HUNT FOR A SUSTAINABLE BOROUGH TO LIVE IN LONDON

Piyush Trivedi

Table of Contents

List of Figures

Section 1: Introduction

London is a leading global city, demonstrating excellence in multiple areas including the arts, commerce, education, entertainment, research and development, fashion, finance, media, and tourism. Placed 2nd in the Global Cities Index Rank 2019 published by Kearney[1], London accounts for 23% of the UK's economic output with just 13% of the national population. It's not only the world's leading financial services centre but also Europe's fastest growing technology hub and several of the world's highest-ranking universities.

With its total population standing just over 8 million, London is one of the most popular destinations for people all over the world. As per Annual population Survey from ONS, 36% of the Londoners are born outside of the UK with 41% comprising of Black and Minority Ethnicity. With regard to its ever-growing population, London has become the most ethnically diverse region in the UK with over 300 languages spoken.

Section 2: Business Problem

London is certainly a land of opportunities for many people from all over the world and as they move to the city for job opportunities and education, they are poised with dilemma of where to find a place to Live. In this report I have tried to provide a solution for the same taking into consideration few basic criteria per borough and comparing them. The criteria considered are as follows –

[1] https://www.kearney.com/global-cities/2019

- Rent Affordability - Borough wise Rent for past 2 years
- House Affordability - House price per earnings for 2019
- Quality of Education – Average Attainment of 8 score
- Crime Safety – Crime Incident per month per 1000 people
- Public Transport Accessibility Levels
- Types of Venues available

Link to GitHub for project code.

Section 3: Data

3.1: Data Sources

The Data Sources are as follows:

- London Boroughs List from Wikipedia Page[2]
- ONS UK website –
    - London Rental Market 2020[3] and 2019[4]
    - House prices to earning data[5]
    - GSCE result data for 2019[6]
    - Borough Level Crime Data for 2019[7]
    - Population Estimate for 2019[8]
    - Public Transport Accessibility Levels data[9]
- Geocoder to get Latitude Longitude data for each Borough
- Foursquare api to get venues in a radius of 900m of each Borough
- Geo data to plot the Boroughs on the London Map[10]

---

[2] https://en.wikipedia.org/wiki/London_boroughs

[3] https://www.ons.gov.uk/peoplepopulationandcommunity/housing/adhocs/12435privaterentalmarketinlondonoctober2019toseptember2020

[4] https://www.ons.gov.uk/peoplepopulationandcommunity/housing/adhocs/11100privaterentalmarketinlondonjanuary2019todecember2019

[5] https://data.london.gov.uk/dataset/index-private-housing-rental-prices-region

[6] https://data.london.gov.uk/dataset/gcse-results-by-borough

[7] https://data.london.gov.uk/dataset/recorded_crime_summary

[8] https://data.london.gov.uk/dataset/land-area-and-population-density-ward-and-borough

[9] https://data.london.gov.uk/dataset/public-transport-accessibility-levels

[10] https://joshuaboyd1.carto.com/tables/london_boroughs_proper/public

3.2: Data Cleaning

The Inner/Outer Boroughs mapping is scrapped from Wikipedia site. The City of London is an independent entity and does not comes under any Borough classification but for the simplicity of classification I have categorized it as “Inner” Borough.

The statistics data collected from the ONS source are collected in data frames and refined further and merged into a single table with relevant field required for analysis per Borough. This data is then normalized using minmax scaling to have values between 0 to 1 for better handling of data. The Rent per month, House price per earnings data, Crime Incidents per 1000 people is further inverted and named as Rent Affordability, House Affordability, Crime Safety respectively.

Figure 3. 1: Snap of Neighbourhood statistics data

	Borough	Avg. Rent 2020	Avg. Rent 2019	House/Earnings	Avg. Attainment 8 Score	Avg. PTAL	Monthly incidents per 1000 people
0	Barking and Dagenham	1208.235294	1202.723404	10.76	46.4	5.583730	7.830753
1	Barnet	1481.329341	1442.595376	14.21	57.1	5.542330	6.541564
2	Bexley	1111.702381	1119.492958	9.68	49.6	4.552927	5.955276
3	Brent	1487.036036	1476.425197	15.74	50.2	9.223249	7.449395
4	Bromley	1317.445087	1331.671429	10.73	50.8	3.592084	6.282681

Figure 3. 2 : Snap of Normalized Neighbourhood statistics data

	Borough	Rent Affordability	House Affordability	Quality Education	Avg. PTAL	Crime Safety
0	Barking and Dagenham	0.956335	0.941304	0.791809	0.032206	0.707374
1	Barnet	0.830493	0.753804	0.974403	0.031736	0.755550
2	Bexley	1.000000	1.000000	0.846416	0.020499	0.777459
3	Brent	0.812745	0.670652	0.856655	0.073538	0.721625
4	Bromley	0.888686	0.942935	0.866894	0.009588	0.765224

The Latitude and Longitude data for each Borough is obtained by using geocoder Library. In order to get venues data, Foursquare api is used to get first 100 most popular venues in a radius of 900m of a given Borough Location. The venues category of the venues is at a granular level for example in case of restaurant we have Indian Restaurant, Chinese and so on. I have used Foursquare's "Get Venue Categories" to classify these on a broader level of 10 Categories, so for example all sorts of restaurants are now classified under Food.

Figure 3. 3: Snap of Broader Category for a given venue category

```
There are 10 uniques categories.
```

	Id	Venue Category	Main Category
0	56aa371be4b08b9a8d5734db	Amphitheater	Arts & Entertainment
1	4fceea171983d5d06c3e9823	Aquarium	Arts & Entertainment
2	4bf58dd8d48988d1e1931735	Arcade	Arts & Entertainment
3	4bf58dd8d48988d1e2931735	Art Gallery	Arts & Entertainment
4	4bf58dd8d48988d1e4931735	Bowling Alley	Arts & Entertainment

Section 4: Methodology

4.1: Exploratory Data Analysis

Once data is collected, cleaned and arranged in a meaningful way, basic charting is done to understand the statistics with respect to other Boroughs.

Figure 4. 1 : Average Rent per Borough

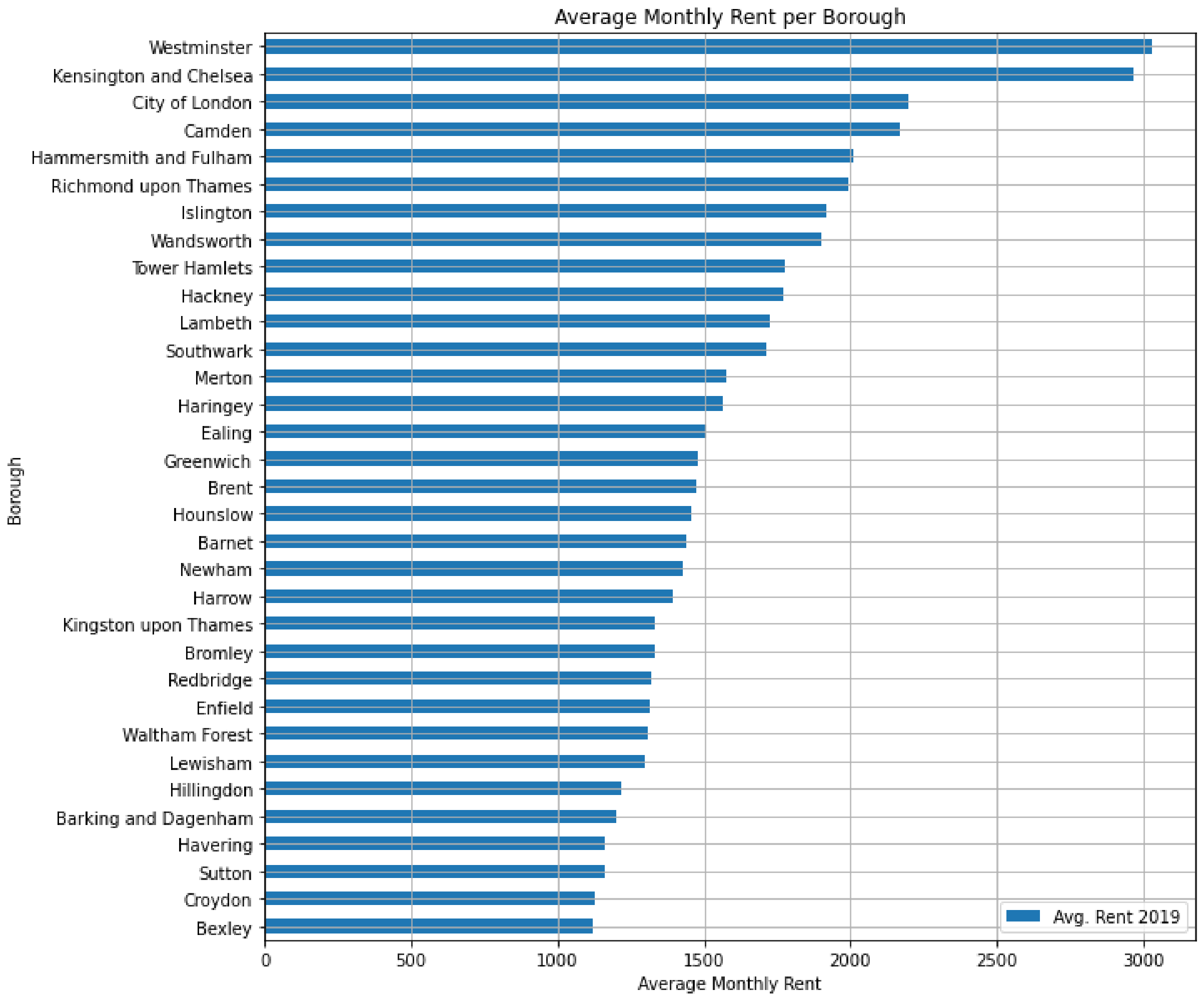


Figure 4. 2: House Price per Earnings

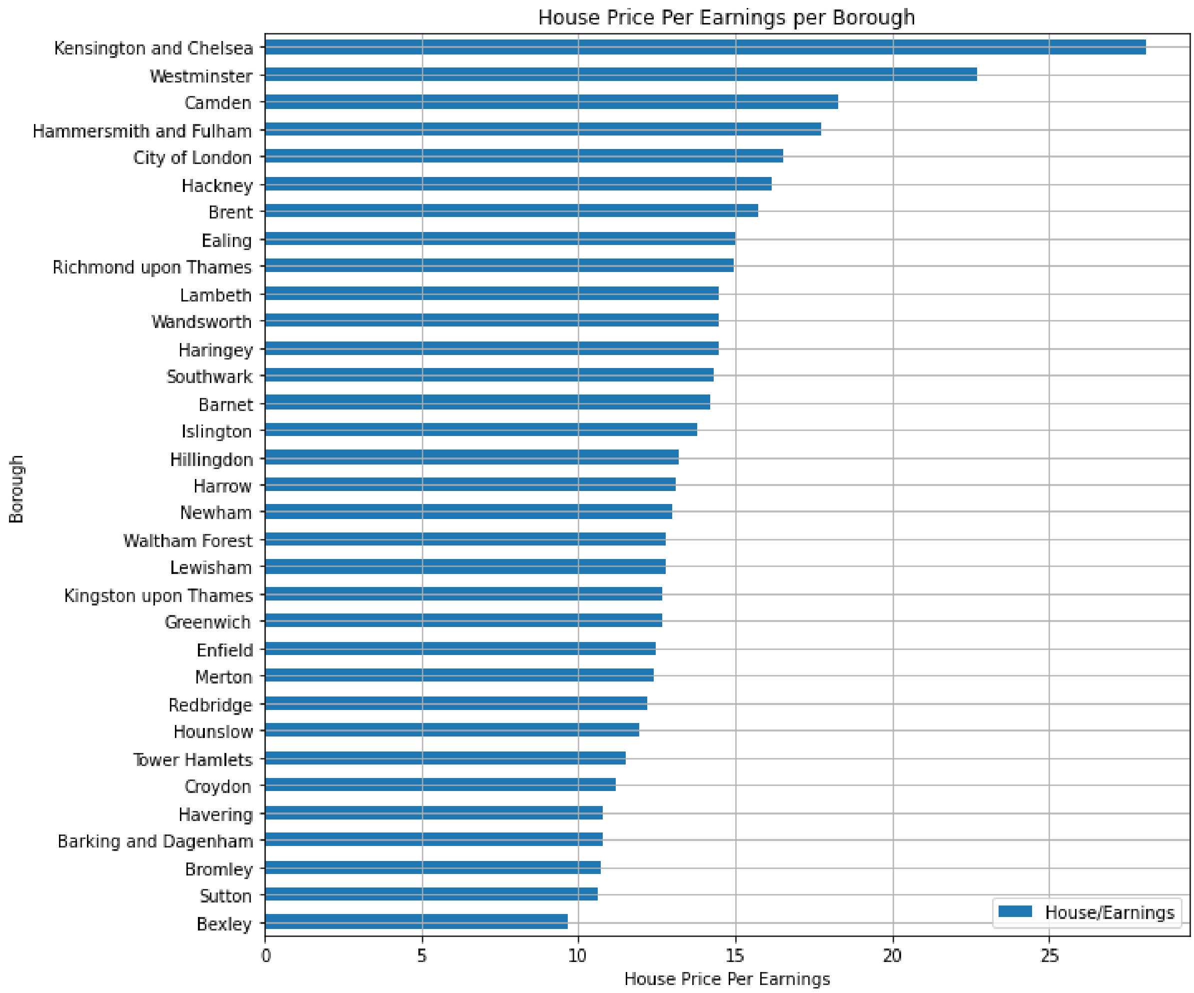


Figure 4. 3: Average Attainment 8 Score Per Student

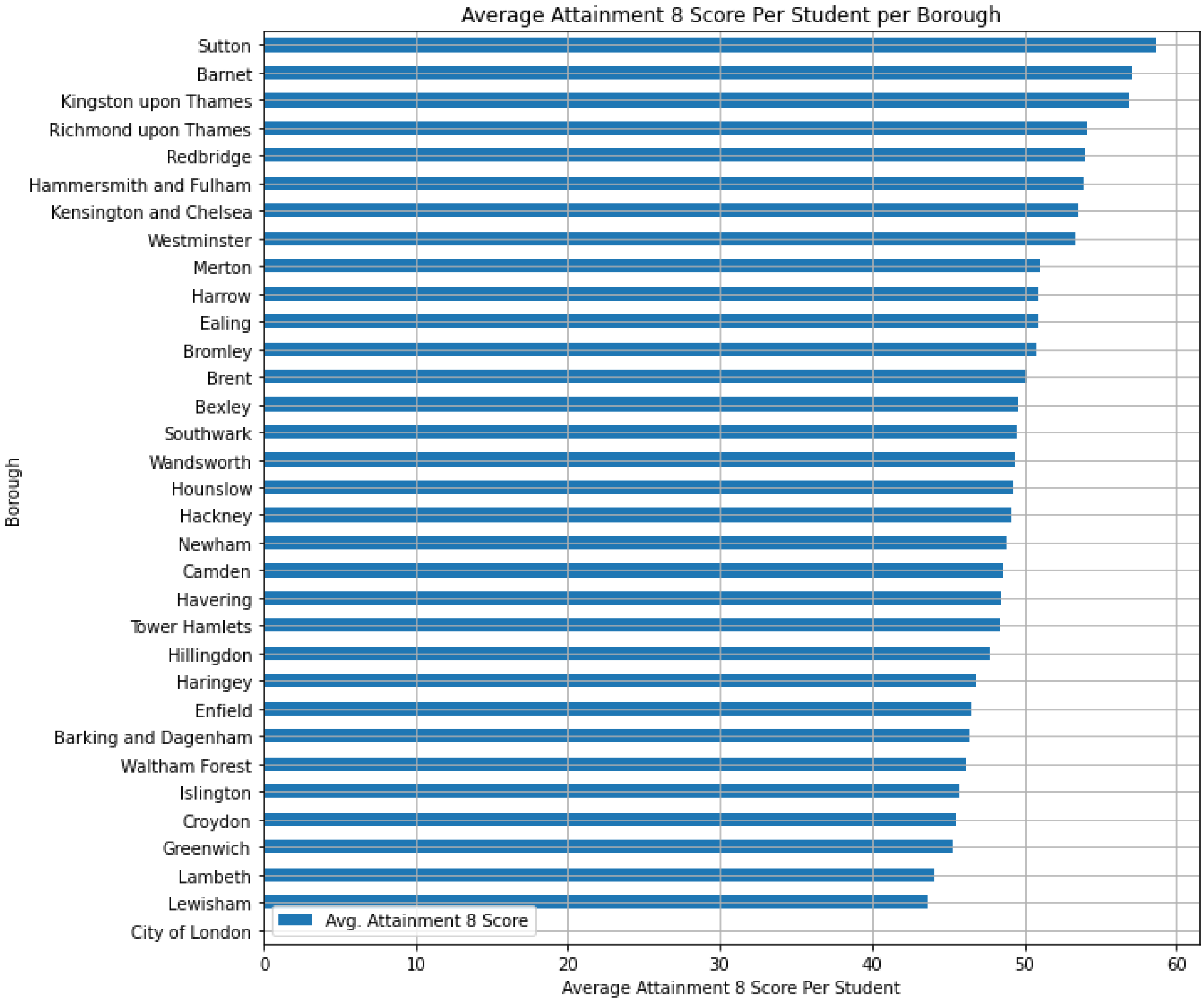


Figure 4. 4: Monthly Crime incidents per 1000 people

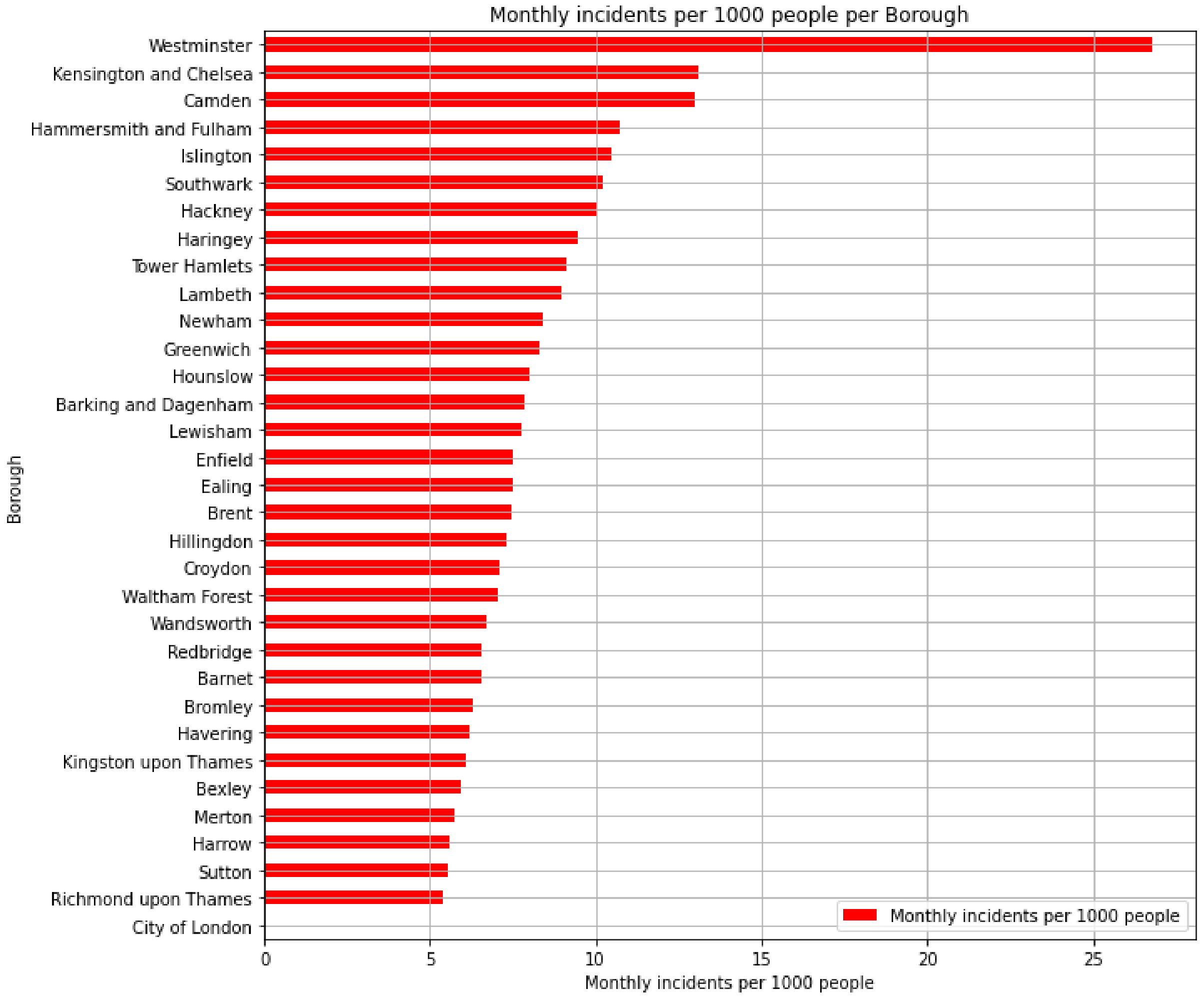


Figure 4. 5: Average Public Transport Accessibility Levels

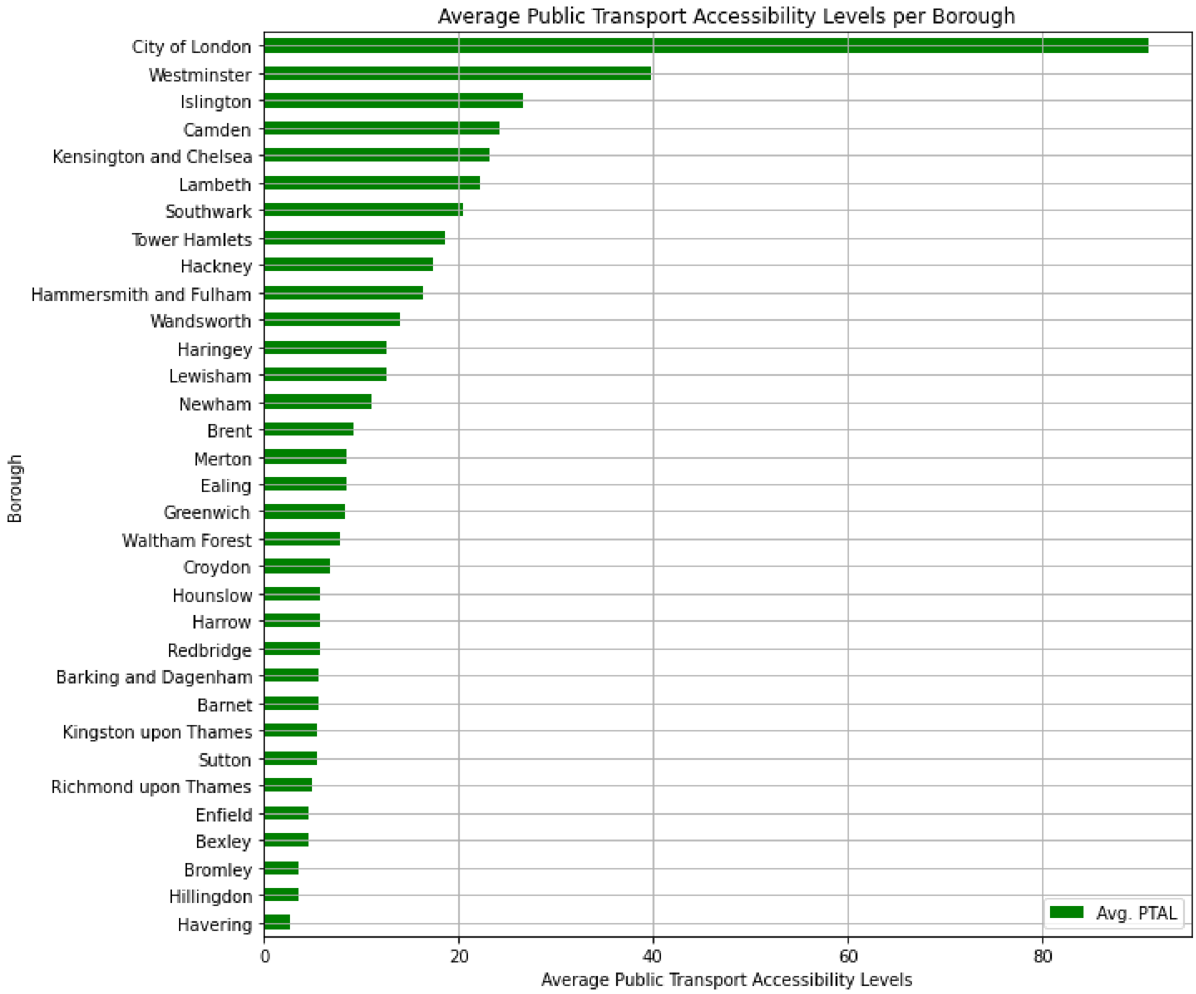
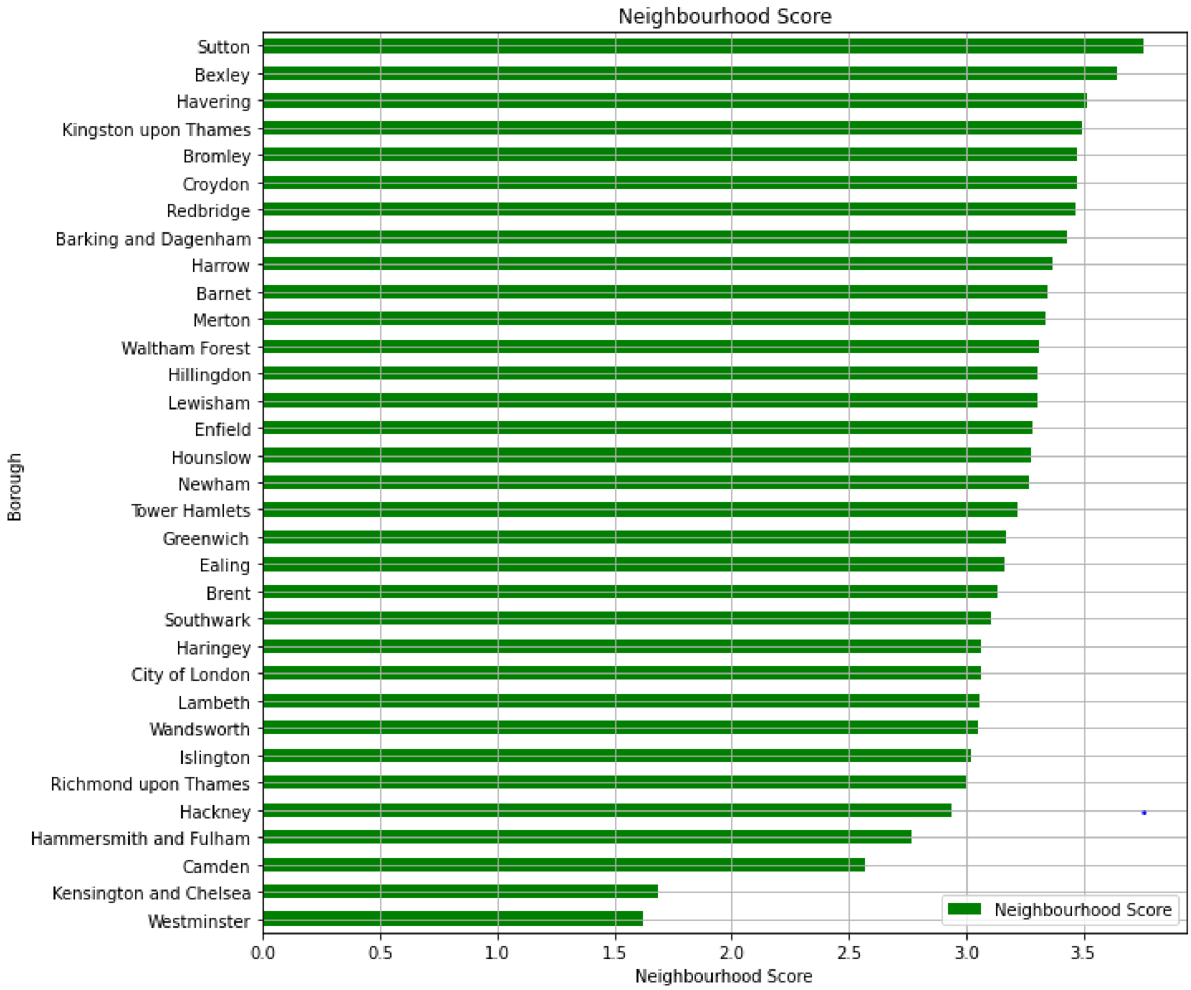


Figure 4. 6: Neighbourhood Score



4.1.a: Pointers from the Exploratory Analysis

From the charts above we can see that – the rent prices and the house price per earnings is Higher for most of the Inner Boroughs where as its lower for the Outer Borough. With respect to Education, its better in most of the areas but Sutton, Barnet and Kingston Upon Thames stands out as the best. Crime incidents reported in Westminster is very high as compared to other Boroughs with Sutton being the Lowest. In Regards to City of London, the Crime incidents as well as Education stats are negligible as the population is very low (only 7700 in 2018), however; the Average Public Transport Accessibility Level (PTAL) is highest which is obvious as most of the people travel to the City from other Boroughs for Work. PTAL decreases as we move away from the central areas of London and into the Outer Boroughs.

The Last chart – Neighbourhood Score is derived by summing the Normalized statistics data - Rent Affordability, House Affordability, Crime Safety, Quality Education and Average PTAL. It gives a Borough wise picture of the how each Borough stands based on the Neighbourhood factors we have considered as important in decision making for choosing a place to Live. Westminster, Kensington and Chelsea and Camden are at the bottom positions which can be attributed to higher rate of Rent and Crime. The Outer Boroughs occupy the top spot as they are cheaper to Live as compared to Inner counter parts with better Education score and Lower Crime incidents.

4.2: Cluster Analysis

Cluster the Boroughs of London based on the Normalized values of Neighbourhood Statistics which were described earlier along broader category of the venues data collected from Foursquare api. This exercise would help in clustering the Boroughs in similar groups based on the factors which we have considered important as a part of selecting a sustainable place to Live. Once the clustering is done, we can analyse characteristics inter as well as intra cluster to come up with a conclusion.

The Categorical venue data is converted to numeric using OneHot coding and is processed further to get the average distribution of the venues per Borough.

Figure 4. 7: OneHot Coding snap of Boroughs

	Borough	Arts & Entertainment	College & University	Food	Nightlife Spot	Outdoors & Recreation	Professional & Other Places	Residence	Shop & Service	Travel & Transport
0	Barking and Dagenham	0	0	0	0	0	0	0	1	0
1	Brent	0	0	0	0	0	0	0	1	0
2	Bromley	0	0	0	0	0	0	0	1	0
3	Camden	0	0	0	0	0	0	0	1	0
4	City of London	0	0	0	0	0	0	0	1	0

Taking the mean of the data obtained from oneHot coding we get the most visited venues in each Borough.

Figure 4. 8: Snap of most visits Venues in a Borough

	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
0	Barking and Dagenham	Shop & Service	Travel & Transport	Outdoors & Recreation	Residence	Professional & Other Places	Nightlife Spot	Food
1	Barnet	Food	Shop & Service	Nightlife Spot	Outdoors & Recreation	Professional & Other Places	Arts & Entertainment	Travel & Transport
2	Bexley	Outdoors & Recreation	Food	Shop & Service	Nightlife Spot	Travel & Transport	Residence	Professional & Other Places
3	Brent	Food	Shop & Service	Outdoors & Recreation	Nightlife Spot	Travel & Transport	Residence	Professional & Other Places
4	Bromley	Nightlife Spot	Shop & Service	Travel & Transport	Food	Outdoors & Recreation	Arts & Entertainment	Residence

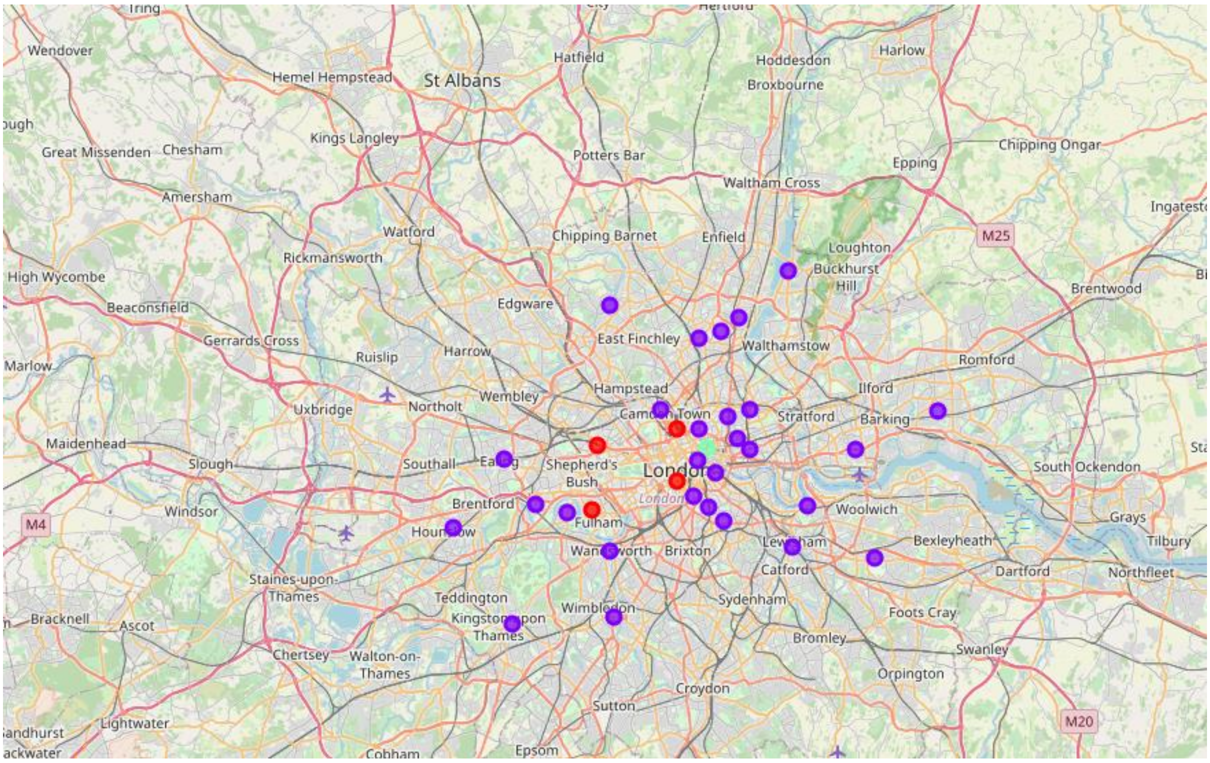
K-means clustering is used to cluster the data. The optimal Level of cluster is decided based on Silhouette Score Method which gave the result as 3. Based on the factors considered, London Boroughs can be divided in 3 clusters.

Figure 4. 9: Silhouette Score graph

Figure 4. 10: Snap of Borough Clustering with most common venues

	Borough	Avg. Rent 2020	Avg. Rent 2019	House/Earnings	Avg. Attainment 8 Score	Avg. PTAL	Monthly incidents per 1000 people	Latitude	Longitude	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
0	Barking and Dagenham	1208.235294	1202.723404	10.76	46.4	5.583730	7.830753	51.543932	0.133157	1	Shop & Service	Travel & Transport	Outdoors & Recreation	Residence	Professional & Other Places	Nightlife Spot	Food
1	Barnet	1481.329341	1442.595376	14.21	57.1	5.542330	6.541564	51.527095	-0.066826	1	Food	Shop & Service	Nightlife Spot	Outdoors & Recreation	Professional & Other Places	Arts & Entertainment	Travel & Transport
2	Bexley	1111.702381	1119.492958	9.68	49.6	4.552927	5.955276	51.452078	0.069931	1	Outdoors & Recreation	Food	Shop & Service	Nightlife Spot	Travel & Transport	Residence	Professional & Other Places
3	Brent	1487.036036	1476.425197	15.74	50.2	9.223249	7.449395	51.609783	-0.194672	1	Food	Shop & Service	Outdoors & Recreation	Nightlife Spot	Travel & Transport	Residence	Professional & Other Places
4	Bromley	1317.445087	1331.671429	10.73	50.8	3.592084	6.282681	51.601511	-0.066365	1	Nightlife Spot	Shop & Service	Travel & Transport	Food	Outdoors & Recreation	Arts & Entertainment	Residence
5	Camden	2027.228916	2171.184615	18.28	48.6	24.316782	12.983349	51.532360	-0.127960	0	Food	Shop & Service	Arts & Entertainment	Outdoors & Recreation	Travel & Transport	Nightlife Spot	Professional & Other Places
6	City of London	2071.000000	2198.600000	16.55	0.0	90.802508	0.000000	51.520500	-0.097430	2	Food	Outdoors & Recreation	Nightlife Spot	Arts & Entertainment	Shop & Service	Travel & Transport	Professional & Other Places
7	Croydon	1149.169399	1126.730924	11.18	45.5	6.757744	7.116415	51.593480	-0.083420	1	Food	Shop & Service	Outdoors & Recreation	Arts & Entertainment	Travel & Transport	Residence	Professional & Other Places

Figure 4. 11: Visualization of the Clusters



These are further grouped by Borough per cluster to Analyse the behaviour of venues and Neighbourhood factors. Bar charts with respect to Neighbourhood factors and venues factors of each clusters

Figure 4. 12: Stats per each Clusters

Cluster Labels	Rent Affordability	House Affordability	Quality Education	Avg. PTAL	Crime Safety	Arts & Entertainment	College & University	Food	Nightlife Spot	Outdoors & Recreation	Professional & Other Places	Residence	Shop & Service	Travel & Transport	Count
0	0.253160	0.346739	0.893771	0.263089	0.406290	0.082331	0.000000	0.484601	0.097304	0.138977	0.022500	0.000000	0.115687	0.058600	4
1	0.818877	0.823292	0.844344	0.077602	0.721073	0.043724	0.000714	0.402287	0.139022	0.142169	0.004662	0.000357	0.203873	0.063193	28
2	0.433875	0.626630	0.000000	1.000000	1.000000	0.110000	0.000000	0.440000	0.150000	0.170000	0.020000	0.000000	0.070000	0.040000	1

Figure 4. 13: Venues per Cluster

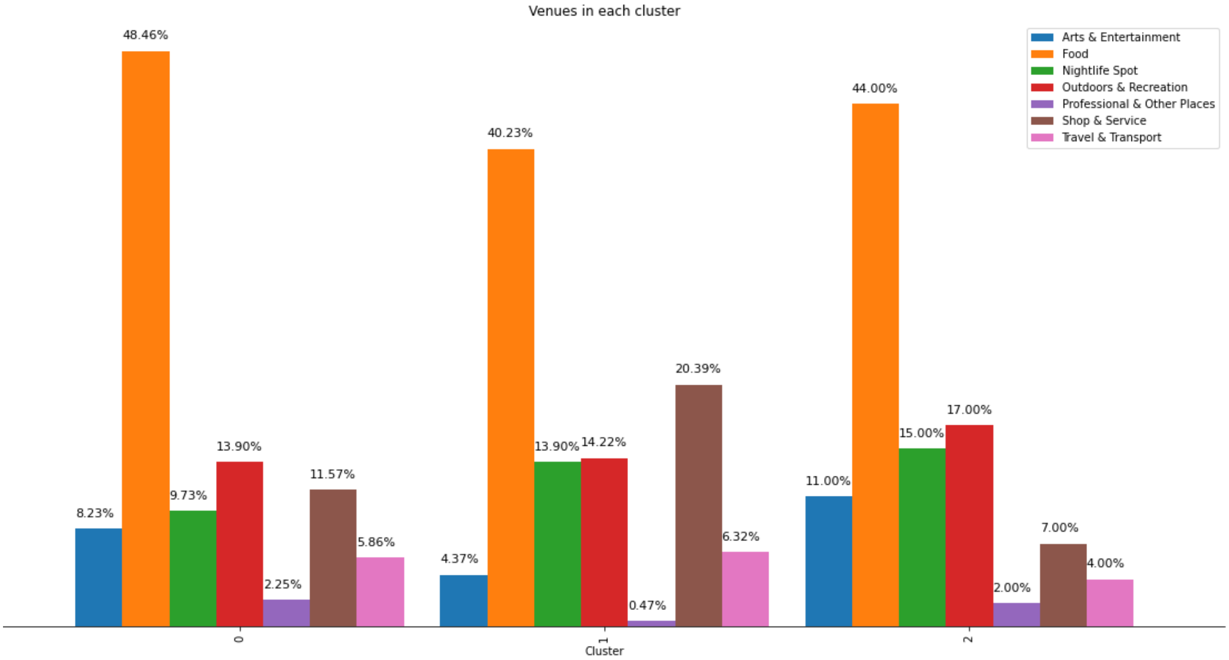
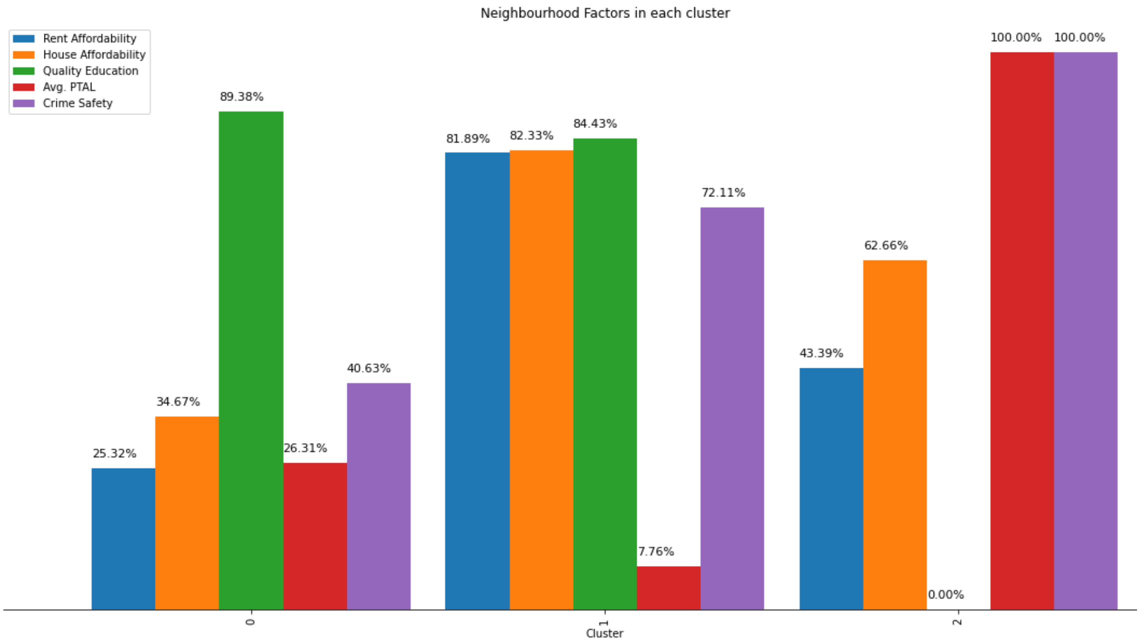


Figure 4. 14: Neighbourhood per Cluster



4.2.a: Pointers from the Cluster Analysis

From the Bar charts following observations can be made –

Figure 4. 15: Cluster Takeaways

Cluster	**Venues**	**Neighbourhood**
2: City of London	Mostly has places for Food, Night Life and Art	Average PTAL and Crime Safety is highest which is expected as it's has lower population and located at heart of the London with commuters coming to work from other Boroughs
1: Comprises of all the Outer and few Inner Boroughs	Has Higher concentration of Food places but lower than other clusters, followed by Shop and services	Best Rent and House Affordability as compared to other 2 clusters. PTAL is Lower than other clusters with better Education quality and Crime Safety
0: Westminster, Kensington and Chelsea, Camden and Hammersmith & Fulham	Better spots for Food. Other venues are not	Renting or purchasing property is costlier but Education quality is a bit better compared to Cluster 1.

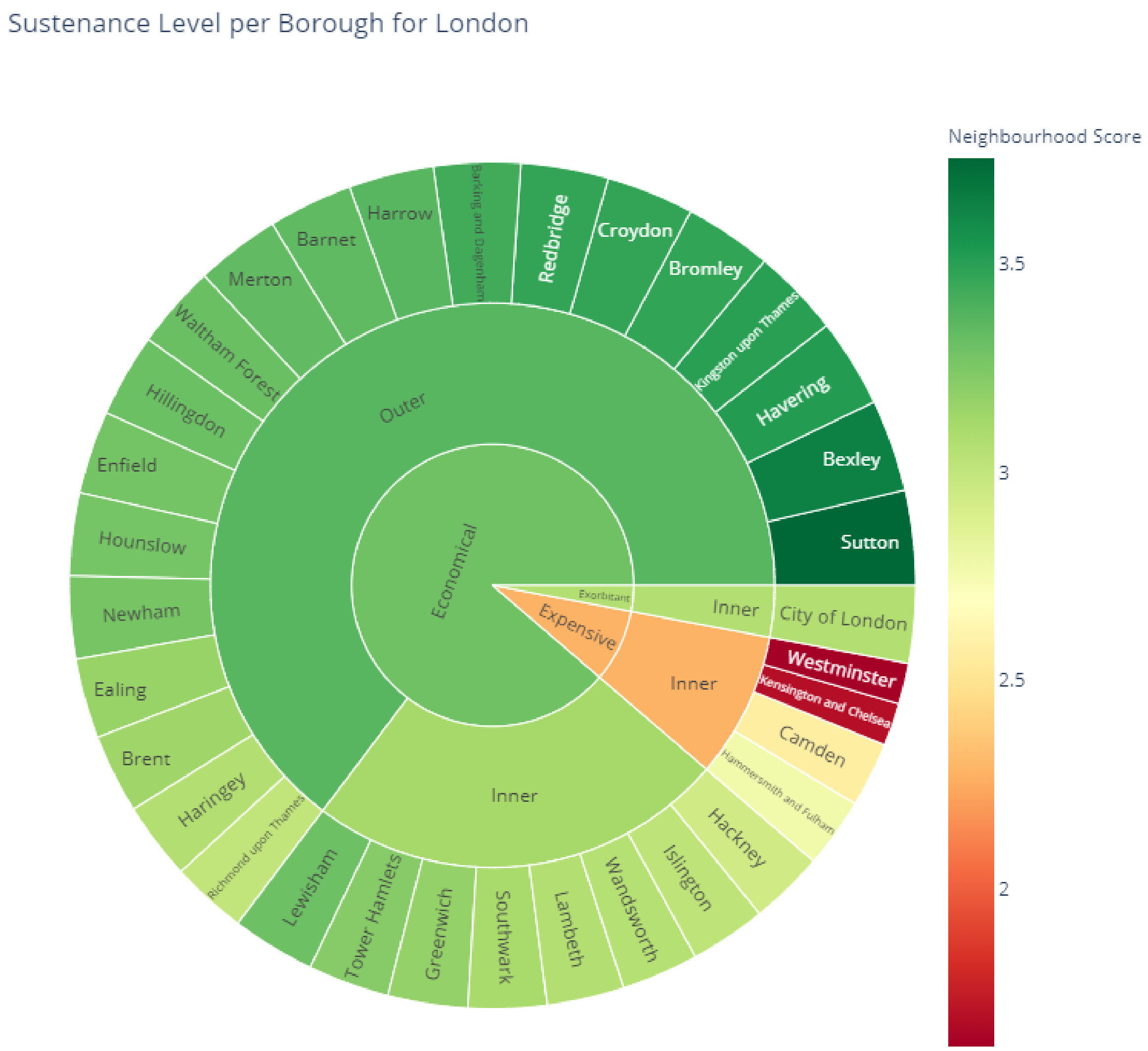
Based on the above observation clusters are labelled by Sustenance Level:

- 0: Exorbitant
- 2: Expensive
- 1: Economical

Section 5: Results

Post classifying the clusters as Sustenance Level based on their characteristics, I have tried to create visualization using sunburst graph to give an overall picture of how Boroughs are divided with respect to clusters, Borough designation – Inner or Outer and Neighbourhood score.

Figure 5. 1: Borough Sunburst



From the sunburst chart we can see that in Economical Boroughs we can further use Neighbourhood score to rank the Borough. All of the Outer Boroughs are under Economical Clusters. Sutton, Bexley, Havering, Kingston Upon Thames and Bromley are Top 5 Boroughs to Live.

Figure 5. 2: Choropleth of Clusters against Neighbourhood score

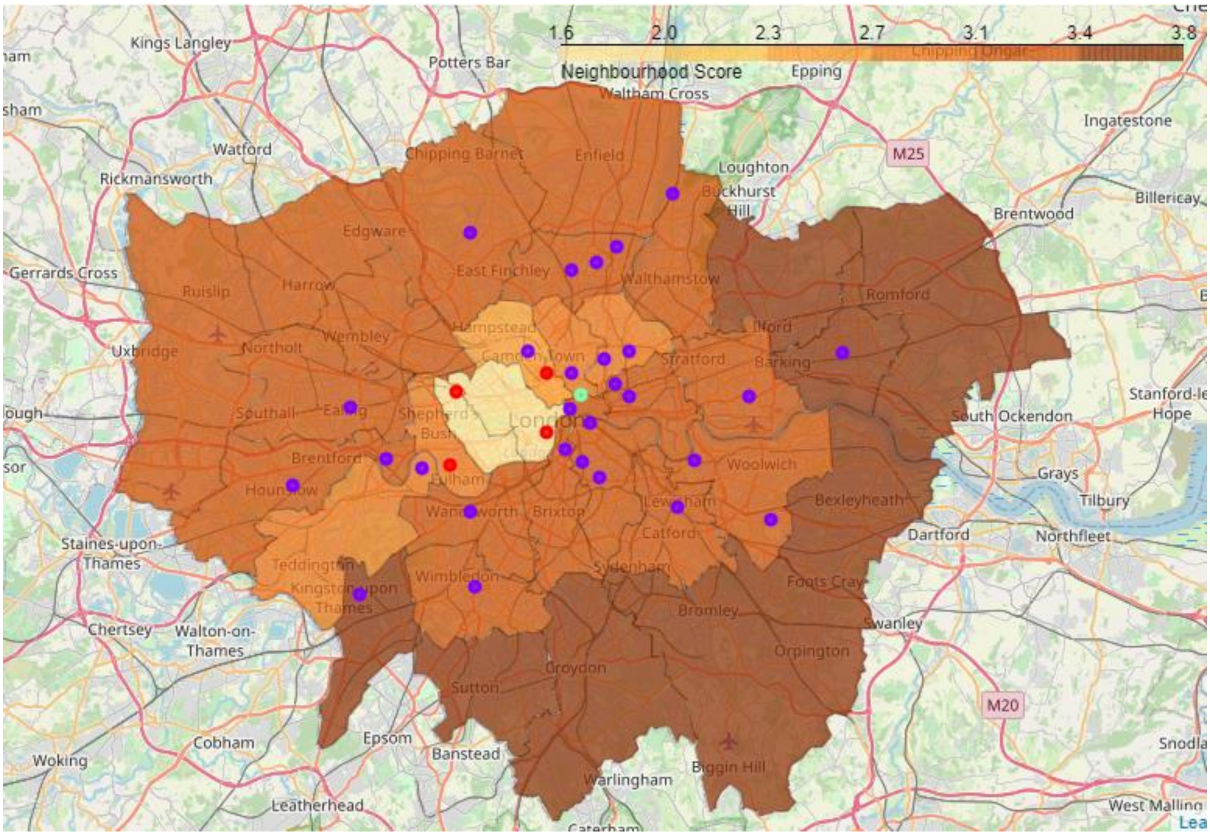
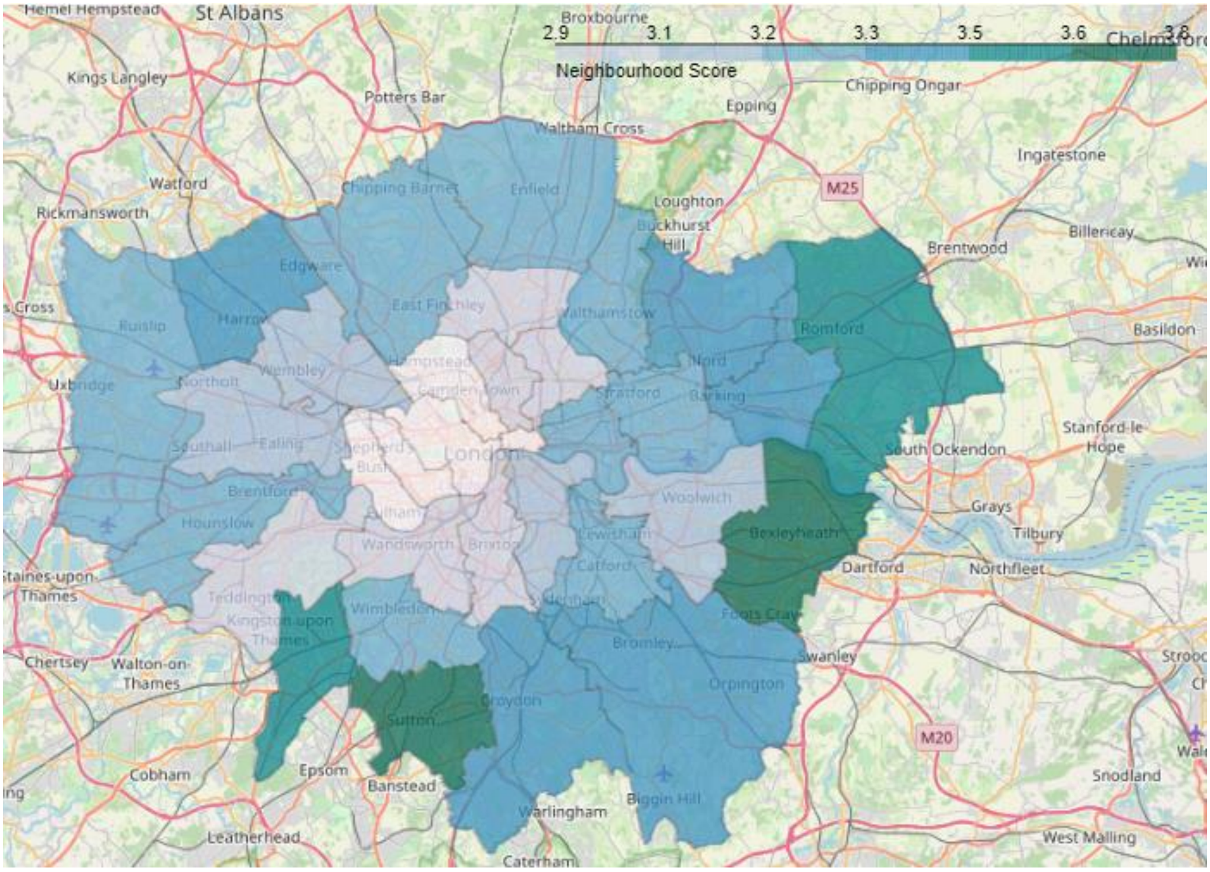


Figure 5. 3: Boroughs in Economical Clusters with Neighbourhood Score

Borough	Designation	Sustenance Level	Neighbourhood Score
Sutton	Outer	Economical	3.752440
Bexley	Outer	Economical	3.644374
Havering	Outer	Economical	3.515389
Kingston upon Thames	Outer	Economical	3.497726
Bromley	Outer	Economical	3.473327
Croydon	Outer	Economical	3.470738
Redbridge	Outer	Economical	3.469958
Barking and Dagenham	Outer	Economical	3.429029
Harrow	Outer	Economical	3.364643
Barnet	Outer	Economical	3.345985
Merton	Outer	Economical	3.337109
Waltham Forest	Outer	Economical	3.313820
Hillingdon	Outer	Economical	3.307263
Lewisham	Inner	Economical	3.305107
Enfield	Outer	Economical	3.281500
Hounslow	Outer	Economical	3.276462
Newham	Outer	Economical	3.270877
Tower Hamlets	Inner	Economical	3.220094
Greenwich	Inner	Economical	3.173046
Ealing	Outer	Economical	3.159853
Brent	Outer	Economical	3.135216
Southwark	Inner	Economical	3.103356
Haringey	Outer	Economical	3.065031
Lambeth	Inner	Economical	3.059469
Wandsworth	Inner	Economical	3.050591
Islington	Inner	Economical	3.019902
Richmond upon Thames	Outer	Economical	3.003035
Hackney	Inner	Economical	2.939945

Figure 5. 4: Neighbourhood Score plot for Economical Borough



Section 6: Discussion

As Observed in the result section, Economical Clusters comprises of all Outer Boroughs including some Inner Boroughs also. The Neighbourhood score gets better as we move away from the Inner Boroughs into the Outer Boroughs which can be seen from Map Plot of Figure 5.4. Sutton, Bexley, Havering, Kingston Upon Thames and Bromley which are Top 5 Boroughs to Live belong to Outer Boroughs. If someone prefers to stay in the Inner Boroughs, we have Lewisham and Tower Hamlets as Top 2 options though they would be a bit costlier as compared to the Top choices.

The analysis is based on the 2019 data only, so accuracy of the analysis might not be spot on. In order to get better accuracy, it would be preferable to consider the time series analysis of the Neighbourhood factors. Further, only 6 factors (namely average rent 2019, house price per earnings, GCSE Average Attainment of 8 score, crime count 2019, Public Transport Accessibility Levels and venues data) were considered per Borough for clustering. In order to increase accuracy and also the level of clustering (Economical cluster has 28 of the 33 Broughs almost 85% of the Total), other important factors like air quality, social integration and access to nature can also be considered. Since most of the population is working from Home due to COVID situation "Internet Broadband Speed" per Borough could also be considered as an important factor.

Section 7: Conclusion

This study dealt with finding out a sustainable place to live in London. Based on the Factors considered and clustering of Boroughs, we can conclude that Outer Boroughs are the better to live as compared to the Inner Boroughs. The results in the study are based on the limited factors as well as limited dataset which have provided a Higher-level result. For better accuracy and get more granular results, its recommended to consider more factors and a detailed data set over a period of time.