



# Capstone Project - The Battle of Neighborhoods Report written by Eric Damalie

# Business Problem

- ▶ I have been approached by a marketing company employed by the London NHS trust (Stakeholder) to produce a report about London to help sell the City to potential foreign doctors and nurses who may be considering coming to work in the UK. This marketing material is to help entice these potential health workers (Target Audience) to choose settling and working in London and also help decide which London Borough to settle

# Few facts about London

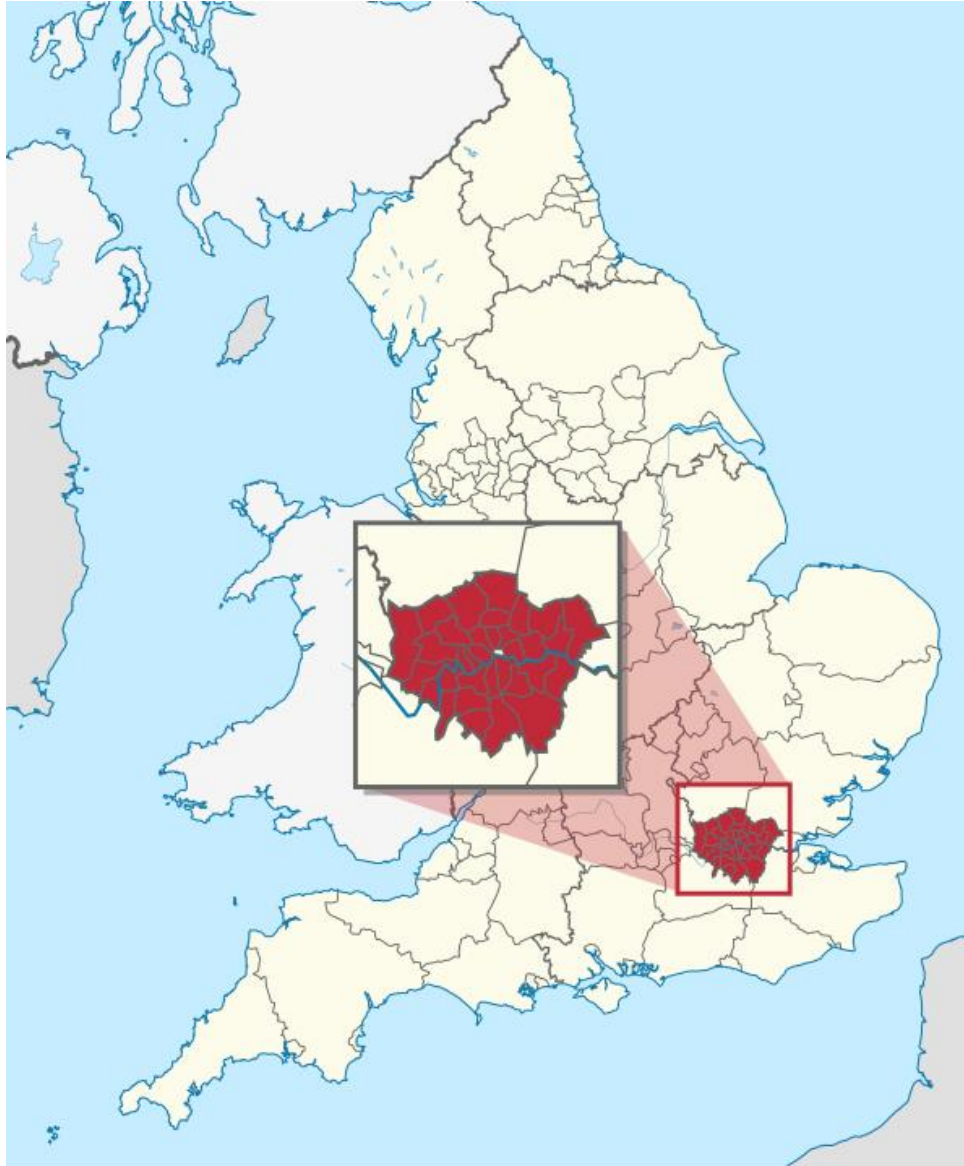
London is the capital and largest city of England and the United Kingdom

London is considered to be one of the world's most important global cities

London ranks 26 out of 300 major cities for economic performance.

It is one of the largest financial centers and has either the fifth or sixth largest metropolitan area GDP

It is the most-visited city as measured by international arrivals and has the busiest city airport system as measured by passenger traffic.



# Objectives

To explore London to show what London has got to offer our target audience in terms of venues.

To show the crime rates of each borough.

To explore the various boroughs to help the potential NHS health worker to decide on which neighborhood he/she would like to settle based on the venues or amenities important to him/her.

# Data Acquisition and Preprocessing

- ▶ In this project, I will be using the following datasets to help solve my problem:
- ▶ The London borough data was scrapped from London Borough Wikipedia page
- ▶ Foursquare and geopy data to list various selected venues that I believe will be important for our target audience.
- ▶ The downloaded crime data for London was from [https://data.london.gov.uk/dataset/recorded\\_crime\\_summary](https://data.london.gov.uk/dataset/recorded_crime_summary)
- ▶ I will then filter the top 10 most common venue types of each neighborhoods.
- ▶ The target audience will be given the top ten venues of each neighborhood so they could look at them in determining which neighborhood/borough to settle
- ▶ Neighborhoods or Boroughs will be clustered into different groups based on common attributes.



# Methodology

- ▶ The crime data will be trimmed and converted from number of crimes per borough per month to crimes per 1000 people per month for better comparison.
- ▶ Foursquare API will be used to retrieve all venues of each neighborhoods.
- ▶ The top 10 most common venue types of each neighborhoods or Borough will be filtered.
- ▶ The neighborhoods will be then put in different clusters based on similar or common attributes or venues using K-means clustering
- ▶ The target audience will be shown the top ten venues of each neighborhood.
- ▶ The target audience will be shown the crime levels of all the boroughs

# Recorded London Crime Data for the past 24 months. Table 1

]:

	MajorText	MinorText	Borough	Sum	201801	201802	201803	201804	201805	201806	...	201903	201904	201905	201906	201907	201908	201909	201910	201911	201912
0	Arson and Criminal Damage	Arson	Barking and Dagenham	125	2	3	6	3	4	12	...	5	5	11	3	5	3	6	9	8	5
1	Arson and Criminal Damage	Criminal Damage	Barking and Dagenham	2818	135	111	115	122	126	123	...	138	130	138	113	134	118	109	109	100	114
2	Burglary	Burglary - Business and Community	Barking and Dagenham	748	38	33	38	36	24	33	...	29	27	22	27	31	35	37	30	30	24
3	Burglary	Burglary - Residential	Barking and Dagenham	2472	143	134	122	75	93	77	...	99	96	114	96	71	67	80	97	114	131
4	Drug Offences	Drug Trafficking	Barking and Dagenham	145	4	5	7	3	8	6	...	6	4	8	6	7	6	6	9	10	2

5 rows x 28 columns


◀ ▶



:

	Borough	Population	Coordinates
0	Barking and Dagenham [note 1]	194,352	51°33'39"N 0°09'21"E / 51.5607°N 0.1557°E / ...
1	Barnet	369,088	51°37'31"N 0°09'06"W / 51.6252°N 0.1517°W / ...
2	Bexley	236,687	51°27'18"N 0°09'02"E / 51.4549°N 0.1505°E / ...
3	Brent	317,264	51°33'32"N 0°16'54"W / 51.5588°N 0.2817°W / ...
4	Bromley	317,899	51°24'14"N 0°01'11"E / 51.4039°N 0.0198°E / ...

Table 2: Top 5 rows of coordinates information obtain from scrapping Wikipedia.



	Borough	Population	Latitude	Longitude
0	Barking and Dagenham	194352	51.5607	0.1557
1	Barnet	369088	51.6252	-0.1517
2	Bexley	236687	51.4549	0.1505
3	Brent	317264	51.5588	-0.2817
4	Bromley	317899	51.4039	0.0198

After some data cleaning and wrangling the information in table 2 was converted into Table 3 below.

Table 4 show the combine information of the cleaner versions of the coordinate, population and crime data

	Borough	MonthlyAvg	Population	Latitude	Longitude	CrimeToPop
0	Barking and Dagenham	1601.750000	194352.0	51.5607	0.1557	8.241490
1	Barnet	2486.500000	369088.0	51.6252	-0.1517	6.736876
2	Bexley	1389.041667	236687.0	51.4549	0.1505	5.868686
3	Brent	2541.333333	317264.0	51.5588	-0.2817	8.010153
4	Bromley	2009.333333	317899.0	51.4039	0.0198	6.320666
5	Camden	3133.208333	229719.0	51.5290	-0.1255	13.639309
6	Croydon	2692.958333	372752.0	51.3714	-0.0977	7.224531
7	Ealing	2496.041667	342494.0	51.5130	-0.3089	7.287841
8	Enfield	2429.541667	320524.0	51.6538	-0.0799	7.579906
9	Greenwich	2267.666667	264008.0	51.4892	0.0648	8.589386
10	Hackney	2724.000000	257379.0	51.5450	-0.0553	10.583614
11	Hammersmith and Fulham	1915.083333	178685.0	51.4927	-0.2339	10.717650
12	Haringey	2633.416667	263386.0	51.6000	-0.1119	9.998317
13	Harrow	1342.750000	243372.0	51.5898	-0.3346	5.517274
14	Havering	1574.958333	242080.0	51.5812	0.1837	6.505942
15	Hillingdon	2176.458333	286806.0	51.5441	-0.4760	7.588608

# Graph 1: Crime rates of the London borough

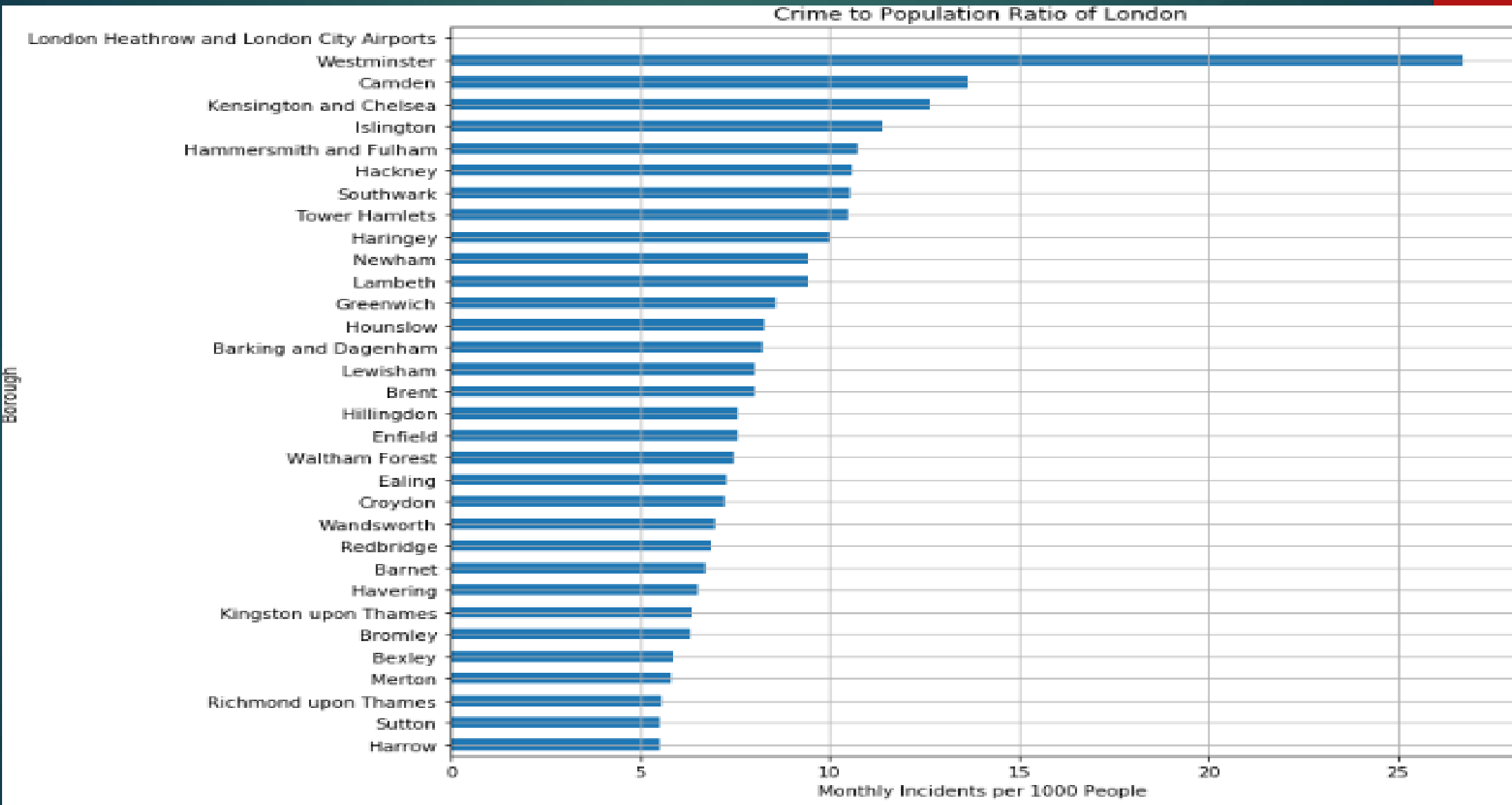


Table 5: Popular venues in and around London boroughs (Top 5 rows of the Table)

	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	10th Most Common Venue
0	Barking and Dagenham	Martial Arts Dojo	Pool	Bus Station	Park	Supermarket	Gym / Fitness Center	Golf Course	Yoga Studio	Fish & Chips Shop	Field
1	Barnet	Café	Bus Stop	Business Service	Yoga Studio	French Restaurant	Flea Market	Fish Market	Fish & Chips Shop	Field	Fast Food Restaurant
2	Bexley	Clothing Store	Coffee Shop	Pub	Italian Restaurant	Supermarket	Fast Food Restaurant	Pharmacy	Portuguese Restaurant	Bakery	Hotel
3	Brent	Coffee Shop	Hotel	Clothing Store	Sporting Goods Shop	Grocery Store	American Restaurant	Sandwich Place	Multiplex	Bubble Tea Shop	Burger Joint
4	Bromley	Clothing Store	Coffee Shop	Bar	Burger Joint	Pizza Place	Gym / Fitness Center	Furniture / Home Store	Burrito Place	Sandwich Place	Café

# Clustering

- ▶ By using the Foursquare data and K-means clustering, the different boroughs were grouped into different clusters based on similarities of their most popular venues. More detailed steps including one hot coding can be found in my data section. After the clustering was done, the boroughs or neighborhoods were put into 5 clusters i.e. 0 to 4. The tables below show the different boroughs and their clusters.



## Cluster 0: showing the borough and top 10 popular venue

	Borough	CrimeToPop	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
16	Hounslow	8.291668	0.0	Bed & Breakfast	Park	Metro Station	Chinese Restaurant	Café	Yoga Studio	Flea Market	Fish Market	Fish & Chips Shop	Field

# Cluster 1: showing the boroughs and the top 10 popular venues

	Borough	CrimeToPop	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
3	Brent	8.010153	1.0	Coffee Shop	Hotel	Clothing Store	Sporting Goods Shop	Grocery Store	American Restaurant	Sandwich Place	Multiplex	Bubble Tea Shop	Burger Joint
5	Camden	13.639309	1.0	Hotel	Café	Coffee Shop	Italian Restaurant	Hotel Bar	Pizza Place	Train Station	Breakfast Spot	Bakery	Modern European Restaurant
12	Haringey	9.998317	1.0	Café	Fast Food Restaurant	Italian Restaurant	Pharmacy	Sandwich Place	Pub	Portuguese Restaurant	Hotel Bar	Convenience Store	Park
18	Kensington and Chelsea	12.621524	1.0	Café	Gym / Fitness Center	English Restaurant	Burger Joint	French Restaurant	Bakery	Juice Bar	Clothing Store	Restaurant	Modern European Restaurant
20	Lambeth	9.416575	1.0	Caribbean Restaurant	Market	BBQ Joint	Beer Bar	Pub	Tapas Restaurant	Restaurant	Burger Joint	Indian Restaurant	Cocktail Bar
23	Merton	5.798843	1.0	Italian Restaurant	Supermarket	Park	Café	Bakery	Sandwich Place	Fast Food Restaurant	Garden Center	Burger Joint	Bar
27	Southwark	10.535330	1.0	Coffee Shop	Pub	Hotel	Art Gallery	English Restaurant	Hotel Bar	Bar	Theater	Gym / Fitness Center	Shopping Mall
29	Tower Hamlets	10.505301	1.0	Italian Restaurant	Pizza Place	Hotel	Coffee Shop	Gym / Fitness Center	Outdoor Sculpture	Chinese Restaurant	Grocery Store	Steakhouse	Asian Restaurant

## Cluster 2: Showing the borough and the top 10 venues

	Borough	CrimeToPop	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
1	Barnet	6.736876	2.0	Café	Bus Stop	Business Service	Yoga Studio	French Restaurant	Flea Market	Fish Market	Fish & Chips Shop	Field	Fast Food Restaurant

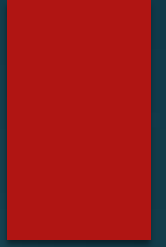
# Cluster 3: showing the borough and the top 10 venue

	Borough	CrimeToPop	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
0	Barking and Dagenham	8.241490	3.0	Martial Arts Dojo	Pool	Bus Station	Park	Supermarket	Gym / Fitness Center	Golf Course	Yoga Studio	Fish & Chips Shop	Field
2	Bexley	5.868686	3.0	Clothing Store	Coffee Shop	Pub	Italian Restaurant	Supermarket	Fast Food Restaurant	Pharmacy	Portuguese Restaurant	Bakery	Hotel
4	Bromley	6.320666	3.0	Clothing Store	Coffee Shop	Bar	Burger Joint	Pizza Place	Gym / Fitness Center	Furniture / Home Store	Burrito Place	Sandwich Place	Café
6	Croydon	7.224531	3.0	Pub	Coffee Shop	Indian Restaurant	Gym / Fitness Center	Asian Restaurant	Portuguese Restaurant	Italian Restaurant	Brewery	Burger Joint	Burrito Place
7	Ealing	7.287841	3.0	Coffee Shop	Italian Restaurant	Hotel	Vietnamese Restaurant	Park	Burger Joint	Pub	Pizza Place	Bar	Portuguese Restaurant
8	Enfield	7.579906	3.0	Coffee Shop	Clothing Store	Pharmacy	Department Store	Shopping Mall	Sandwich Place	Restaurant	Pub	Supermarket	Optical Shop
9	Greenwich	8.589386	3.0	Supermarket	Coffee Shop	Clothing Store	Pub	Plaza	Grocery Store	Fast Food Restaurant	Discount Store	Sandwich Place	African Restaurant
10	Hackney	10.583614	3.0	Pub	Café	Coffee Shop	Cocktail Bar	Bakery	Brewery	Hotel	Vietnamese Restaurant	Clothing Store	Grocery Store
11	Hammersmith and Fulham	10.717650	3.0	Pub	Indian Restaurant	Italian Restaurant	Gastropub	Café	Japanese Restaurant	Clothing Store	German Restaurant	Portuguese Restaurant	Polish Restaurant

## Cluster 4: Showing the borough and the top 10 venues

	Borough	CrimeToPop	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
24	Newham	9.427102	4.0	Hotel	Light Rail Station	Pharmacy	Sandwich Place	Rafting	Chinese Restaurant	Airport	Airport Service	Dance Studio	Dumpling Restaurant

# Results Section





## Cluster 0

- ▶ As can be seen from the Cluster 0 (Healthy area) table, there is only one Borough which is Hounslow in this Cluster.
- ▶ Some of the most common venues in Hounslow are the metro station, café and the park.
- ▶ There is a yoga studio within the top ten common venues and the crime rate is relatively low at 8.3.
- ▶ Based on the most common venues in this cluster, this cluster could be described as an average area

# Cluster 1

- ▶ Some of the Boroughs found in cluster 1 include Camden, Kensington and Chelsea, Westminster and Haringey.
- ▶ Common venues in this cluster include various kinds of restaurants, gym/fitness centers, coffee shops, pubs and grocery shops.
- ▶ This cluster can be seen as a cluster of boroughs with lots of lively venues hence this cluster could be described as a lively area.
- ▶ There is a massive difference between the crime rates in this cluster with crime rates ranging from 5.7 in Merton to 26.7 in Westminster

## Cluster 2

- ▶ Like cluster 0, there is only one Borough found in Cluster 2.
- ▶ Popular venues includes cafe, business service area, flea and fish markets.
- ▶ The crime rate is among the lowest in London at 6.7.
- ▶ There seem to be lots of business and trading activities taken place in this borough hence describing this cluster as a business area

## Cluster 3

- ▶ Cluster 3 is made up of are Bexley, Barking and Dagenham, Brent, Bromley and Croydon Boroughs just to mention a few.
- ▶ Cluster 3 is the largest Cluster which is made up of 20 Boroughs. The common venues in this cluster include coffee shops, different types of restaurants, clothing store, bars and hotels.
- ▶ The crime rate range is not as wide as that for cluster 2, ranging from 5.5 in Sutton and Richmond up Thames and 11.4 in Islington.
- ▶ It is safe to say that this cluster is fairly safe as the crime rate of 11.4 in Islington is half the crime rate of Westminster in cluster 2.
- ▶ Like Cluster 2, based on the common venues found in this Cluster, Cluster 3 could also be described as a lively area.

## Cluster 4

- ▶ The only Borough in this cluster is Newham with a crime rate of 9.4. The most popular venues in Newham include hotels, airport services, duty free shop, light rail station and an airport. It can be deduce that Newham is near a London airport.

# Discussion Section

- ▶ The venues obtained from exploring the data for London boroughs, shows how lively and busy the city is as majority of the boroughs explored showed lots of lively and busy venues. When it comes to crime rates in the City, it could be very high in Westminster and low in Merton and Sutton
- ▶ At the end of this project, London boroughs have been explored by using Foursquare API to give a picture about how each borough looks like in terms of the types and number of venues that exist in that borough
- ▶ The objective of my task was not to choose a Neighborhood or Borough in London for my target audience but rather show them what different parts of London has to offer and depending on their taste, they choose where to settle after making the decision to move to London.



# Conclusion

- ▶ In conclusion I believe the target audience in this case the potential NHS worker will have some information about the various neighborhoods/ boroughs in London. The crime rate data will also be vital in determining where he/she would like to settle. This project was done to provide some information about London to help the target audience make a decision and not done to choose settlement location for the target audience.

# Conclusion

- ▶ No matter the individual's preference, one thing all humans desire is safety as a result I will say the safest boroughs to live in London are Sutton and Harrow in cluster 3 and Richmond upon Thames in cluster 1 with crime rates of about 5.5 crimes per 1000 people per month. The two most dangerous places to live in London are Camden and Westminster both in cluster 1 with crime rates of 13.6 and 26.7 per 1000 people per month respectively.