

Identifying the best ward to purchase a property in within Greater London

By Charlotte Fettes

Identifying suitable locations within London to buy a property in is beneficial for house buyers

- ▶ London is a highly diverse city with many unique areas within it. Most people do not know much about many areas within London to make an informed decision on which area would best meet their needs and preferences
- ▶ Property buyers base their decisions on where to buy on more than just price. Other important factors they consider include crime rate, property type, educational quality (for those with children), access to public transport, open space, and the character of the area in general.
- ▶ The diversity of London and the numerous factors important in choosing a location makes researching for suitable locations within which to buy a property is a very time- and effort-intensive activity.
- ▶ Estate agents have an interest in providing their clients with a personalized service to improve their chance of making a sale and of guaranteeing return business.

The Case Study

To demonstrate the use of this project a case study was used. Our case study is a family of 5 - husband and wife, and their 3 children aged 12, 9, and 5 - moving into London due to the wife accepting a job position in central London, the husband wanting to build a personal training business and access more clients, and wanting to provide their children with more to do. Ideally, a location would fulfil the following requirements:

- ▶ The property price to be around their budget of £800,000.
- ▶ The husband is a personal trainer looking to build up and expand his business, and so access to gyms and outdoor areas is important.
- ▶ The wife will be working in central London, so good access to the public transport network would be beneficial.
- ▶ With one child now in secondary school, and GCSE's on the horizon, the quality of secondary education is particularly important. With another two children in primary school, the quality of primary education is also important.
- ▶ They are looking for a house - ideally semi-detached or detached, with a terraced house being an option in the right area. They do not want a flat.
- ▶ With children, safety is always a concern. They want to be sure that they are in a safe area with low crime levels.
- ▶ An area with parks and open space would be great, but not so much that facilities and amenities of London are just as inaccessible as in their current location. One of the reasons they want to be in London so there is more to do around them at a closer distance.
- ▶ A sociable area with cafes, food stores, and things for the children to do.

Data Acquisition and Cleaning

Data on London wards was obtained from London DATASTORE. This included:

- ▶ Ward profiles and atlas, from which the features selected were crime rate, average GCSE point score, % area open space, public transport accessibility score, % detached houses, % semi-detached houses, % terraced house, % flat, maisonette or apartment
- ▶ Key stage 1 and key stage 2 2018 data for a measure of primary education
- ▶ Land registry average house prices for the year ending December 2017

Geocoder was used with the provider Location IQ to find latitude and longitude coordinates for all wards

Foursquare was used to explore ward locations and return popular venues for each ward.

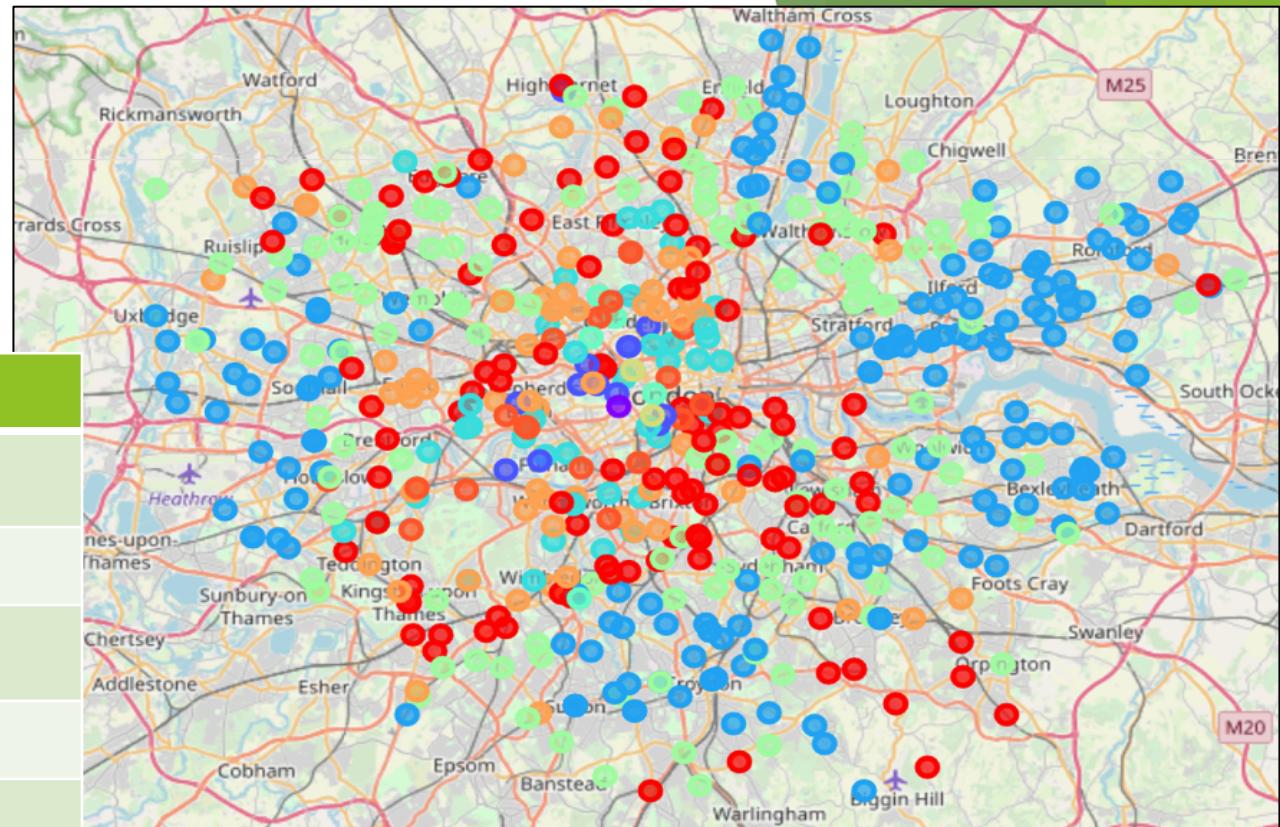
The resulting dataframe contained a total of 571 wards from which to choose from.

Exploratory Analysis

- ▶ K-Means clustering was used for all features in order to group wards in a way that limited intra-cluster variability and maximized inter-cluster variability.
- ▶ Elbow curves were devised for features to find the optimal number of clusters for each feature.
- ▶ Descriptive statistics were generated for the entire datasets for each feature. These were used to understand the datasets better and determine, based on means, what could be considered an average. For example, mean crime rate informed which clusters had an average, above average and below average crime rate.
- ▶ Comparative plots were used to visualize the clusters in comparison to each other.
- ▶ Scores were assigned to each cluster based on the case study preferences

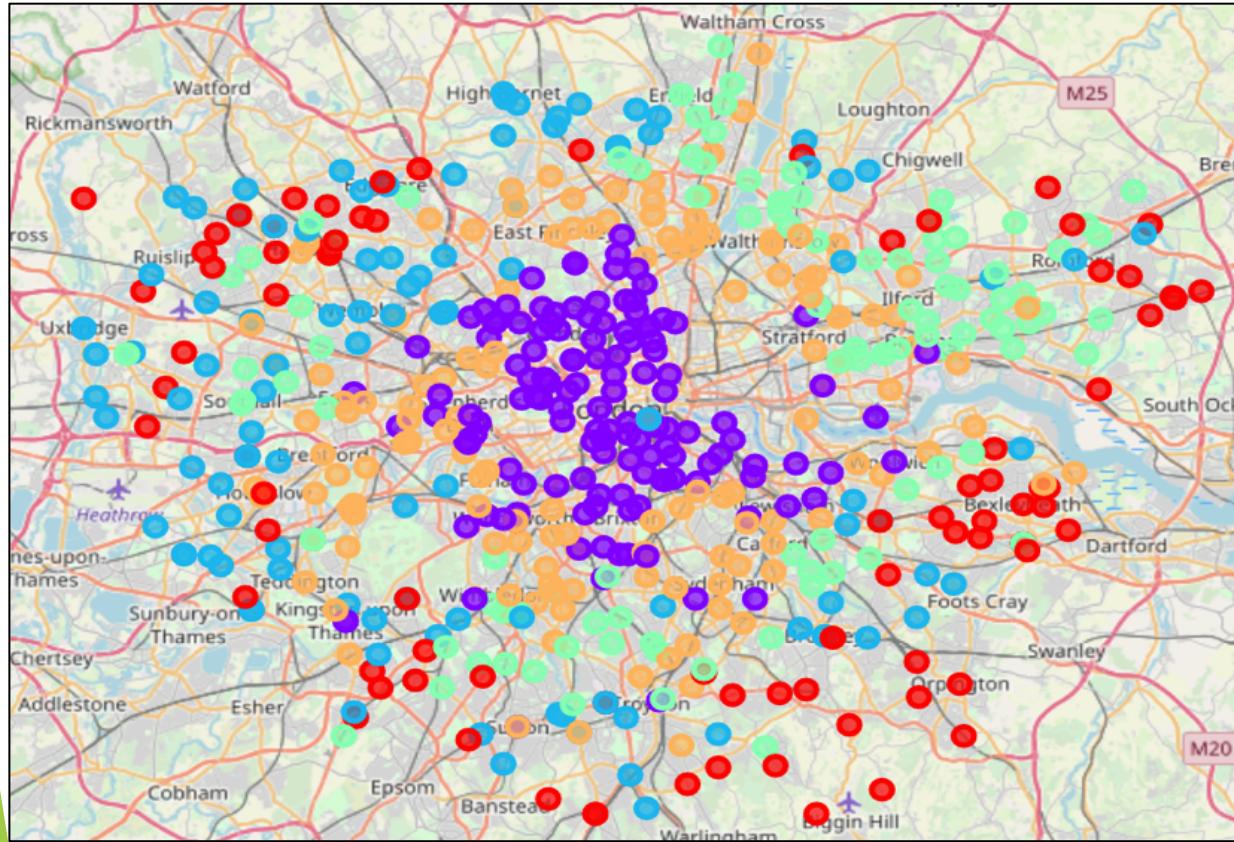
Property Price Clusters

Cluster	Average price range	Score
Cluster0 (red)	£524,100.61 - £648,806.41	5
Cluster1 (purple)	£4,416,659.40	-1000
Cluster2 (dark blue)	£1,296,800.86 - £1,585,562.76	-1000
Cluster3 (light blue)	£259,486.62 - £401,320.31	1
Cluster4 (turquoise)	£823,462.88 - £1,008,300.62	2
Cluster5 (green):	£1,891,715.94 - £2,186,666.67	-1000
Cluster6 (lime green):	£403,917.61 - £518,567.23	4
Cluster7 (yellow):	£2,856,435.71 - £3,033,756.54	-1000
Cluster8 (light orange):	£649,877.67 - £817,646.99	5
Cluster9 (dark orange):	£1,021,486.54 - £1,238,295.89	-1000

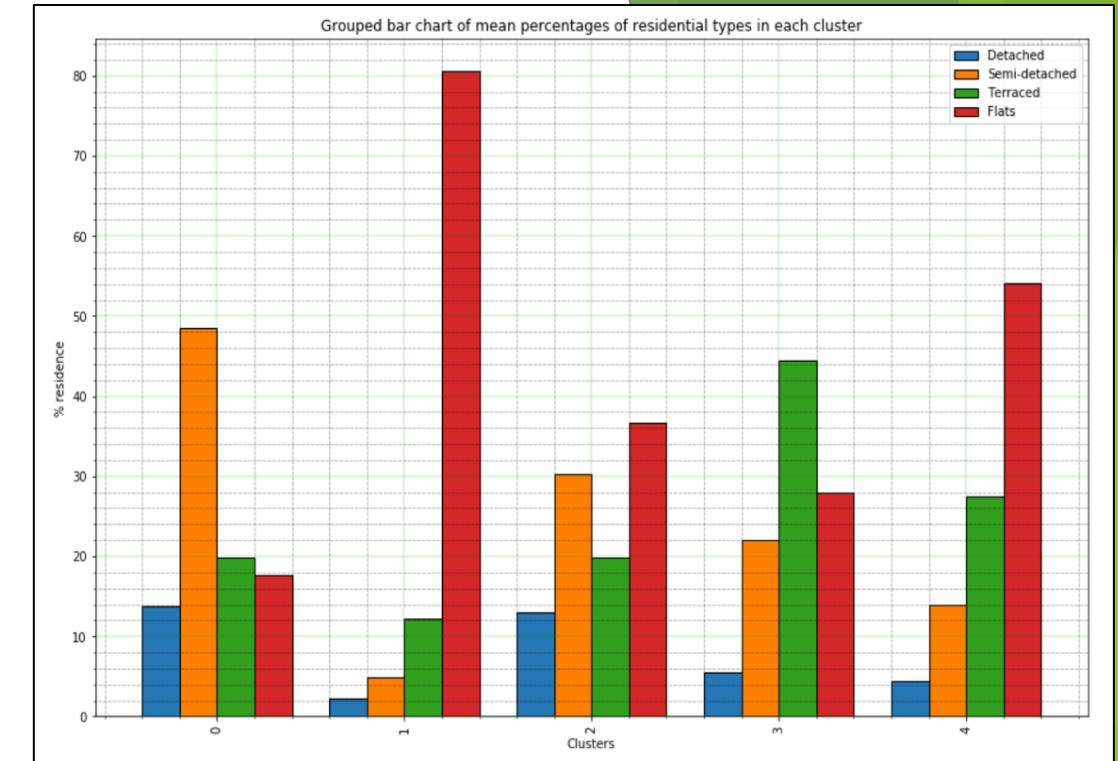


The buyer has a budget of up to £800,000. With clusters based on the mean property price within wards, with expected variability within wards, and with the potential for price negotiation, all clusters containing wards with a mean property price of up to £1m were considered. All clusters containing wards with a mean property price above £1m were excluded by assigning a large negative value (-1000). Clusters containing average property prices slightly below the mean were awarded a high score as property prices will vary around the means, and so the buyer will be able to afford a property at the upper end of the market in these locations.

Property Type



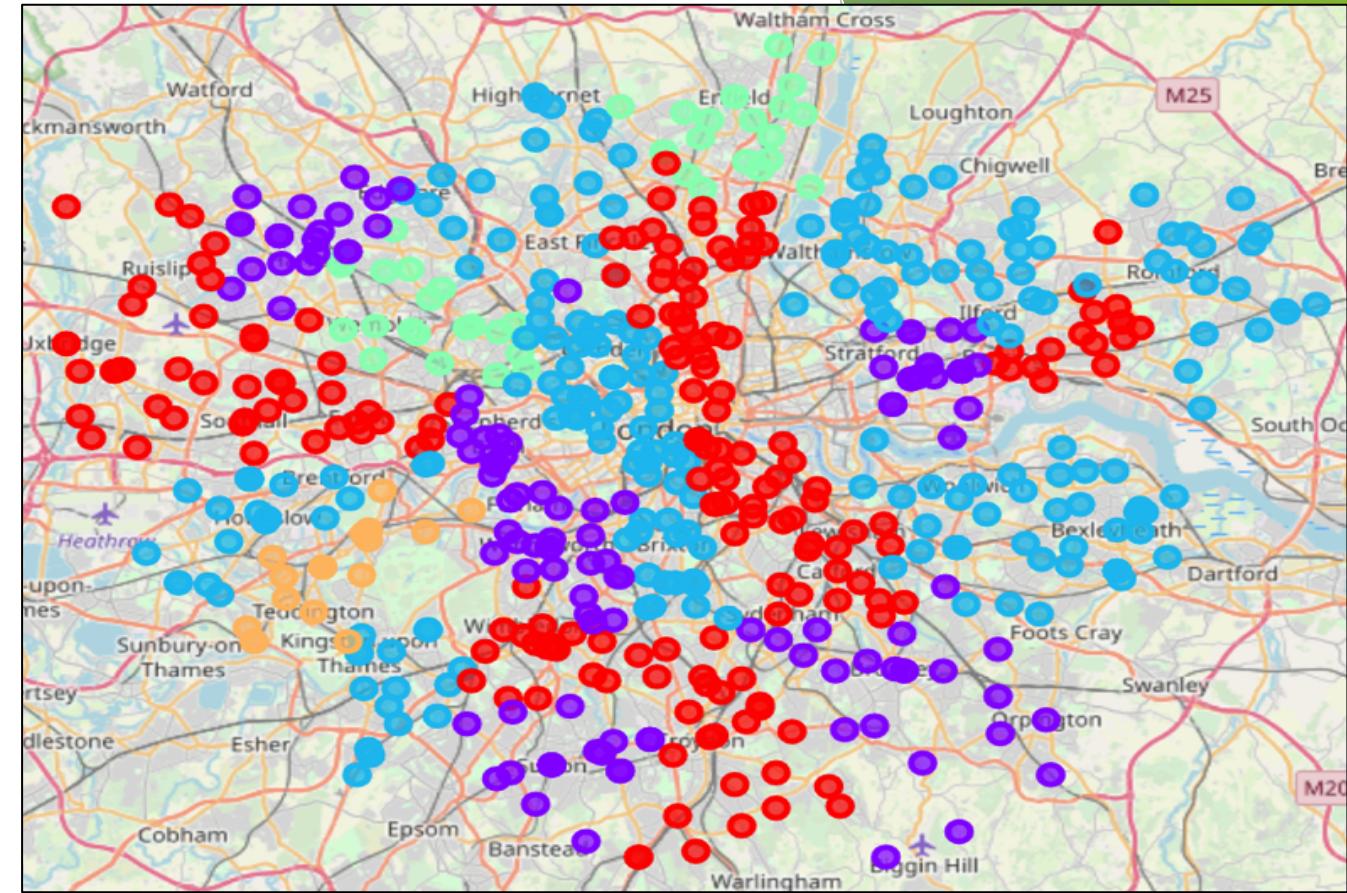
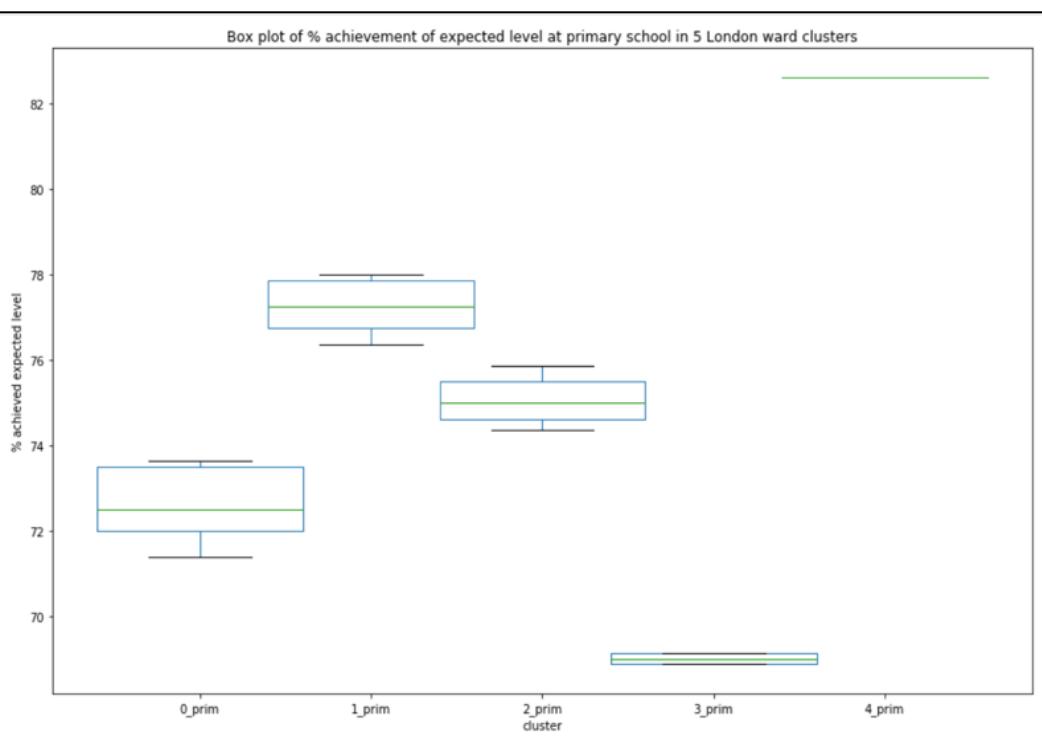
The house buyer expressed a preference for a house over a flat, apartment or maisonette. They are indifferent between detached and semi-detached as their first choice, with terraced houses being the second choice.



Cluster (colour for map)	Property types	Scores
0 (red)	detached (1), semi-detached (1), terraced (4), flats(5)	5
1 (purple)	detached (5), semi-detached (5), terraced (5), flats(1)	0
2 (blue)	detached (2), semi-detached (2), terraced (3), flats(3)	4
3 (green)	detached (3), semi-detached (3), terraced (1), flats(4)	2
4 (orange)	detached (4), semi-detached (4), terraced (2), flats(2)	1

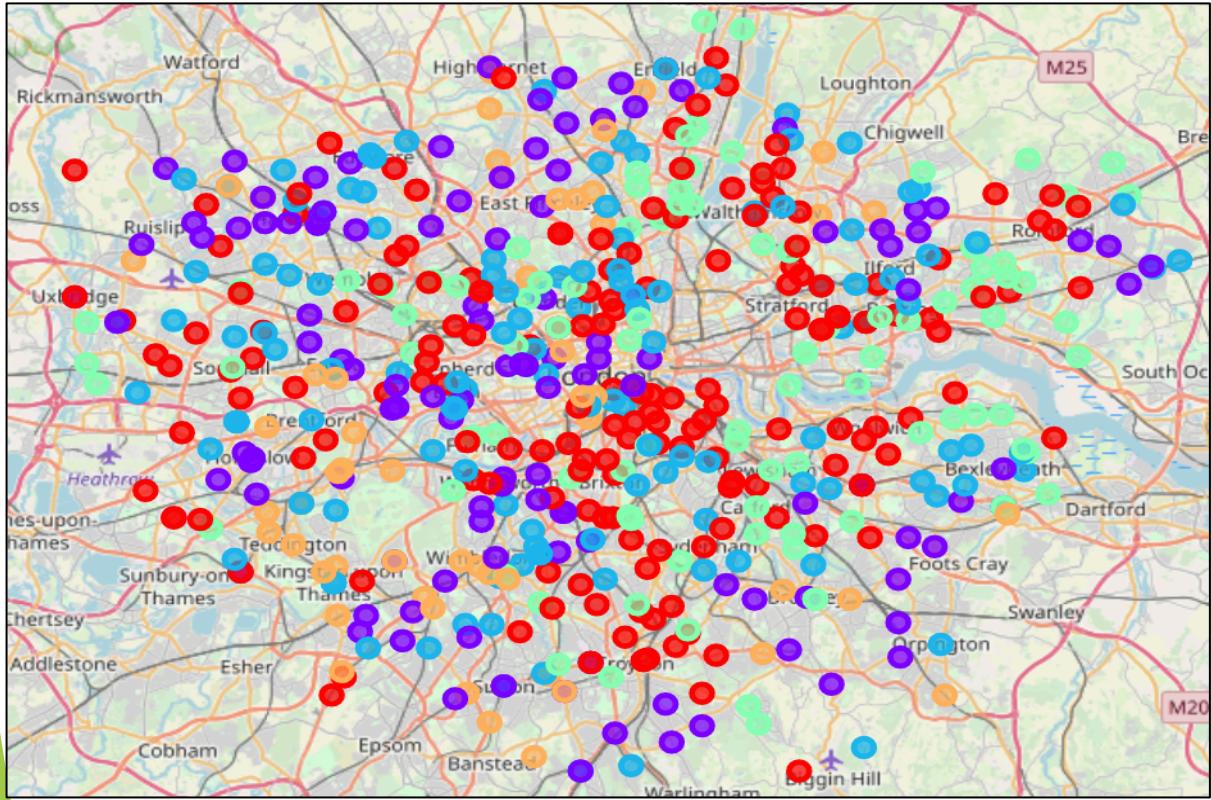
Primary Education

With two children in primary school, the quality of primary education is important to the property buyer. The higher the percentage of children achieving the expected level the better. The mean percentage (75%) and above is taken as indication of the primary education being of an acceptable standard; any clusters with wards below the mean level are assigned a value of 0.

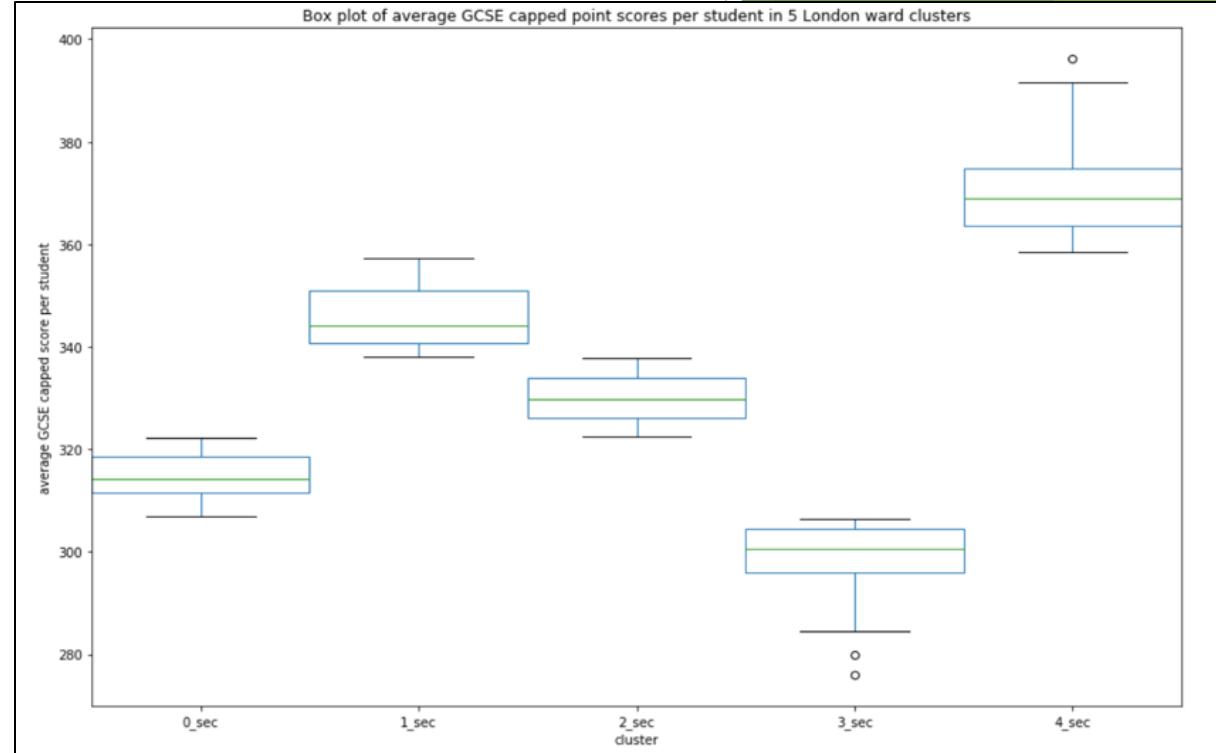


Cluster (colour for map)	Range of % of children achieving expected primary level	Score
0 (red)	71.38-73.62% - below average	0
1 (purple)	76.38-78% - above average	4
2 (blue)	74.38-75.88% - low average	0
3 (green)	68.88-69.12% - lowest level	0
4 (orange)	82.62% - highest level	5

Secondary Education



With one child in secondary school, the quality of secondary education is important to the case study property buyers. The higher the average GCSE point score of a ward the better. The mean point score (327) and above is taken as indication of the secondary education being of an acceptable standard; any clusters with wards below the mean level are assigned a value of 0.

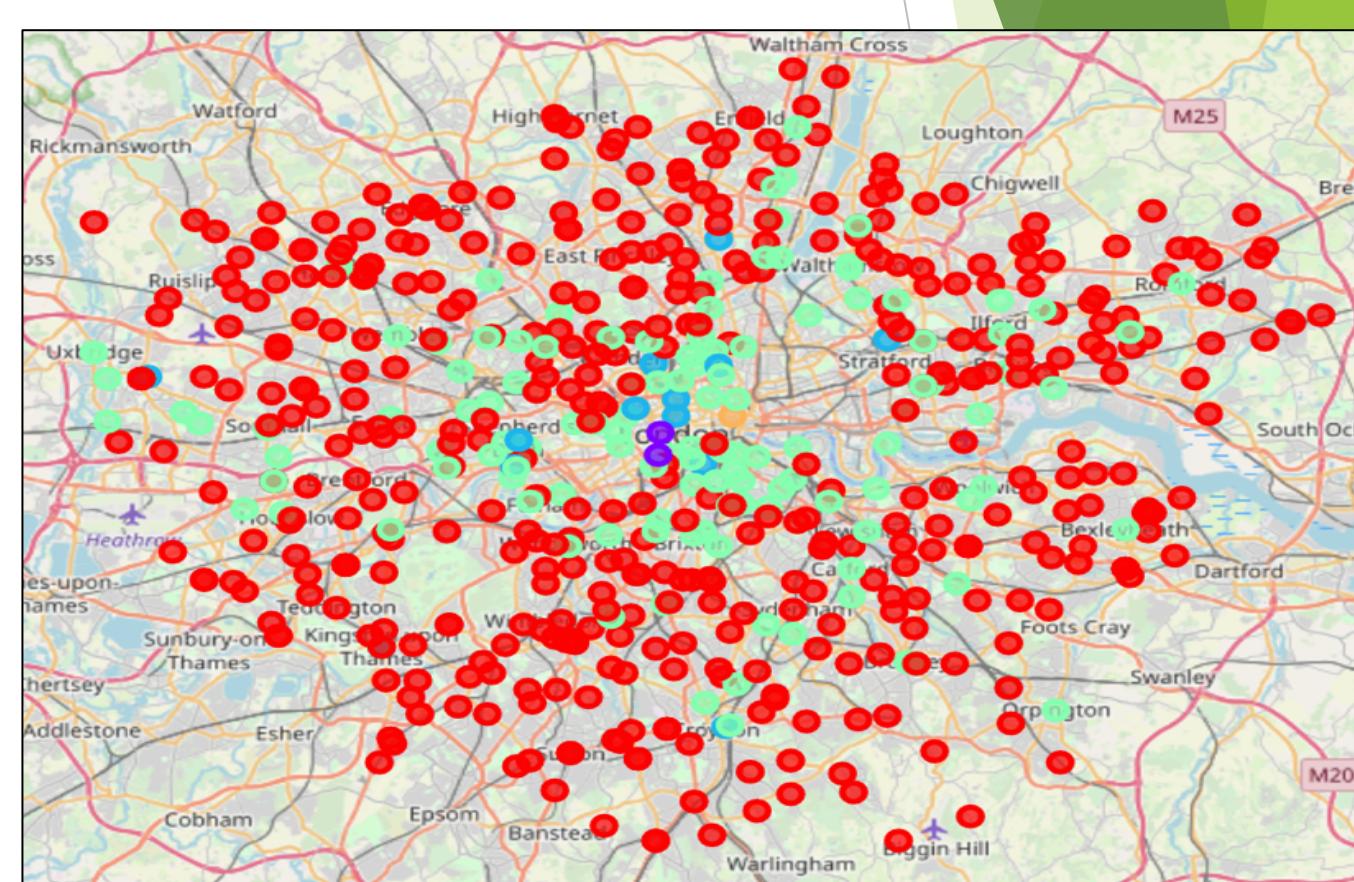
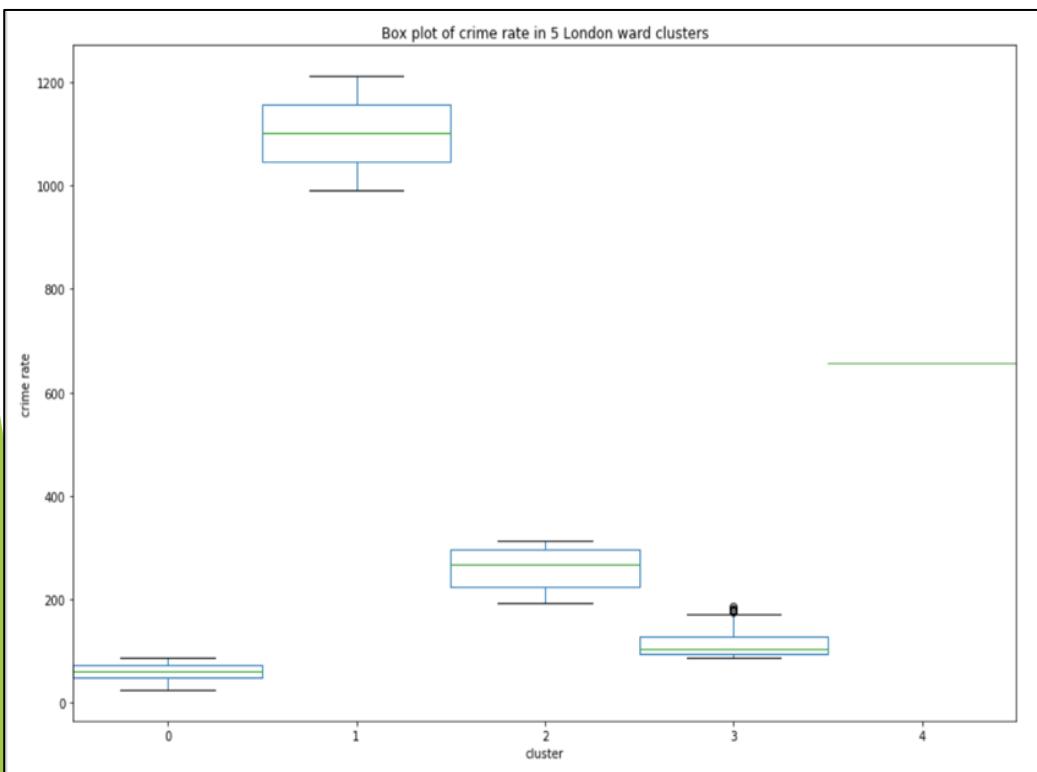


Cluster (colour for map)	Range of average GCSE point scores	Score
0 (red)	306.94 - 322.14 - below average	0
1 (purple)	338.09 - 357.23 - above average	4
2 (blue)	322.44 - 337.78 - low to average	2
3 (green)	275.95 - 306.49 - lowest level	0
4 (orange)	358.47 - 396.21 - highest level	5

Crime Rate

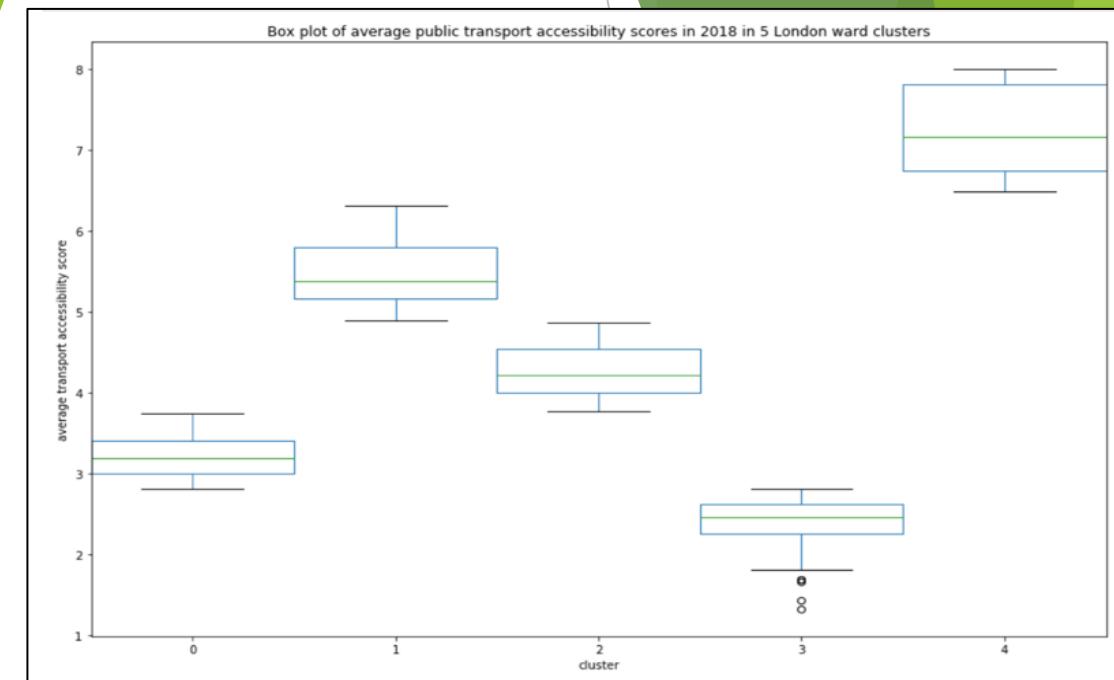
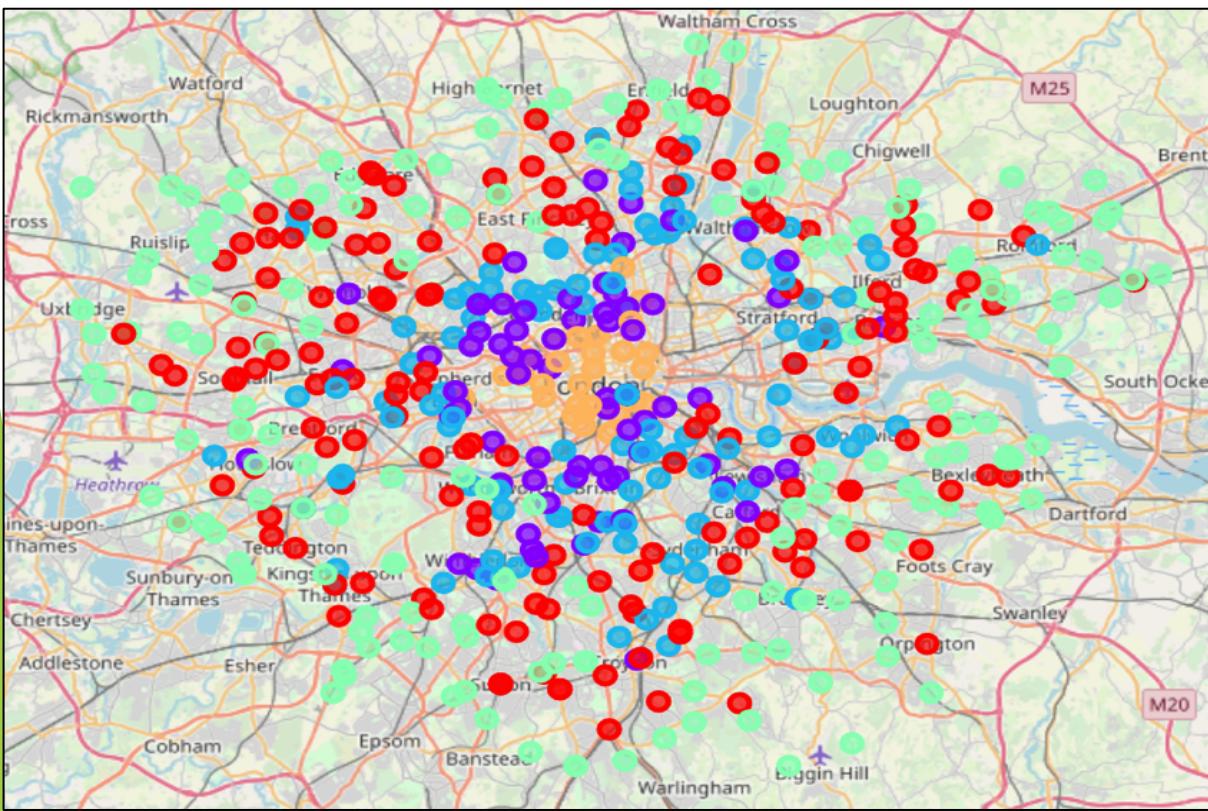
With children, safety is important. Crime rate has been used as a proxy for safety - the higher the crime rate, the less safe the ward. A crime rate of at or below the mean is considered acceptable and will generate a score for the cluster. Clusters that contain wards predominantly above the mean crime rate (80.87) will be assigned a value of 0

Cluster	Range of crime rates	Score
0 (red)	24.50 - 86.42 - lowest to average crime rate	5
1 (purple)	990.00 - 1212.13 - very high crime rate	0
2 (blue)	192.22 - 314.13 - high crime rate	0
3 (green)	86.64 - 184.91 - above average crime rate	0
4 (orange)	656.38 - highest crime rate	0



Public Transport Accessibility

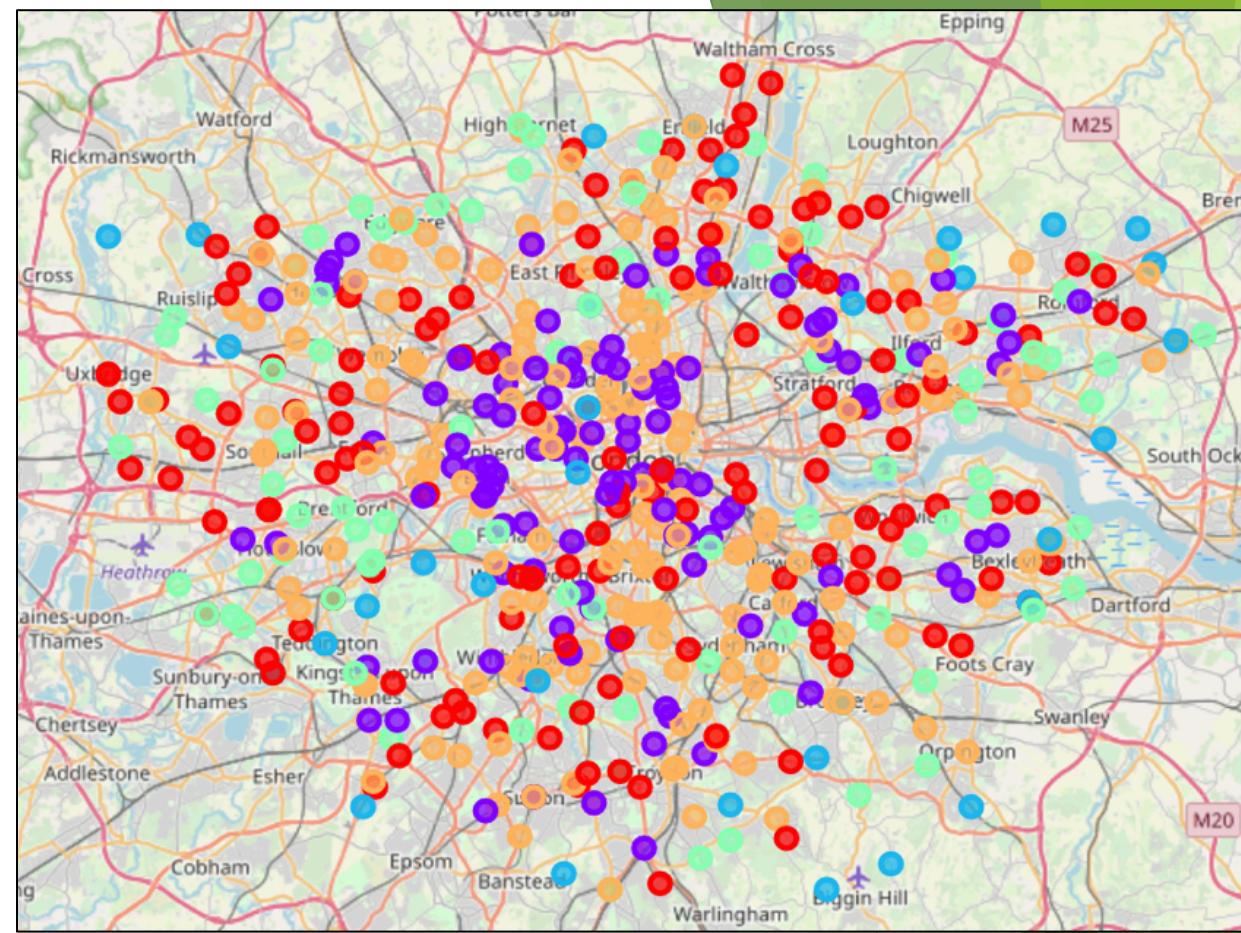
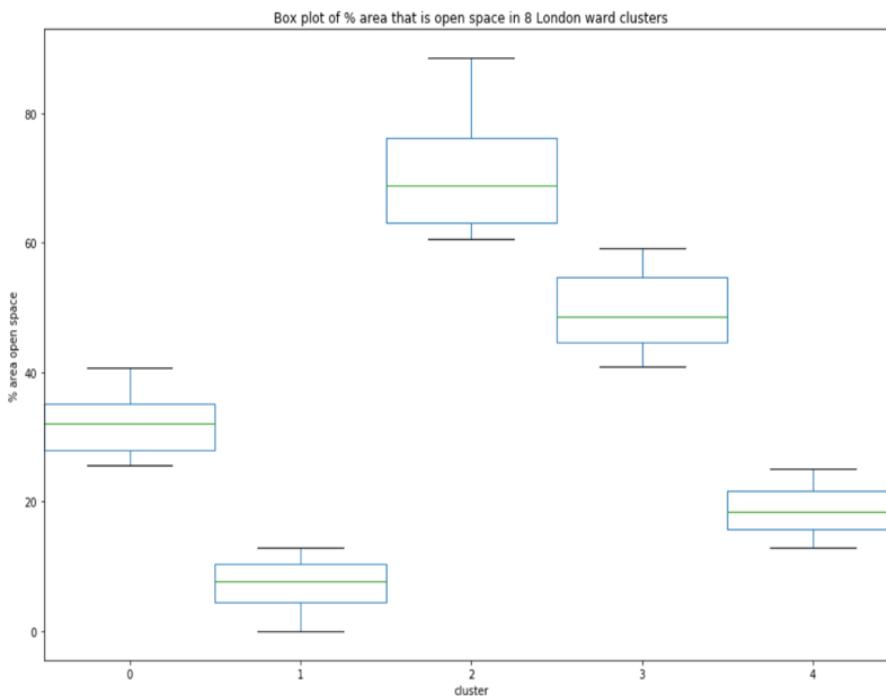
The wife has her main work office in central London. The requirement to commute into central London for work means access to public transport is necessary. Each area has been given an average score out of 8, with 8 being the highest level of accessibility. Clusters containing wards with the mean (3.63) or above is considered an acceptable level, with clusters below this being assigned a score of 0.



Cluster	Range of public transport accessibility scores	Score
0	2.82 - 3.74 - low to average	2
1	4.89 - 6.30 - high accessibility	4
2	3.77 - 4.86 - above average accessibility	3
3	1.32 - 2.81 - lowest accessibility	0
4	6.48 - 8.00 - highest accessibility	5

Open Space

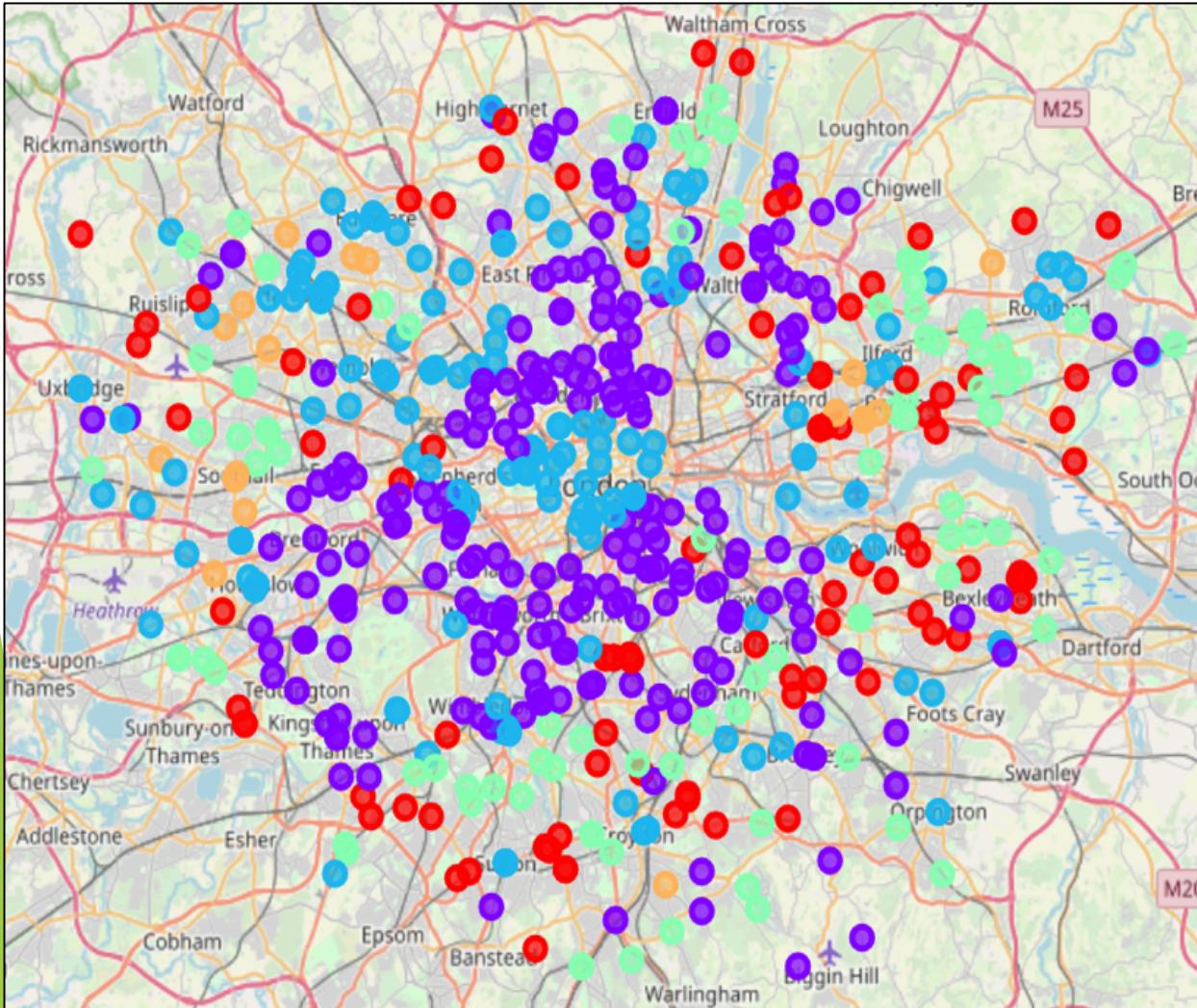
With three children and the husband being a personal trainer (offering the option of training clients outdoors), the property buyer expressed an interest in areas with open space. As they are moving from a rural area to London because they want to be in a more accessible area with options for the children, they want enough open space to meet their needs but not so much that they are just moving to yet another rural area. The mean percentage open space of London wards (27.6%) will be considered optimal.



Cluster	Range of % open space of wards	Score
0 (red)	25.62 - 40.61	5
1 (purple)	0.00 - 12.87	1
2 (blue)	60.63 - 88.53	1
3 (green)	40.95 - 59.08	3
4 (orange)	12.99 - 25.17	4

Venues

The property buyer is looking for a family area. They want an area with things to do with the children. With the husband being a personal trainer, an active area with plenty of gym and park access would be ideal.



Cluster	Venues/area character	Score
0 (red)	active/busy family cluster - a lot of places families are likely to be found, such as grocery stores/supermarkets, parks, fitness centres, garden centres, as well as family- or sports-type activities (e.g. Go Kart Track, soccer field, athletics and sports, ferry, pools, golf)	5
1 (purple)	young/business professional cluster young/business professional cluster - a plethora of places to eat, drink and socialise as well as a lot of gyms/fitness centres	0
2 (blue)	tourist/traveller/visitor cluster - a concentration of hotels, restaurants, cafes, shops, attractions (e.g. zoo, theatre, theme parks, galleries), and transportation options (e.g. airport lounge, bus stations, rail stations)	0
3 (green)	urban family cluster - a lot of places families are likely to be found, such as supermarkets/grocery stores, parks, family-type shops (e.g. furniture and clothing stores) and activities (e.g. multiplex, movie theatres, indoor play areas)	5
4 (orange)	eating in and dining out cluster - a concentration of restaurants and take away establishments	0

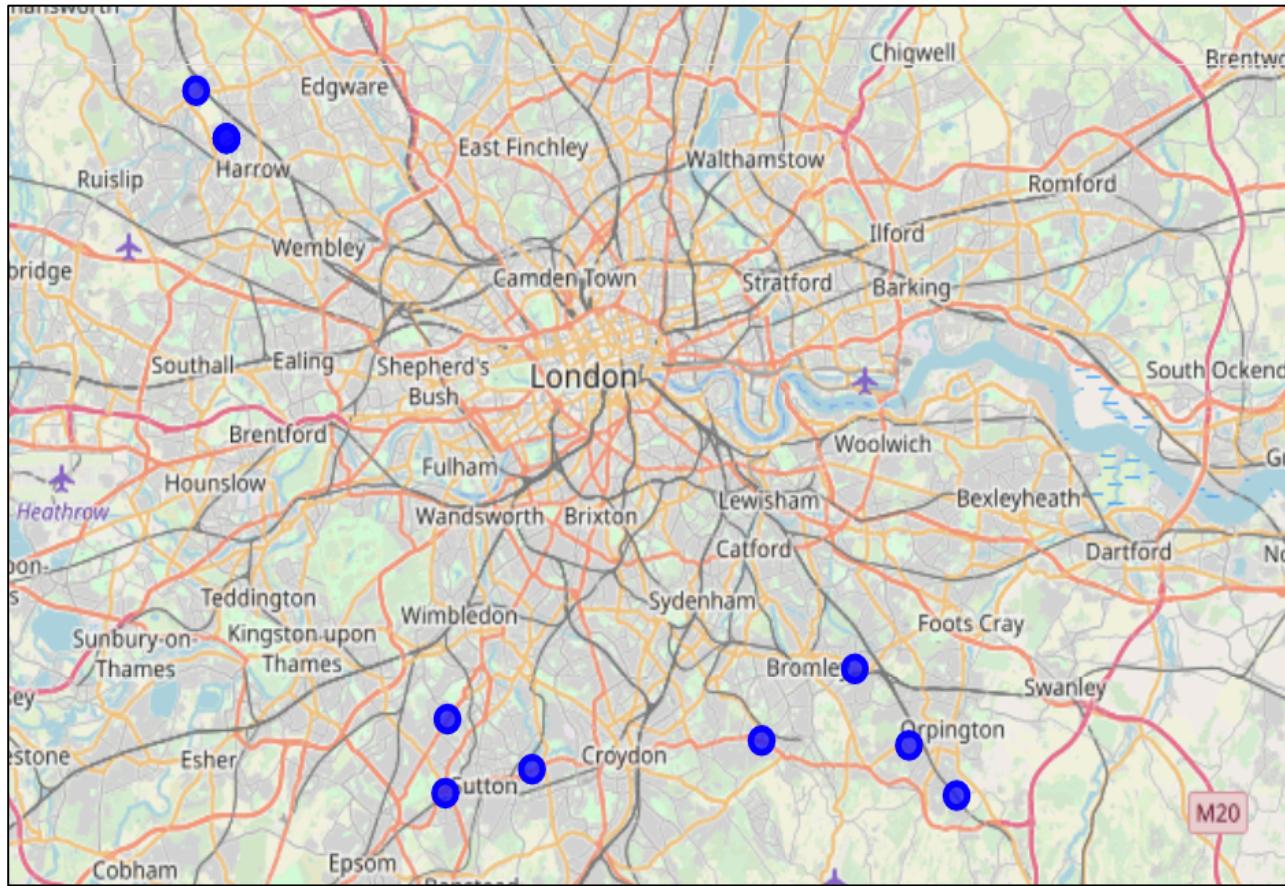
Results - applying the case study to the clusters

We assumed that the case study property buyers, when providing a list of criteria that are important to them when choosing a location to buy a property in, these were accompanied with a score of 1-10 to indicate the relative importance of each feature to them - the higher the score, the more important the criteria. These scores were used to generate a weighted score for each ward. The scores for features are as follows:

- ▶ Price - for those within budget 0.7
- ▶ Property type - 0.7
- ▶ Secondary education - 0.9
- ▶ Primary education - 0.9
- ▶ Safety/lack of crime - 0.8
- ▶ Transport accessibility - 0.8
- ▶ Open space - 0.5
- ▶ Area type (venues) - 1.0

To calculate the score for each ward, the score for individual features was multiplied by the weight of that feature and weighted feature scores were added together.

Recommendations - top 10 scoring wards



Ward loc	Score
Carshalton Central, Sutton	26.8
West Wickham, Bromley	26.6
Headstone North, Harrow	25.6
Bickley, Bromley	25.4
Headstone South, Harrow	25.4
Cheam, Sutton	25.4
Hatch End, Harrow	25.0
Farnborough and Crofton, Bromley	24.7
Chelsfield and Pratts Bottom, Bromley	24.6
Stonecot, Sutton	24.5

Recommendations - top 10 scoring wards

Observations

- All of the locations recommended are on the outskirts of Greater London. As distance from central London was not a feature specified by the buyer, this still means that the recommendations would be satisfactory to the buyer.
- The majority score highly for all features, except for accessibility to public transport. This points to a trade-off for those living in London, especially those with children - areas more suited to families, with better educational performances and lower crime rates have lower access to public transport. With the case study buyer requiring public transport access, low scores for transport accessibility may make these recommendations unsuitable. This would require further research for travel times from specific areas within wards.
- Open space was not as consistent as most other features. This is due to it being of a lower priority to the buyer. Given that the recommendations are all in outer Greater London, there is more open space, rather than less, than declared as optimal.

	Ward loc	Score	Latitude	Longitude	Price score	Property type score	Primary score	Secondary score	Crime score	Transport score	Space score	Venue score
0	Carshalton Central, Sutton	26.8	51.365788	-0.161086	4	4	4	5	5	2	5	5
1	West Wickham, Bromley	26.6	51.375804	-0.014684	5	5	4	5	5	0	5	5
2	Headstone North, Harrow	25.6	51.591883	-0.354217	5	5	4	5	5	0	3	5
3	Bickley, Bromley	25.4	51.401740	0.043712	5	4	4	5	5	0	4	5
4	Headstone South, Harrow	25.4	51.591883	-0.354217	4	4	4	4	5	2	4	5
5	Cheam, Sutton	25.4	51.357616	-0.216241	5	4	4	5	5	0	4	5
6	Hatch End, Harrow	25.0	51.608440	-0.373548	5	4	4	4	5	0	5	5
7	Farnborough and Crofton, Bromley	24.7	51.374000	0.077803	5	5	4	4	5	0	3	5
8	Chelsfield and Pratts Bottom, Bromley	24.6	51.356611	0.108328	5	5	4	5	5	0	1	5
9	Stonecot, Sutton	24.5	51.383526	-0.214193	4	5	4	4	5	0	4	5

Conclusion, limitations and future directions

This project identified some key variables that are important features people consider when deciding on a location within which to purchase a property - property price, property type, quality of education, crime rate, public transport accessibility, open space and the character of the area.

Feature clusters were applied to a case study property buyer and their preferences to generate a top ten list of recommended London wards to search for a property in. The top ten recommended wards were all characterised as family areas, with good primary and secondary education performance, low to average crime rate, within price range, and containing property types to the buyer's preference. Open space scores were more variable as it was considered less important by the buyer, and public transport accessibility consistently scored low for the recommendations.

Ideas for improving the project include:

- Adding in distance from central London.
- A limit to prevent features of importance that have been rated zero from being in the top recommendations.
- Further testing of the model could improve its performance. For example, additional cases with different needs and preferences could be tested.
- Additional variables could be added that may be important to other buyers; for example, some buyers may think healthcare services are an important consideration.
- Although the most up-to-date that could be obtained was used, it was still dated (most was 2014/15). For certain features, changes will have occurred between recording of the data and the use of it in this project. More recent data would provide more accurate results