Capstone Project
Battle of Neighborhoods
London

Introduction

With Covid-19 and Brexit firmly taken route in the UK, it is imperative to sell London as an international city which is safe for skilled foreign immigrants to come and settle. Especially those who will work for the NHS. The Covid19 crisis has shown how fractured our NHS and society is since the Brexit Vote. With a large out flowing of foreign professionals going back to their country of origin.

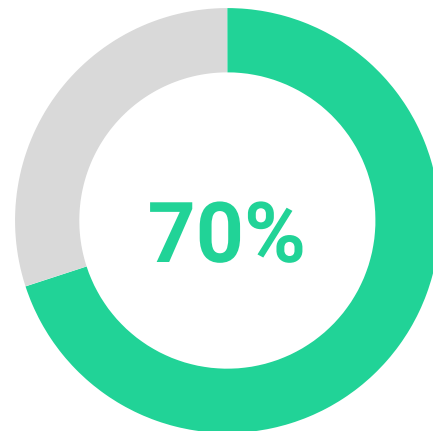In this presentation we set out what boroughs are the best to live in

Data Scrapping and Cleansing

In this project, I will be using the following datasets to help solve my problem:

- London Borough and Coordinates data scrapped from Wikipedia page

- Mapping using foursquare

- Crime data from:

https://data.london.gov.uk/dataset/recorded_crime_summary

70%

70% of time is taken analysing and cleansing the data

Data Validation and Visualization

Summing most recent 24 months of London Crime data analysed via csv

Data was scrapped from Wikipedia using beautiful soup in python. Extracting unstructured data into a structured form merging both data sets

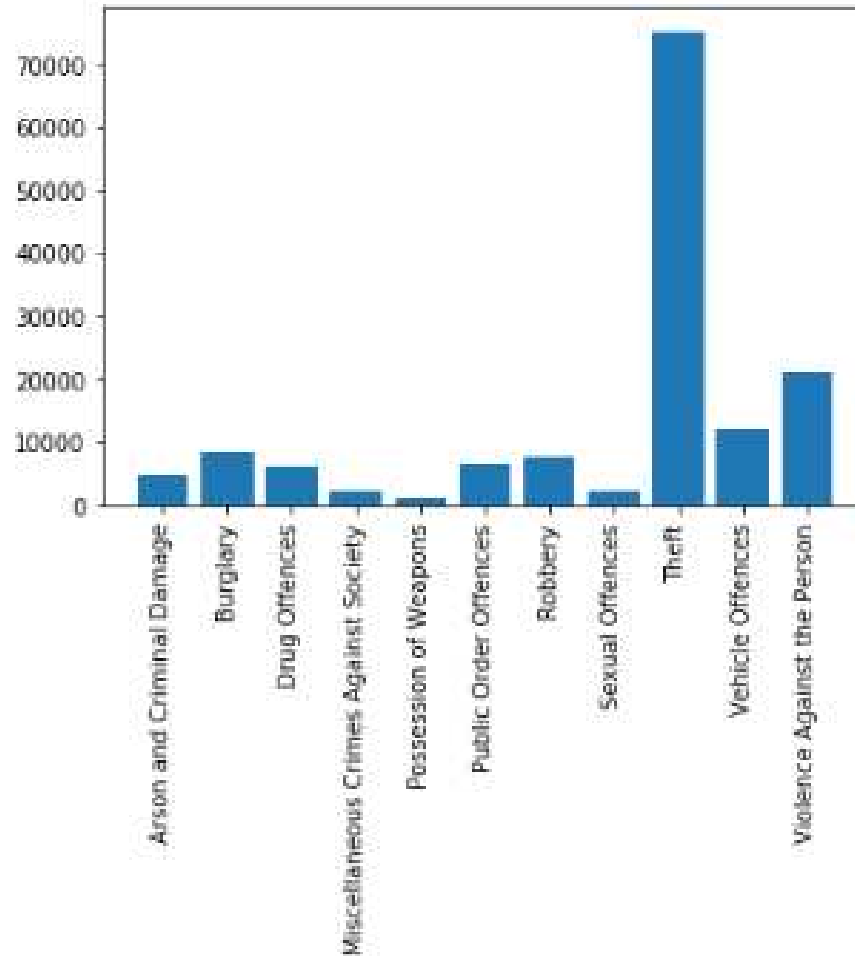K means clustering was used to cluster similar boroughs

Finally graphical and mapping visualization as carried out
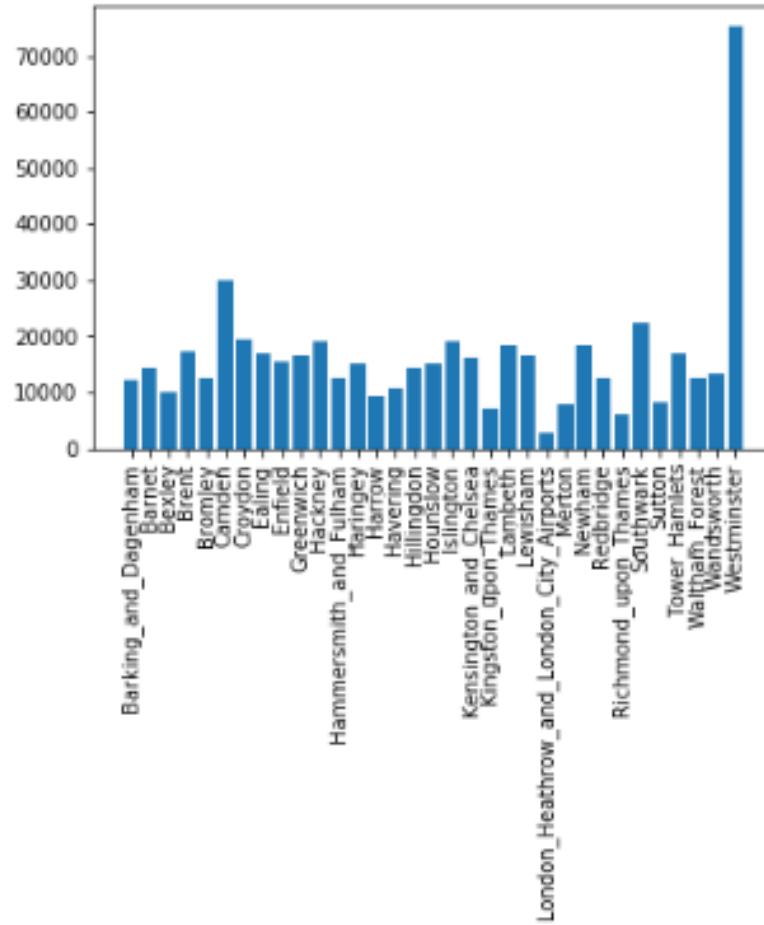
Methodology

- Using Foursquare API to retrieve all venues of each neighborhoods.
- Analyse crime rates
- Segmenting boroughs using k-means clustering
- Top 10 most common venue types of each neighborhoods or Borough.
- The neighborhoods will be then put in different clusters based on similar or common atributes using K-means clustering.
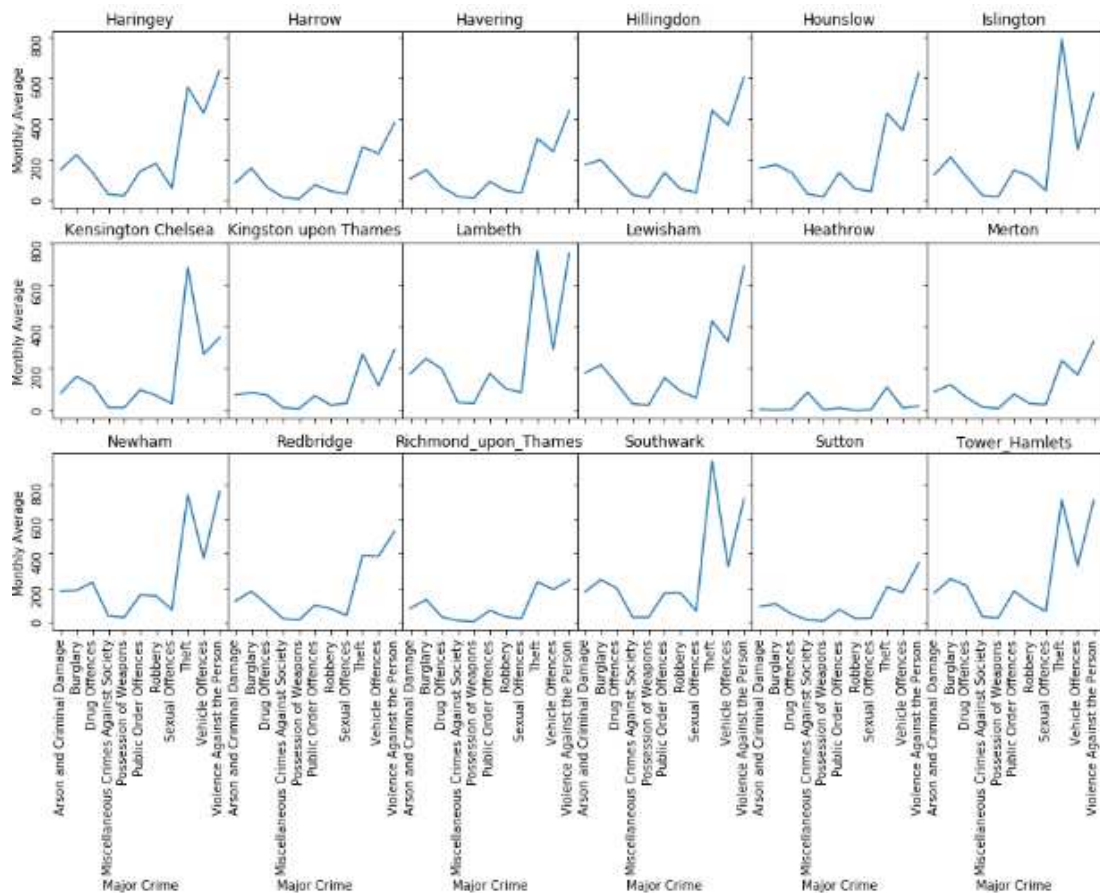
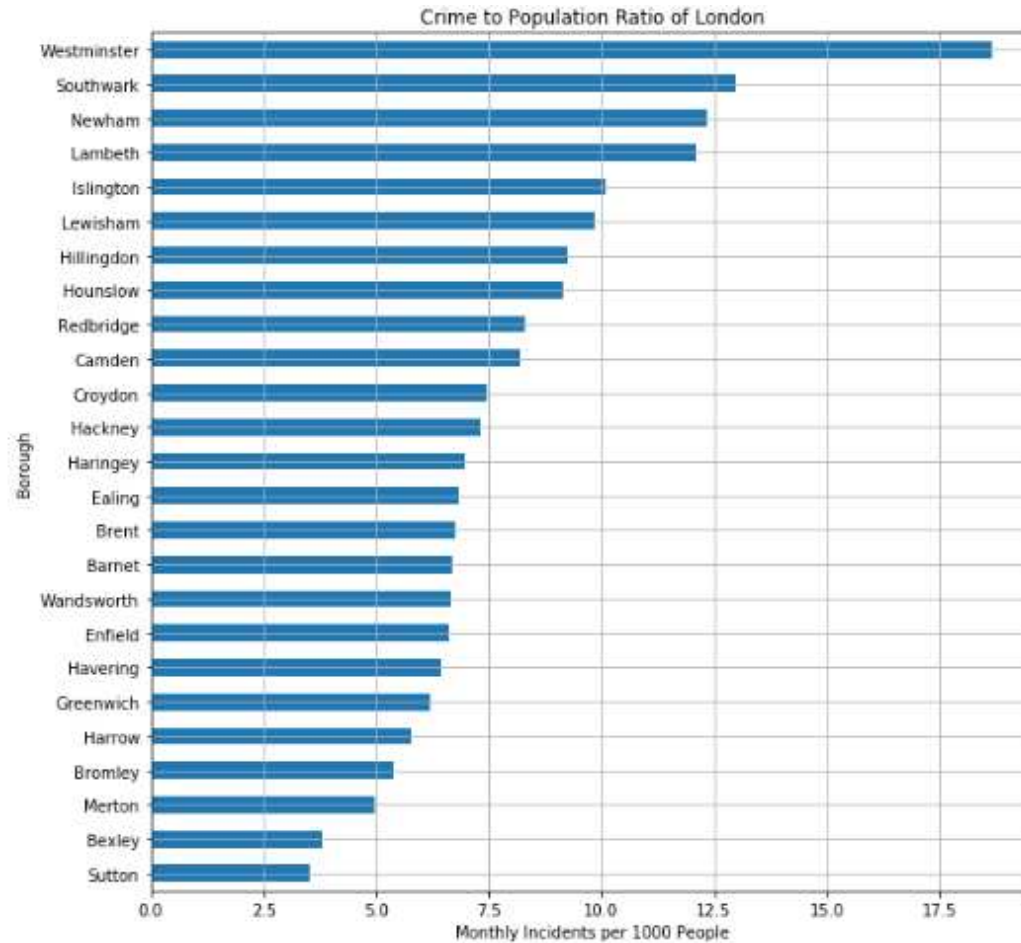# London Crime Breakdown by Type of Crime

Total Crime Breakdown - Borough

Total Crime Breakdown - Borough

Data Scraping - Coordinates

	Borough	MonthlyAvg	Population	Latitude	Longitude	CrimeRate
0	Barnet	2467.208333	369088	51.6252	-0.1517	6.684607
1	Bexley	1403.833333	236687	51.4549	0.1505	3.803519
2	Brent	2496.708333	317264	51.5588	-0.2817	6.764534
3	Bromley	1983.833333	317899	51.4039	0.0198	5.374960
4	Camden	3031.833333	229719	51.5290	-0.1255	8.214392
5	Croydon	2754.875000	372752	51.3714	-0.0977	7.464006
6	Ealing	2521.750000	342494	51.5130	-0.3089	6.832381
7	Enfield	2443.666667	320524	51.6538	-0.0799	6.620824
8	Greenwich	2285.458333	264008	51.4892	0.0648	6.192177
9	Hackney	2705.958333	257379	51.5450	-0.0553	7.331472
10	Haringey	2577.875000	263386	51.6000	-0.1119	6.984445
11	Harrow	1369.000000	243372	51.5898	-0.3346	5.784010
12	Havering	1527.958333	242080	51.5812	0.1837	6.455607
13	Hillingdon	2186.875000	286806	51.5441	-0.4760	9.239523
14	Hounslow	2167.291667	262407	51.4746	-0.3680	9.156784
15	Islington	2390.291667	215667	51.5416	-0.1022	10.098956
16	Lambeth	2862.541667	314242	51.4607	-0.1163	12.094207
17	Lewisham	2329.458333	286180	51.4452	-0.0209	9.841936
18	Merton	1171.416667	203223	51.4014	-0.1958	4.949223
19	Newham	2924.250000	318227	51.5077	0.0469	12.354924
20	Redbridge	1966.541667	288272	51.5590	0.0741	8.308617
21	Southwark	3071.916667	298464	51.5035	-0.0804	12.978814
22	Sutton	1120.333333	195914	51.3618	-0.1945	3.531234
23	Wandsworth	2115.000000	310516	51.4567	-0.1910	6.666372
24	Westminster	5920.666667	226841	51.4973	-0.1372	18.661640

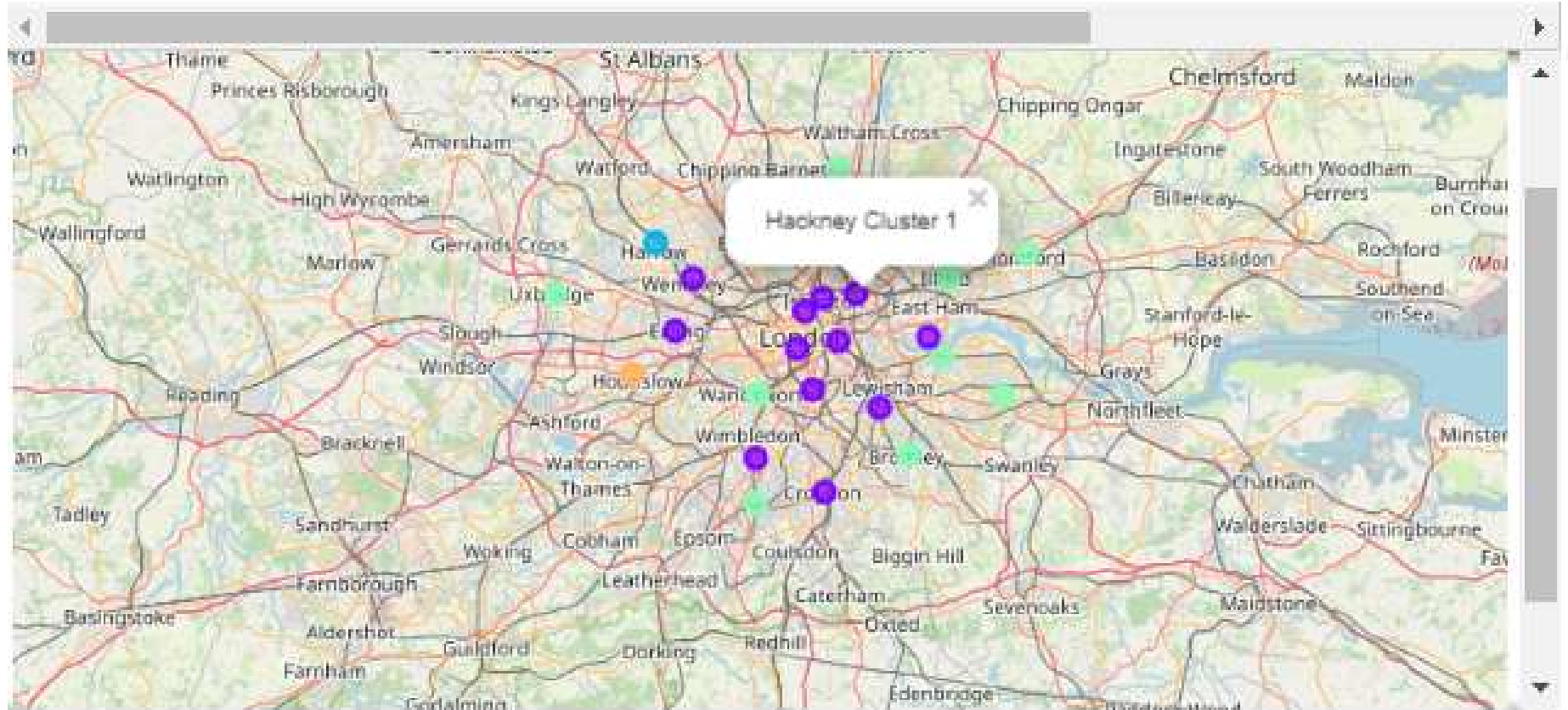# Crime to Population Ratio

Crime to Population Ratio of London

k means Clustering

Borough	MonthlyAvg	Population	Latitude	Longitude	CrimeRate	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
Barnet	2467.208333	369088	51.6252	-0.1517	6.684607	0	Café	Bus Stop	Business Service	Yoga Studio
Bexley	1403.833333	236687	51.4549	0.1505	3.803519	3	Coffee Shop	Pub	Clothing Store	Supermarket
Brent	2496.708333	317264	51.5588	-0.2817	6.764534	1	Coffee Shop	Hotel	Clothing Store	Grocery Store
Bromley	1983.833333	317899	51.4039	0.0198	5.374960	3	Clothing Store	Coffee Shop	Burger Joint	Pizza Place
Camden	3031.833333	229719	51.5290	-0.1255	8.214392	1	Café	Coffee Shop	Hotel	Burger Joint

FourSquare Map

Result - Segment 0

	Borough	CrimeRate	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 Mo Commo Venu
0	Barnet	6.684607	0	Café	Bus Stop	Business Service	Yoga Studio	Electronics Store	Fish Market	Film Studio	Fast Food Restaurant	Farme Mark

Result - Segment 1

	Borough	CrimeRate	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
2	Brent	6.764534	1	Coffee Shop	Hotel	Clothing Store	Grocery Store	Sporting Goods Shop	Bar	American Restaurant
4	Camden	8.214392	1	Café	Coffee Shop	Hotel	Burger Joint	Garden	Breakfast Spot	Deli / Bodega
5	Croydon	7.464006	1	Pub	Coffee Shop	Malay Restaurant	Food Court	Supermarket	Strip Club	Gaming Cafe
6	Ealing	6.832381	1	Coffee Shop	Italian Restaurant	Pizza Place	Clothing Store	Park	Vietnamese Restaurant	Hotel
9	Hackney	7.331472	1	Pub	Coffee Shop	Café	Bakery	Brewery	Modern European Restaurant	Hotel
10	Haringey	6.984445	1	Park	Bus Stop	Turkish Restaurant	Mediterranean Restaurant	Café	Fast Food Restaurant	Bar
15	Islington	10.098956	1	Pub	Bakery	Turkish Restaurant	Boutique	Ice Cream Shop	Theater	Burger Joint
16	Lambeth	12.094207	1	Caribbean Restaurant	Market	Pub	Pizza Place	Cocktail Bar	Tapas Restaurant	Restaurant
17	Lewisham	9.841936	1	Supermarket	Grocery Store	Coffee Shop	Platform	Italian Restaurant	Train Station	Cocktail Bar
18	Merton	4.949223	1	Café	Supermarket	Park	Indian Restaurant	Bakery	Bus Stop	Fast Food Restaurant
19	Newham	12.354924	1	Hotel	Chinese Restaurant	Airport	Airport Lounge	Airport Service	Pharmacy	Rafting
21	Southwark	12.978814	1	Coffee Shop	Hotel	Pub	Thai Restaurant	Hotel Bar	Art Gallery	English Restaurant
24	Westminster	18.661640	1	Coffee Shop	Hotel	Sandwich Place	Theater	Sporting Goods Shop	Juice Bar	Sushi Restaurant

Conclusion and Discussion

- This report aims to enable tourists and Migrants decide which area is best to visit when coming to London.

- Analysis of crime data is provided to enable the safest choice to locate too

- The clusters also provide details venues like - supermarkets, restaurants, cinemas etc

- Segment 3 offers the most lively area in terms of venues available it also has relatively low rates of crime

- The choice of neighborhood chosen will depends on the requirements of an individual

- Further analysis can be done to look at property and rent prices