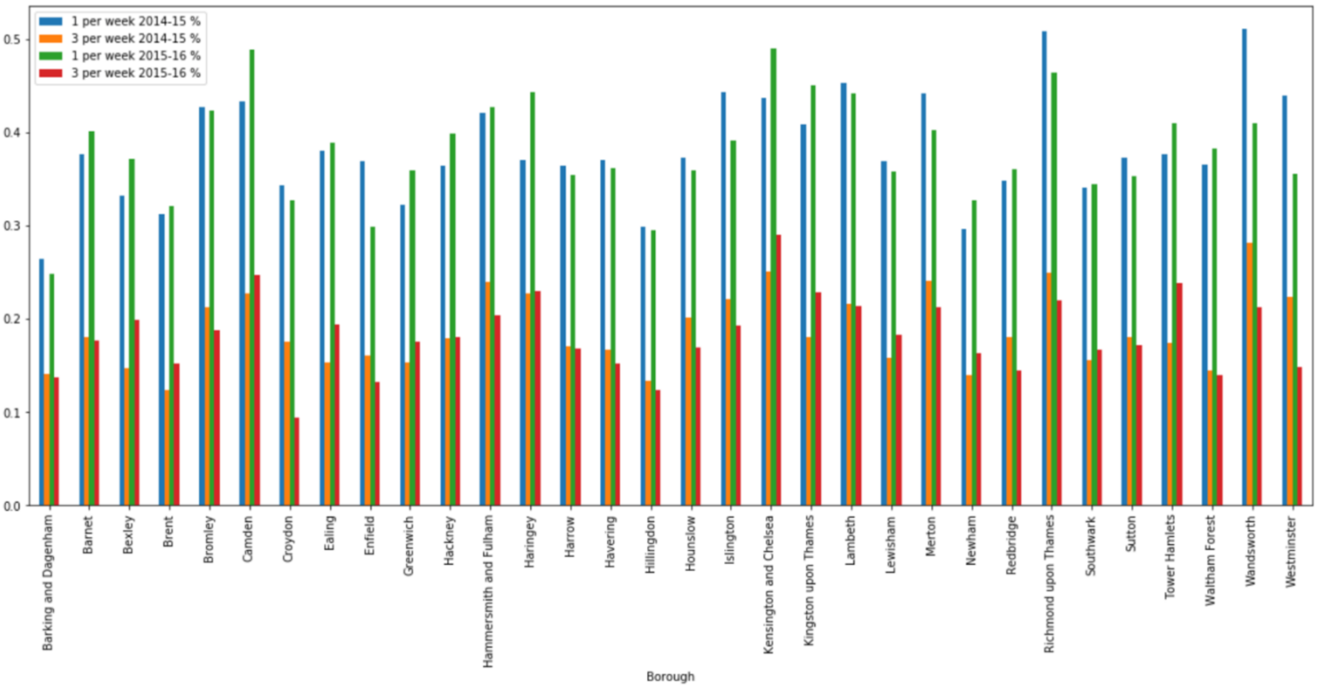
Top 10 boroughs – least active smoker percentage

	Rate % 2015	Rate % 2016	Rate % 2017
Borough			
Merton	15	13	11
Ealing	15	15	11
Redbridge	12	10	12
Richmond upon Thames	12	12	12
Bexley	14	13	12
Croydon	18	13	12
Hillingdon	17	15	12
Southwark	16	15	12
Harrow	13	13	13
Sutton	14	13	13

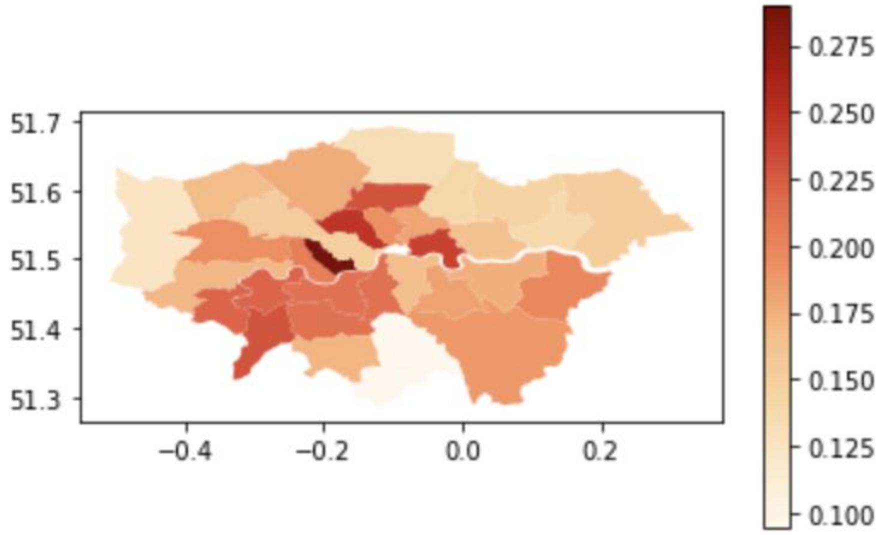
Sports participation

I personally consider an active population to be a healthy population. For the purpose of this exercise, I have taken the sports participation as one of the proxies for a healthy population. This is done by considering the percentage of population in each of the boroughs that participate in sports actively as represented by once or three times per week and ranking the boroughs.

Bar chart



Heat map



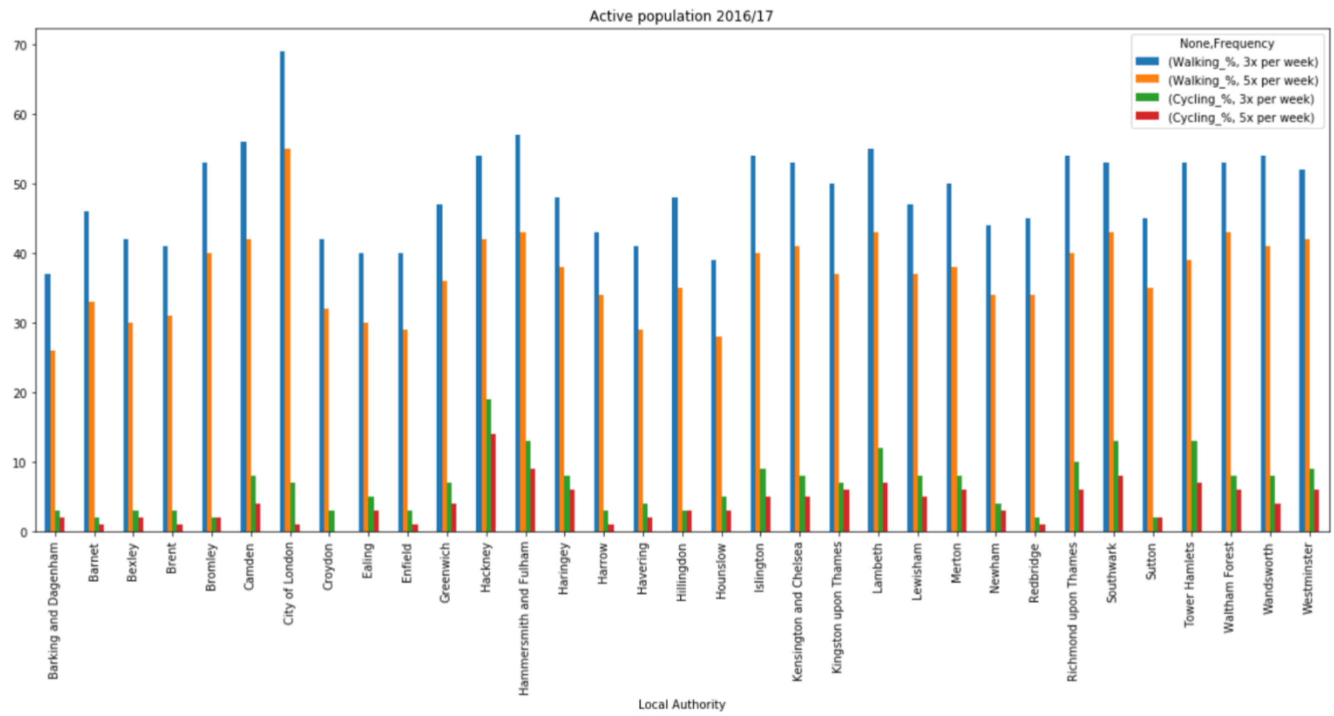
Top 10 boroughs – active sports participation

	1 per week 2014-15 %	3 per week 2014-15 %	1 per week 2015-16 %	3 per week 2015-16 %
Borough				
Kensington and Chelsea	0.437	0.251	0.490	0.290
Camden	0.433	0.228	0.489	0.247
Tower Hamlets	0.377	0.174	0.410	0.239
Haringey	0.371	0.228	0.444	0.230
Kingston upon Thames	0.409	0.181	0.451	0.229
Richmond upon Thames	0.509	0.250	0.465	0.220
Lambeth	0.453	0.216	0.442	0.214
Wandsworth	0.511	0.282	0.410	0.213
Merton	0.442	0.241	0.403	0.213
Hammersmith and Fulham	0.421	0.240	0.427	0.204

Walking / Cycling frequency

Another proxy for active population is the frequency of population that walk and / or cycle frequently. This is more pertinent than active sports participation due to lack of sports facilities or grounds or equipment in certain areas / neighbourhoods.

Bar chart



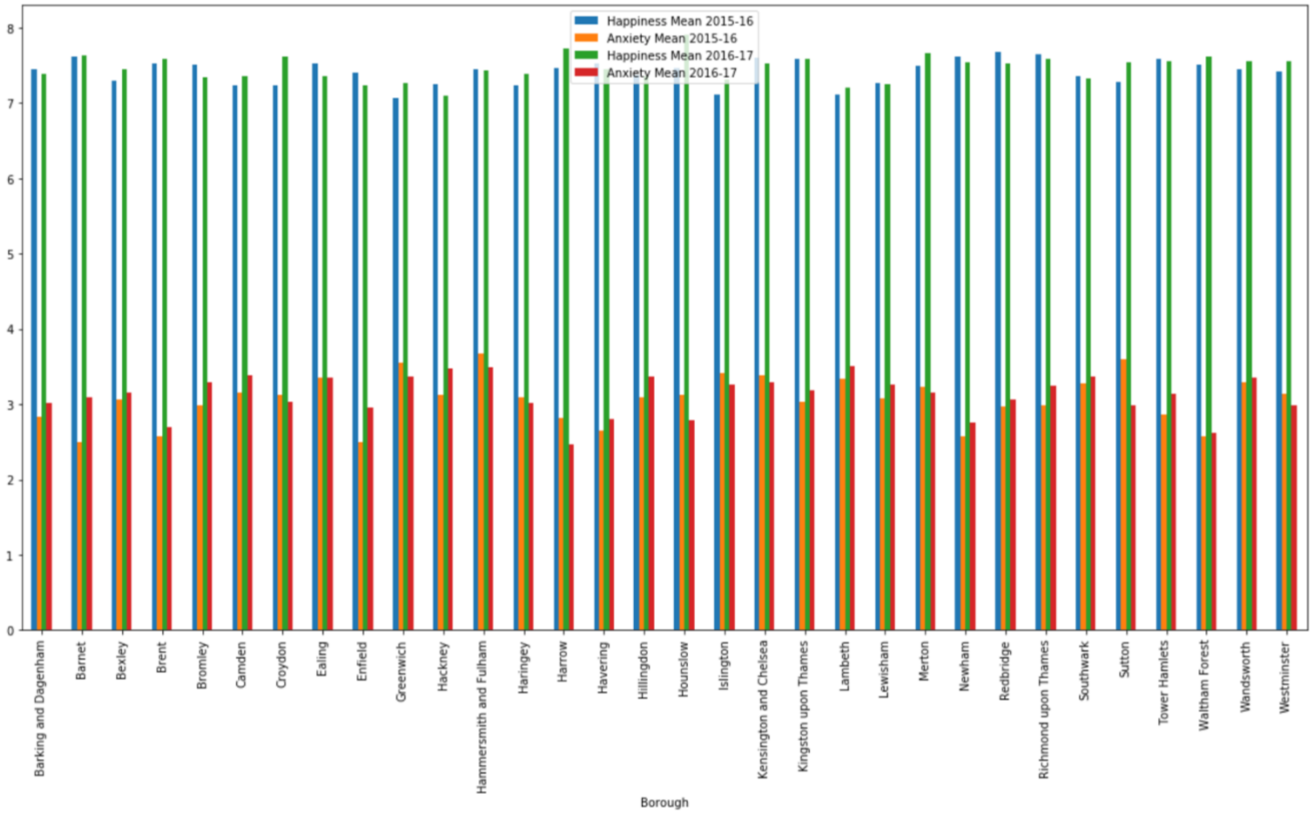
Top 10 boroughs – based on Walking 5x per week

	Walking_%	**Cycling_%**
Local Authority		
City of London	55	1
Waltham Forest	43	6
Southwark	43	8
Lambeth	43	7
Hammersmith and Fulham	43	9
Hackney	42	14
Westminster	42	6
Camden	42	4
Wandsworth	41	4
Kensington and Chelsea	41	5

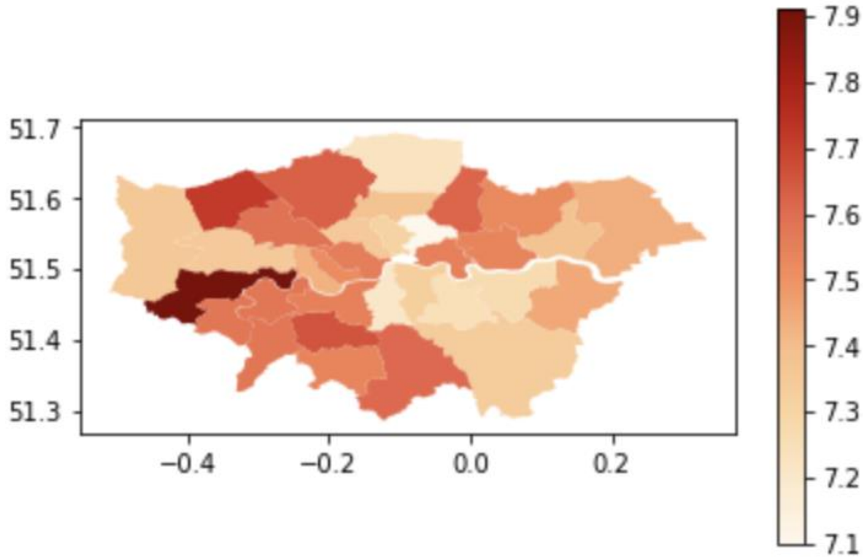
Mental Health

Mental health is one of the most important and often ignored statistic. This characterizes the mental health across 4 factors: Life Satisfaction, Worthwhile, Happiness and Anxiety and taking the mean across based on ratings out of 10. For this exercise, I have considered the happiness and anxiety measures with high happiness and low anxiety being the perfect measure for a mentally healthy population.

Bar chart



Heat map



Top 10 boroughs – highest Happiness scores

	Happiness Mean 2015-16	Anxiety Mean 2015-16	Happiness Mean 2016-17	Anxiety Mean 2016-17
Borough				
Hounslow	7.47	3.13	7.91	2.79
Harrow	7.46	2.82	7.72	2.47
Merton	7.50	3.23	7.66	3.15
Barnet	7.62	2.49	7.63	3.09
Waltham Forest	7.51	2.58	7.62	2.62
Croydon	7.23	3.12	7.61	3.03
Brent	7.53	2.57	7.59	2.69
Richmond upon Thames	7.65	2.99	7.58	3.25
Kingston upon Thames	7.59	3.03	7.58	3.19
Westminster	7.41	3.14	7.56	2.99

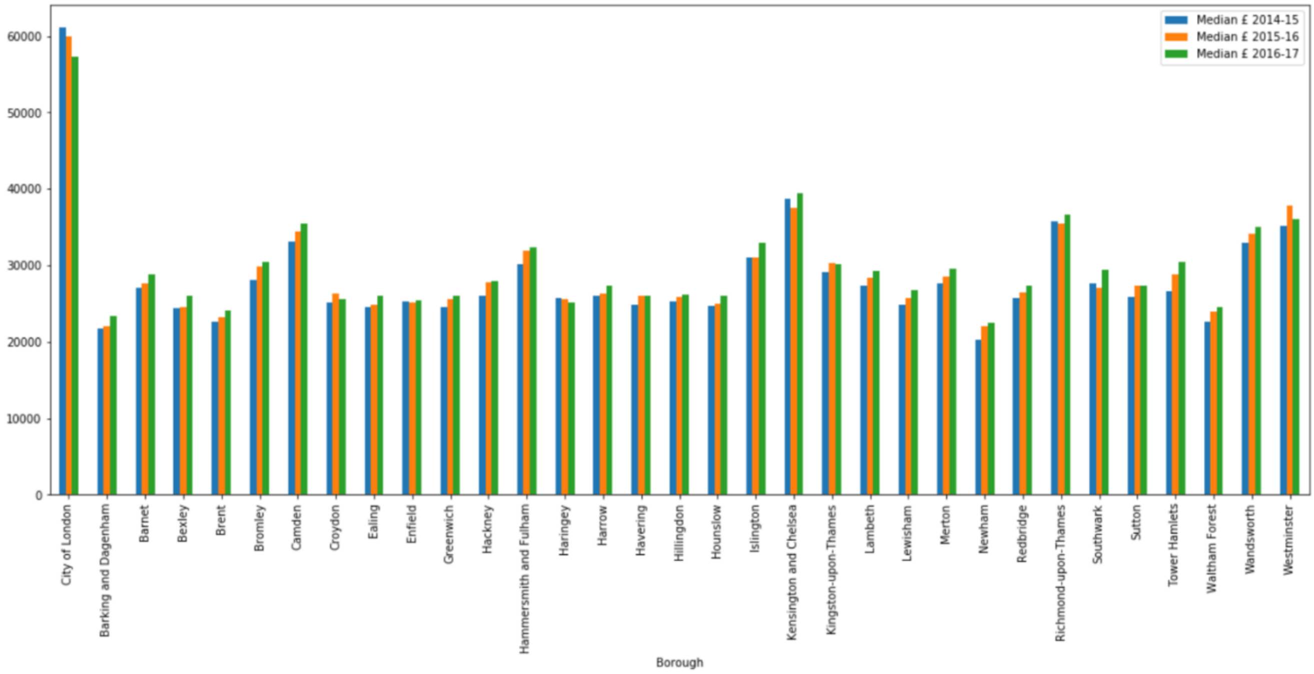
Socio-economic indicators

One of the important factors that make-up selection is the socio-economic snapshot of the boroughs. This can comprise of a variety of factors – but for the purpose of this exercise, I have considered the following:
Earnings of the working age population (Median Income)
Homeless population
Vehicles

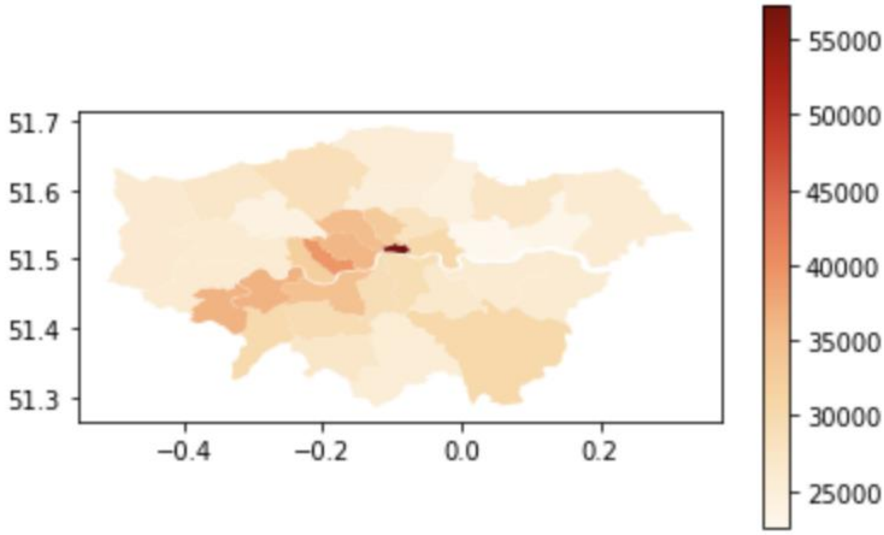
Earnings (Median Income)

The median income is one of the most frequently used statistic to compare the neighbourhoods in the makeup of the population. For the purpose of this exercise, I have considered the median income over the last 3 years.

Bar chart



Heat map



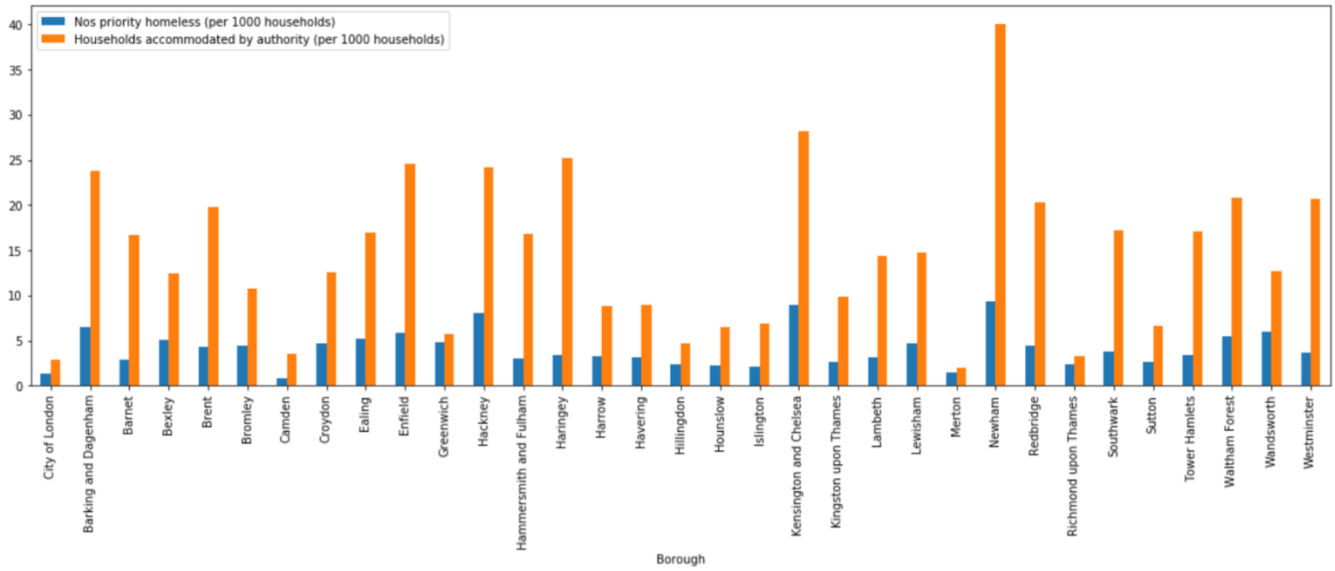
Top 10 boroughs – highest median income

	Median £ 2014-15	Median £ 2015-16	Median £ 2016-17
Borough			
City of London	61100	60000	57300
Kensington and Chelsea	38700	37500	39500
Richmond-upon-Thames	35800	35500	36600
Westminster	35100	37800	36100
Camden	33100	34400	35500
Wandsworth	32900	34200	35000
Islington	31000	31100	32900
Hammersmith and Fulham	30100	32000	32300
Tower Hamlets	26600	28900	30500
Bromley	28100	29800	30400

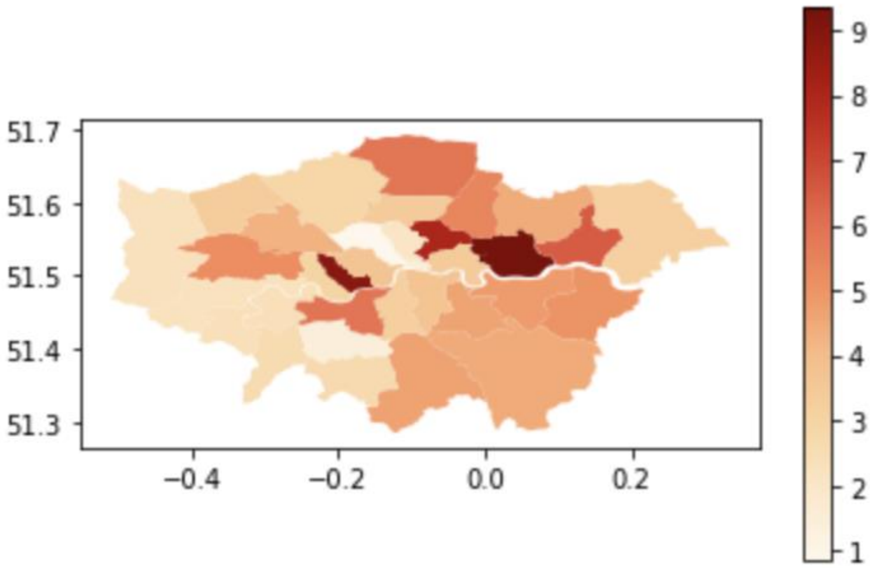
Homeless population

The disadvantage of any global city like London is the unavoidable migration of people from all walks of life to try and make it here to get access to better opportunities. This sometimes leads to homelessness and / or sleeping rough which Local Authorities or Boroughs try to alleviate through measures such as: Bed and Breakfast accommodation, Hostels, Local Authority council housing, Leasing from private sector and other types of housing. For the purpose of this exercise, I have considered two metrics: Number of priority homeless people per 1000 households and the households accommodated by the Local Authority per 1000 households.

Bar chart



Heat map



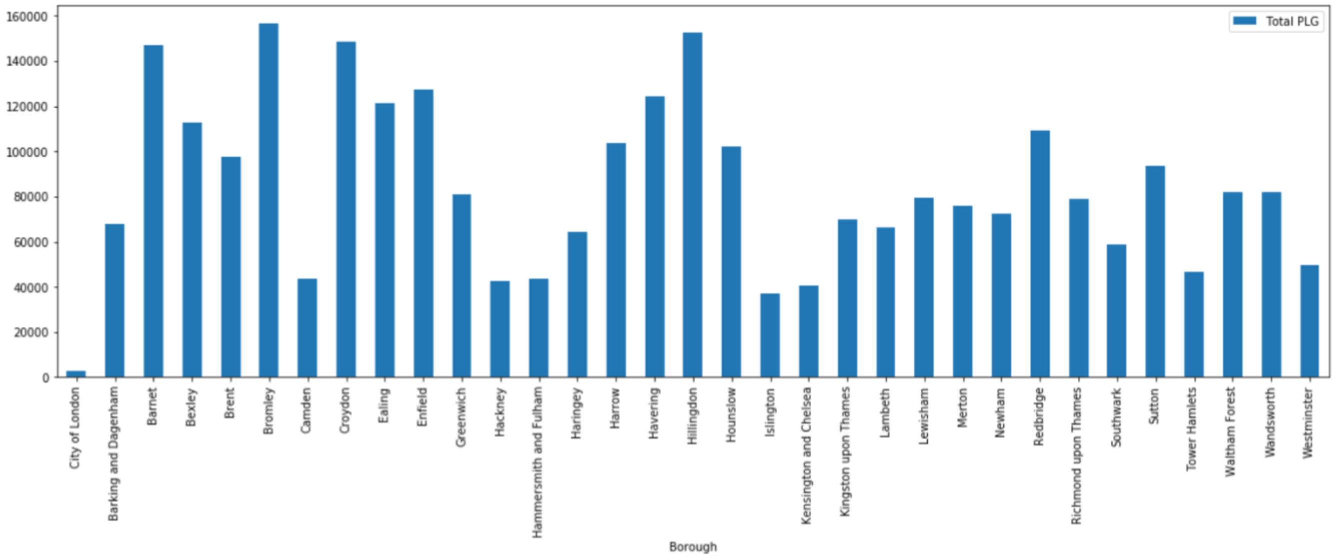
Top 10 boroughs – least numbers of priority homelessness

	Nos priority homeless (per 1000 households)	Households accommodated by authority (per 1000 households)
Borough		
Camden	0.84	3.53
City of London	1.38	2.95
Merton	1.44	1.94
Islington	2.06	6.90
Hounslow	2.29	6.51
Hillingdon	2.32	4.69
Richmond upon Thames	2.40	3.27
Kingston upon Thames	2.67	9.91
Sutton	2.69	6.64
Barnet	2.88	16.70

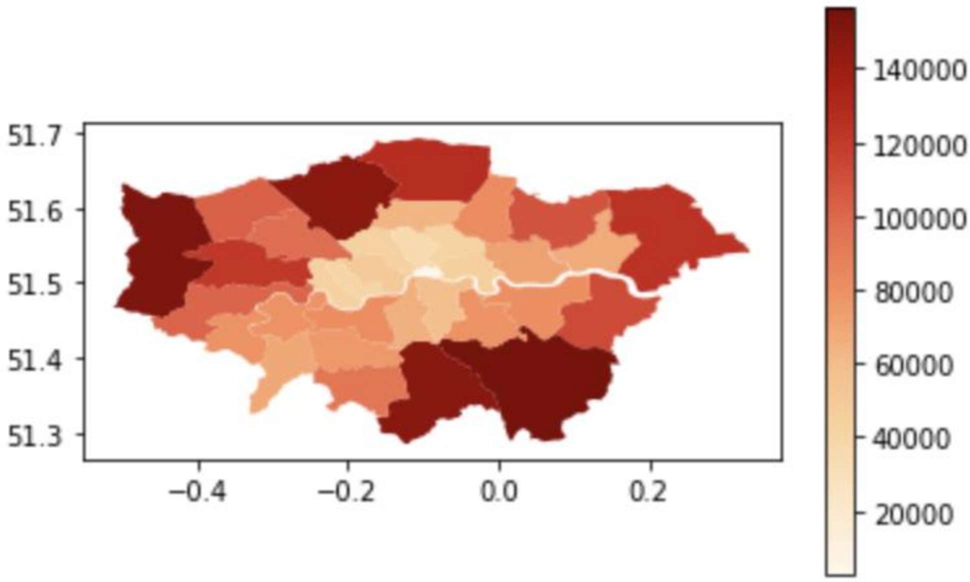
Vehicles

Vehicle ownership is a mixed criterion for analysis. On one hand the greater number of vehicles in a borough point to higher standard of living but on the other hand it may also disguise the lack of public transport accessibility as the reason for the number of vehicles. For this exercise, I have considered the total private or light goods vehicles (PLGs) in a borough and have ranked the boroughs based on the number of vehicles.

Bar chart



Heat map



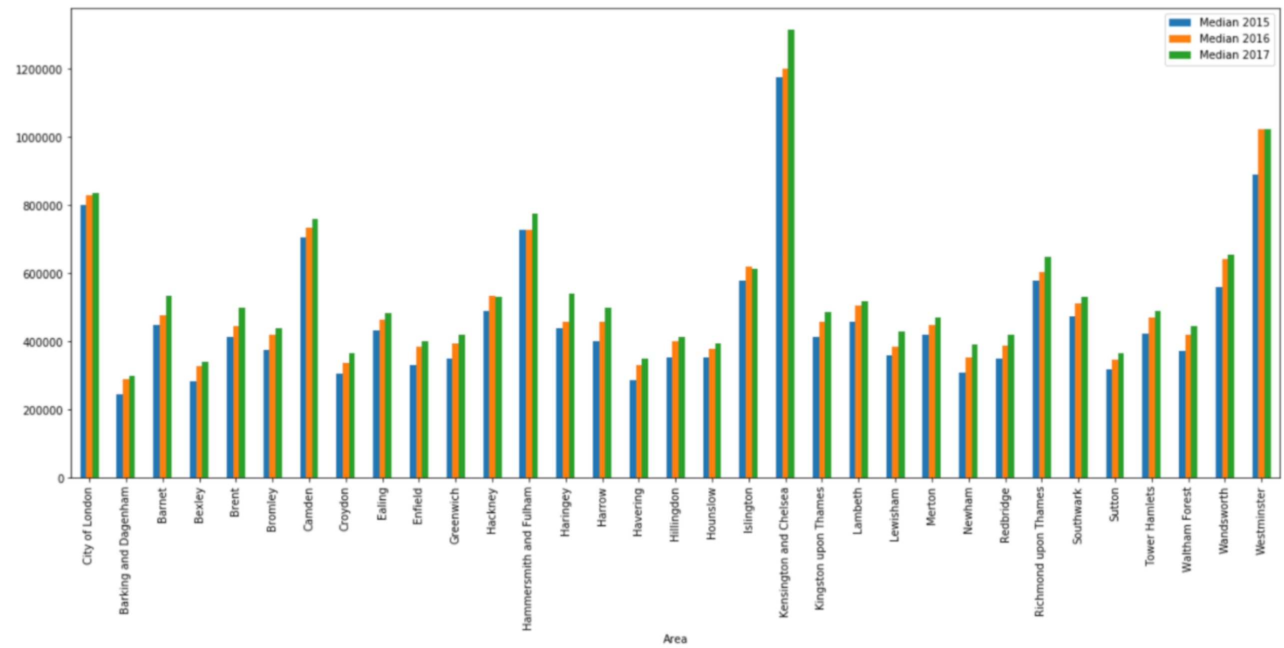
Top 10 boroughs – based on vehicle ownership

	Total PLG
Borough	
Bromley	156765
Hillingdon	152602
Croydon	148620
Barnet	147282
Enfield	127375
Havering	124564
Ealing	121257
Bexley	112563
Redbridge	109045
Harrow	103568

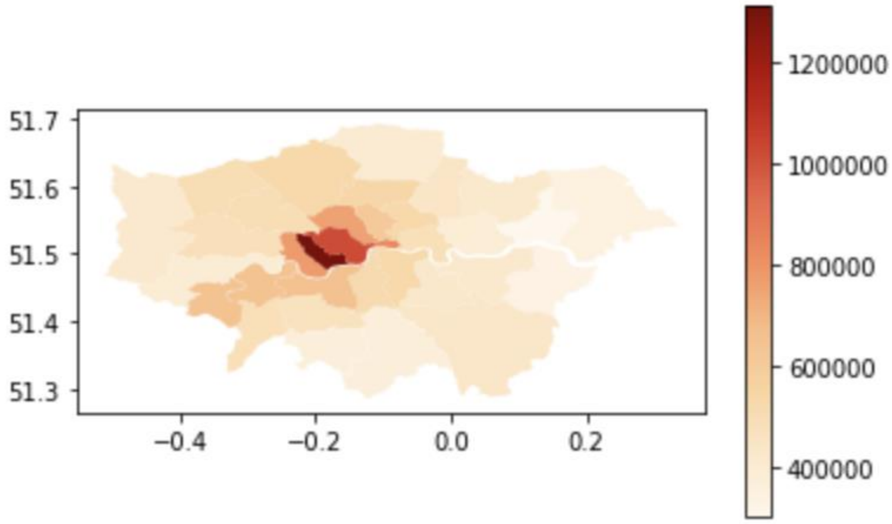
Economic Indicator

Ultimately the most important indicator for relocation for most people would be the house prices. Everyone would prefer a locality with all the good features around crime, education, socio-economic factors etc., but if the housing prices are not affordable then there is no further point to it. For this exercise, I have considered the median house prices for the last 3 year period and have ranked the boroughs based on the most affordable housing options.

Bar chart



Heat map



Top 10 boroughs – Least median house prices

	Value
Area	
Barking and Dagenham	300000
Bexley	342500
Havering	350000
Croydon	365000
Sutton	367000
Newham	390500
Hounslow	395000
Enfield	402500
Hillingdon	415000
Greenwich	420000

Further Data Analysis

Beyond the analysis of these categories and the associated datasets, I have also used FourSquare API to draw the location analysis of the boroughs. I have used Folium map to represent the different London boroughs based on the latitude and longitude. Further, I have taken 20 nearby venues at a radius of 350 metres from the centre of the borough and rank the top 5 most common venues for each borough. This is mainly to draw out wider perspective of the London boroughs in terms of amenities, avenues for entertainment and other such characteristics.

This is not fundamental in ranking the boroughs but can be taken as an additional layer of analysis on top of the basic analysis done earlier.

The details of the 5 most common type of venues from the 350m radius of the latitude and longitude of the boroughs are listed below:

	Borough	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Barnet	Café	Diner	Bus Stop	Women's Store	Cupcake Shop
1	Bexley	Sports Club	Lake	Fish & Chips Shop	Cycle Studio	Deli / Bodega
2	Brent	Food Truck	Fish & Chips Shop	Cycle Studio	Deli / Bodega	Diner
3	Bromley	Construction & Landscaping	Fish & Chips Shop	Cycle Studio	Deli / Bodega	Diner
4	Camden	Gastropub	Gym / Fitness Center	French Restaurant	Museum	Women's Store
5	City of London	Café	Coffee Shop	Hotel	Wine Bar	French Restaurant
6	Croydon	Auto Garage	Gym	South Indian Restaurant	Motorcycle Shop	Women's Store
7	Ealing	Fish & Chips Shop	Grocery Store	Pharmacy	Event Service	Women's Store
8	Enfield	Pub	Auto Workshop	Bus Stop	Fish & Chips Shop	Deli / Bodega
9	Greenwich	Park	Women's Store	Fish & Chips Shop	Deli / Bodega	Diner
10	Hackney	Hotel	Pharmacy	Bakery	Brazilian Restaurant	Pizza Place
11	Hammersmith and Fulham	Hotel	Ice Cream Shop	Gym	Chaat Place	Fish & Chips Shop
12	Haringey	Food Truck	Deli / Bodega	Coffee Shop	Bakery	Restaurant
13	Harrow	Health & Beauty Service	Bed & Breakfast	Food Truck	Deli / Bodega	Diner
14	Havering	Pub	Fish & Chips Shop	Cycle Studio	Deli / Bodega	Diner
15	Hillingdon	Soccer Field	Martial Arts Dojo	Women's Store	Fast Food Restaurant	Cycle Studio
16	Hounslow	Middle Eastern Restaurant	Indian Restaurant	Eastern European Restaurant	Bus Stop	Pizza Place
17	Islington	Café	Pub	Coffee Shop	French Restaurant	Middle Eastern Restaurant
18	Kensington and Chelsea	Café	Clothing Store	Women's Store	Burger Joint	Pharmacy
19	Kingston upon Thames	Indian Restaurant	Women's Store	Food Truck	Deli / Bodega	Diner
20	Lambeth	Cocktail Bar	Supermarket	Bus Stop	Women's Store	Fish & Chips Shop
21	Lewisham	Coffee Shop	Grocery Store	Fried Chicken Joint	Women's Store	Fast Food Restaurant
22	Merton	Grocery Store	Pizza Place	Bakery	Train Station	Park
23	Newham	Fish & Chips Shop	Coffee Shop	Grocery Store	Pharmacy	Fast Food Restaurant
24	Redbridge	Grocery Store	Supermarket	Pizza Place	Coffee Shop	Café
25	Richmond upon Thames	Italian Restaurant	Coffee Shop	Argentinian Restaurant	Restaurant	Gift Shop
26	Southwark	Photography Studio	Park	Lake	Café	Building
27	Sutton	Indian Restaurant	Grocery Store	Café	Women's Store	Fish & Chips Shop
28	Tower Hamlets	Go Kart Track	Recreation Center	Thrift / Vintage Store	Soccer Field	Women's Store
29	Waltham Forest	Concert Hall	Sports Bar	Gym	Bus Line	Deli / Bodega
30	Wandsworth	Bus Stop	Indian Restaurant	Grocery Store	Women's Store	Fish & Chips Shop
31	Westminster	Hotel	Sandwich Place	Hotel Bar	Coffee Shop	Gift Shop

Conclusion

As mentioned earlier, there is no one size fits all methodology to identify the best London borough. However, what I have done through this exercise is to use different categories and the associated datasets to draw out certain key factors for comparison against other boroughs and for ranking them. The suitability of the categories and the weightage given to each category would vary from person to person and hence the outcome of the most suitable London borough would also differ considerably.

However, I would like to give some pointers on my particular methodology in selecting the most suitable London borough to relocate. From looking at almost all the datasets and the top 10 boroughs, the boroughs of Richmond upon Thames and Kingston upon Thames have some of the most desirable traits that I would value. For e.g. they have the least crime rate, more highly qualified populace, quite high performance in educational indicators like GCSE, KS1 etc., low childhood obesity and smoking rates, high median incomes, a smaller number of priority households, quite active population and most importantly house prices are not significantly higher to the average house prices unlike boroughs like Kensington and Chelsea.

There are certain factors where they do not score highly – for e.g. public transport accessibility, vehicle ownership etc., however no borough is likely to be in top for all factors. So taking that into consideration, I would vote for **Richmond upon Thames** as the number one London borough followed by **Kingston upon Thames** as the second choice.

Further Analysis

This exercise can be extended and analyzed further in a myriad of ways which are beyond the scope of this activity. Some of the following could be:

1. Analyzing the correlation of the key factors through linear or polynomial regression to identify which factors strongly impact the house price or other independent variables
2. Clustering the boroughs into a number of clusters featuring some common characteristic so that if a particular borough is not suitable, then another borough in the same cluster could be chosen as an alternative
3. Creating a weightage map with key factors and weights so as to create a unique final value for each borough that corresponds to the final ranking.
4. Analyzing further into specific ward and street level detail of the particular borough once that is identified to reduce the search area to find the most suitable property.

There can be quite a lot of other analysis that can be done given the rich nature of the data that is available for London.