Analysing London's Housing Market: Price Trends, Best Areas, and Future Predictions

1. Overview

1.1 Aim

The real estate market in London has seen significant changes over the years, with prices fluctuating due to various economic, social, and environmental factors. Understanding these trends is crucial for prospective homeowners, investors, and policymakers. This analysis aims to delve into the housing market data of London to uncover the underlying reasons for the increase in housing prices, identify the best and cheapest places to live, and predict future costs for purchasing a property.

1.2 Datasets Overview

Monthly Dataset: This dataset includes monthly records of housing market metrics such as average prices, the number of houses sold, and borough-specific identifiers. It allows for detailed trend analysis and understanding of short-term fluctuations in the market.

Yearly Dataset: This dataset provides annual socio-economic and demographic metrics, including median and mean salaries, recycling percentages, population size, and the number of jobs. These factors are crucial for understanding the broader economic context and living conditions in various boroughs.

2. Data Cleaning and Preparation

2.1 Initial Inspection

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 13549 entries, 0 to 13548
Data columns (total 7 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   date           13549 non-null  object
 1   area           13549 non-null  object
 2   average_price  13549 non-null  int64
 3   code           13549 non-null  object
 4   houses_sold    13455 non-null  float64
 5   no_of_crimes   7439 non-null   float64
 6   borough_flag   13549 non-null  int64
dtypes: float64(2), int64(2), object(3)
memory usage: 741.1+ KB
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1071 entries, 0 to 1070
Data columns (total 12 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   code              1071 non-null   object
 1   area              1071 non-null   object
 2   date              1071 non-null   object
 3   median_salary     1049 non-null   float64
 4   life_satisfaction 352 non-null    float64
 5   mean_salary       1071 non-null   object
 6   recycling_pct     860 non-null    object
 7   population_size   1018 non-null   float64
 8   number_of_jobs    931 non-null    float64
 9   area_size         666 non-null    float64
 10  no_of_houses      666 non-null    float64
 11  borough_flag      1071 non-null   int64
dtypes: float64(6), int64(1), object(5)
memory usage: 100.5+ KB
```

Image 1: Monthly Dataset Information   Image 2: Yearly Dataset Information

```
date               0
area               0
average_price      0
code               0
houses_sold       94
no_of_crimes    6110
borough_flag       0
dtype: int64
```

```
code                 0
area                 0
date                 0
median_salary       22
life_satisfaction  719
mean_salary          0
recycling_pct      211
population_size     53
number_of_jobs     140
area_size          405
no_of_houses       405
borough_flag         0
dtype: int64
```

**Image 3: Monthly Dataset
Missing Values**

**Image 4: Yearly Dataset
Missing Values**

```
Outliers in Monthly Data (IQR):
 average_price     180
houses_sold        138
dtype: int64
```

```
Outliers in Yearly Data (IQR):
 median_salary      21
recycling_pct        0
population_size      0
number_of_jobs      50
area_size           54
no_of_houses         3
dtype: int64
```

Image 5: Monthly Dataset Outliers

Image 6: Yearly Dataset Outliers

During my time understanding my dataset, in **Image 1** and **Image 2**, there is one key data type that needed to be changed which was the date. If this was not changed, it would have several impacts on my data analysis and manipulation. It would affect date-based operations, sorting, visualisations and performance due to the fact that it would require more complex string comparisons and times series plots would not work properly.

Additionally with **Image 3** and **Image 4**, I was able to identify that my data set for monthly and yearly had missing values. These missing values would cause my analysis to be biassed if we had non-random missing data or my variables may be underrepresented. Missing values could also affect my sample size which will lower the statistical power of my analysis, leading me to not detecting any significant relationships. This could further cause my visualisations to be misleading or incomplete. Properly addressing missing values ensures more reliable and valid results in my data analysis.

As seen in **Image 5** and **Image 6,** there are many outliers within my dataset. Outliers would affect my mean, standard deviation and range. It would skew my mean, which means my central tendency would be unrepresentative, or it would increase my standard deviation

which indicates more variability, or increase my range which would give me the wrong spread for my data. This would lead to incorrect interpretations of relationships between variables if not dealt with. Properly detecting and managing outliers ensures more reliable and accurate analysis.

2.2 Cleaning Steps

The data cleaning process involved several crucial steps to ensure the datasets were properly prepared for analysis. Here's a detailed description of the steps taken:

2.2.1 Handling Missing Values

Monthly Data:

- **no_of_crimes**: Missing values were filled with 0, assuming no crimes were reported in those instances.
- **houses_sold** - Missing values were filled with the median value of the houses_sold column to maintain consistency and avoid data loss.

Yearly Data:

- **number_of_jobs**: Missing values were filled with the median value of the number_of_jobs column for consistency.
- **area_size**: Missing values were filled with the median value of the area_size column.
- **no_of_houses**: Missing values were filled with the median value of the no_of_houses column.
- **recycling_pct:** Missing values were filled with the median value of the houses_sold column to maintain consistency and avoid data loss.
- **population_size:** Missing values were filled with the median value of the houses_sold column to maintain consistency and avoid data loss.
- **median_salary:** Missing values were filled with the median value of the houses_sold column to maintain consistency and avoid data loss.

2.2.2 Converting Data Types

- **`date` Column**:
    - For both monthly and yearly datasets, the `date` column was converted to `datetime64[ns]` to enable proper date-based operations, filtering, and visualisations.
- All numerical columns were converted into numeric format that ensures that your data is correctly and efficiently handled throughout the analysis process.

2.2.3 Column Drop

Monthly Data:

- **borough_flag**: The column has been removed because it is not pertinent to my hypothesis.

Yearly Data:

- **borough_flag**: The column has been removed because it is not pertinent to my hypothesis.
- **life_satisfaction:** The column was removed because the variable showed little variation across all areas.

2.2.4 Filtering

- **area -** The dataset included areas outside of London, so I filtered it to include only London boroughs.

2.2.5 Removing Outliers

Monthly Data:

- Outliers in the numerical columns (average_price, houses_sold, no_of_crimes) were detected and removed using the Interquartile Range (IQR) method:
    - **Detection**: Calculated the first quartile (Q1) and third quartile (Q3) for each numerical column, computed the IQR (Q3 - Q1), and determined the lower and upper bounds (Q1 - 1.5 * IQR, Q3 + 1.5 * IQR).
    - **Removal**: Filtered out rows with values outside these bounds.

Yearly Data:

Outliers in the numerical columns (median_salary, life_satisfaction, mean_salary, recycling_pct, population_size, number_of_jobs, area_size, no_of_houses) were detected and removed using the Interquartile Range (IQR) method:

- Detection: Calculated the first quartile (Q1) and third quartile (Q3) for each numerical column, computed the IQR (Q3 - Q1), and determined the lower and upper bounds (Q1 - 1.5 * IQR, Q3 + 1.5 * IQR).
- Removal: Filtered out rows with values outside these bounds.

The data cleaning process involved handling missing values, converting data types, dropping columns, filtering and removing outliers to ensure accurate and reliable analysis. These steps are essential to prepare the data for meaningful insights and visualisations.

3. Exploratory Data Analysis (EDA)

3.1 Summary Statistics

	average_price	houses_sold	no_of_crimes	year
count	7697.000000	7697.000000	7697.000000	7697.000000
mean	256660.623490	333.780694	1656.272574	2006.919059
std	134858.143076	127.429524	1103.186998	7.251734
min	53154.000000	58.000000	0.000000	1995.000000
25%	150244.000000	236.000000	857.000000	2001.000000
50%	238119.000000	318.000000	1979.000000	2007.000000
75%	339849.000000	416.000000	2480.000000	2013.000000
max	660512.000000	704.000000	4875.000000	2020.000000

Image 7: Monthly Key Statistics

Average Price

- **Mean**: 256660.62
  - The average house price across all recorded months and areas is approximately £256,661. This value provides a general sense of the house price level in the dataset.
- **Median**: 238119.00
  - The middle value of the average price is £238,119, indicating that half of the months have an average house price below this value and half above. The median being slightly lower than the mean suggests a slight right skew in the price distribution.
- **Standard Deviation**: 134858.14
  - This value indicates the variability of house prices around the mean. A standard deviation of £134,858.14 shows a significant spread, meaning there is considerable variation in house prices among different months and areas.

Houses Sold

- **Mean**: 333.78
  - On average, 333.78 houses are sold per month across all areas.
- **Median**: 318.00
  - The middle value of the houses sold is 318, indicating that half of the months have fewer than 318 houses sold and half have more. The median being close to the mean suggests a relatively symmetric distribution.
- **Standard Deviation**: 127.43
  - This value shows the variability in the number of houses sold. A standard deviation of 127.43 indicates significant differences in sales volumes across different months or areas.

Number of Crimes

- **Mean**: 1656.27

- ○ On average, 1656.27 crimes are reported per month across all areas.
- **Median**: 1979.00
  - ○ The middle value of the number of crimes is 1979, indicating that half of the months have fewer than 1979 crimes reported and half have more. The median being higher than the mean suggests a left skew in the crime distribution.
- **Standard Deviation**: 1103.19
  - ○ This value shows the variability in the number of crimes reported. A standard deviation of 1103.19 indicates considerable differences in crime rates among different months or areas.

	median_salary	recycling_pct	population_size	number_of_jobs	area_size	no_of_houses
count	409.000000	409.000000	409.000000	409.000000	409.000000	409.000000
mean	27890.440098	24.053790	256670.650367	118894.865526	4278.953545	102887.985330
std	4604.194507	12.410939	42568.260208	42524.228366	1456.911480	13754.188259
min	16720.000000	1.000000	175717.000000	66000.000000	1486.000000	77701.000000
25%	24626.000000	14.000000	222015.000000	81000.000000	3532.000000	94682.000000
50%	28005.000000	25.000000	256269.000000	110000.000000	4323.000000	102402.000000
75%	30718.000000	33.000000	291933.000000	157000.000000	5046.000000	110145.000000
max	40303.000000	55.000000	352005.000000	256000.000000	8220.000000	136558.000000

Image 8: Yearly Key Statistics

Median Salary

- **Mean**: 27890.44
  - ○ The average median salary across all recorded areas is approximately £27,890. This value provides a general sense of the income level in the dataset.
- **Median**: 28005.00
  - ○ The middle value of the median salary is £28,005, indicating that half of the areas have a median salary below this value and half above. The median being close to the mean suggests a relatively symmetric distribution.
- **Standard Deviation**: 4604.19
  - ○ This value indicates the variability of the median salaries around the mean. A standard deviation of £4604.19 shows a moderate spread, meaning there is some variability in median salaries among different areas.

Recycling Percentage

- **Mean**: 24.05
  - ○ On average, 24.05% of waste is recycled across the areas.
- **Median**: 25.00

- - The middle value of the recycling percentage is 25%, indicating that half of the areas have a recycling rate below 25% and half above.
  - **Standard Deviation**: 12.41
    - This value shows the variability in recycling rates. A standard deviation of 12.41 indicates a moderate spread, suggesting differences in recycling practices among areas.

Population Size

- **Mean**: 256670.65
  - The average population size across all recorded areas is approximately 256,671 people.
- **Median**: 256629.00
  - The middle value of the population size is 256,629, indicating that half of the areas have a population below this number and half above. The median being very close to the mean suggests a relatively symmetric distribution.
- **Standard Deviation**: 42568.26
  - This value indicates the variability in population sizes. A standard deviation of 42,568.26 shows significant differences in population sizes among different areas.

Number of Jobs

- **Mean**: 118894.87
  - On average, there are 118,894 jobs in each area.
- **Median**: 110000.00
  - The middle value of the number of jobs is 110,000, indicating that half of the areas have fewer than 110,000 jobs and half have more.
- **Standard Deviation**: 42524.23
  - This value shows the variability in the number of jobs. A standard deviation of 42,524.23 indicates considerable differences in job numbers among different areas.

Area Size

- **Mean**: 4278.95
  - The average area size across all recorded areas is approximately 4278.95 square kilometres.
- **Median**: 4323.00
  - The middle value of the area size is 4323 square kilometres, indicating that half of the areas have an area size below this value and half above.
- **Standard Deviation**: 1456.91
  - This value indicates the variability in area sizes. A standard deviation of 1456.91 shows considerable differences in area sizes among different regions.

Number of Houses

- **Mean**: 102887.99
  - On average, there are 102,887 houses in each area.
- **Median**: 102402.00
  - The middle value of the number of houses is 102,402, indicating that half of the areas have fewer than 102,402 houses and half have more. The median being close to the mean suggests a relatively symmetric distribution.
- **Standard Deviation**: 13754.19
  - This value shows the variability in the number of houses. A standard deviation of 13,754.19 indicates significant differences in the number of houses among different areas.
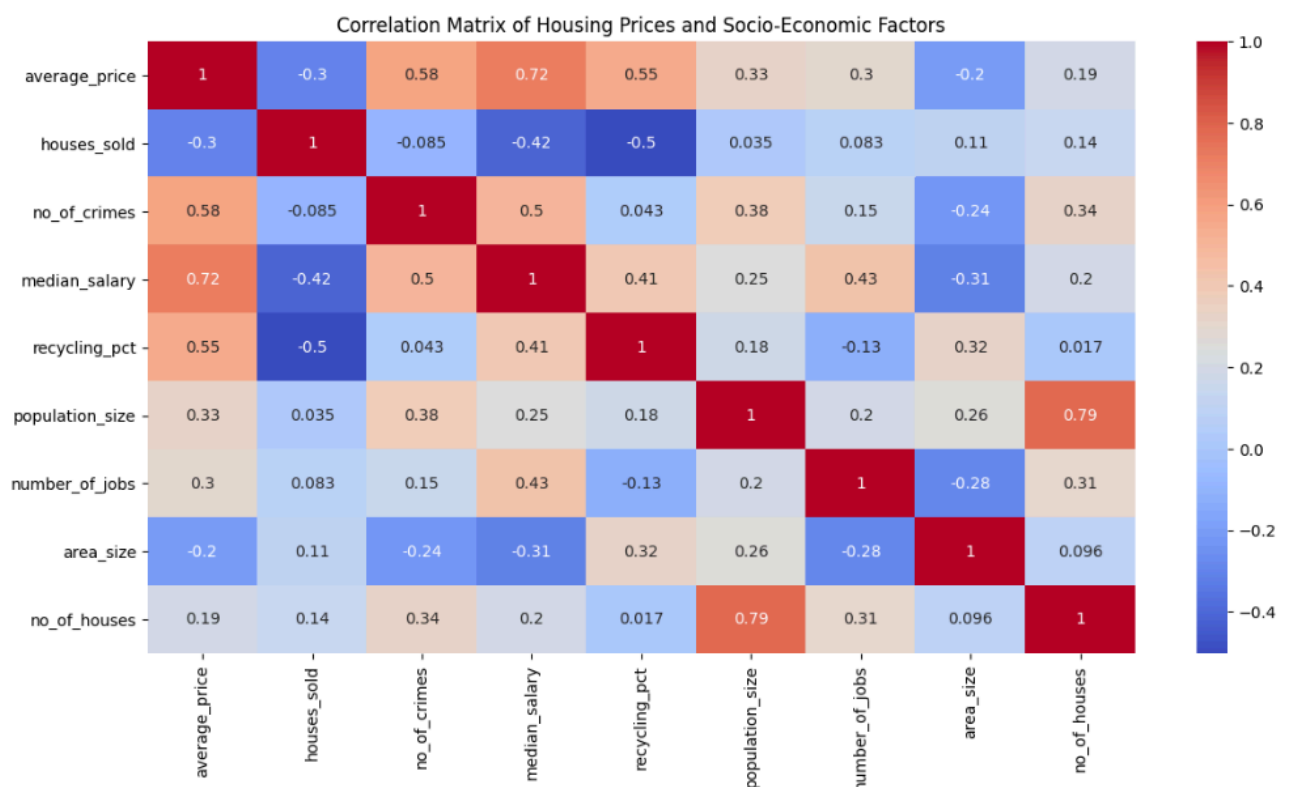
3.2 Correlation Analysis

Correlation Matrix of Housing Prices and Socio-Economic Factors

	average_price	houses_sold	no_of_crimes	median_salary	recycling_pct	population_size	number_of_jobs	area_size	no_of_houses
average_price	1	-0.3	0.58	0.72	0.55	0.33	0.3	-0.2	0.19
houses_sold	-0.3	1	-0.085	-0.42	-0.5	0.035	0.083	0.11	0.14
no_of_crimes	0.58	-0.085	1	0.5	0.043	0.38	0.15	-0.24	0.34
median_salary	0.72	-0.42	0.5	1	0.41	0.25	0.43	-0.31	0.2
recycling_pct	0.55	-0.5	0.043	0.41	1	0.18	-0.13	0.32	0.017
population_size	0.33	0.035	0.38	0.25	0.18	1	0.2	0.26	0.79
number_of_jobs	0.3	0.083	0.15	0.43	-0.13	0.2	1	-0.28	0.31
area_size	-0.2	0.11	-0.24	-0.31	0.32	0.26	-0.28	1	0.096
no_of_houses	0.19	0.14	0.34	0.2	0.017	0.79	0.31	0.096	1

Image 9: Monthly & Yearly Heat Map Correlation

3.2.1 Key Correlations

Average Price

- **Median Salary (0.72)**: There is a strong positive correlation between average house prices and median salaries. This suggests that areas with higher median salaries tend to have higher house prices, which is expected as higher incomes can drive up housing demand and prices.
- **No of Crimes (0.58)**: There is a moderate positive correlation between average house prices and the number of crimes, which is somewhat counterintuitive as higher crime rates are typically associated with lower property values. This might suggest

that areas with higher house prices have better reporting and recording of crimes or other underlying factors.

- **Recycling Percentage (0.55)**: There is a moderate positive correlation between average house prices and recycling percentages. This could imply that wealthier areas are more engaged in environmentally friendly practices.
- **Population Size (0.33)**: There is a weak positive correlation between average house prices and population size, indicating that larger populations might slightly contribute to higher house prices.
- **Houses Sold (-0.3)**: There is a weak negative correlation between average house prices and houses sold. This suggests that areas with higher house prices might see fewer transactions, possibly due to affordability issues.

Houses Sold

- **Recycling Percentage (-0.5)**: There is a moderate negative correlation between houses sold and recycling percentages. This might indicate that areas with higher recycling rates have fewer house sales, possibly due to stable populations and lower turnover.
- **Median Salary (-0.42)**: There is a moderate negative correlation between houses sold and median salaries, suggesting that areas with higher median salaries might have fewer transactions, potentially due to higher housing stability and lower mobility.

No of Crimes

- **Median Salary (0.5)**: There is a moderate positive correlation between the number of crimes and median salaries, which again is counterintuitive and might suggest other underlying factors.
- **Population Size (0.38)**: There is a weak to moderate positive correlation between the number of crimes and population size, which is expected as larger populations might naturally have higher crime rates due to sheer numbers.
- **Number of Houses (0.34)**: There is a weak positive correlation between the number of crimes and the number of houses, indicating that areas with more houses might experience slightly more crimes.

Median Salary

- **Number of Jobs (0.43)**: There is a moderate positive correlation between median salaries and the number of jobs, suggesting that areas with more job opportunities tend to have higher median salaries.
- **Recycling Percentage (0.41)**: There is a weak to moderate positive correlation between median salaries and recycling percentages, indicating that wealthier areas might be more environmentally conscious.
- **Area Size (-0.31)**: There is a weak negative correlation between median salaries and area size, indicating that larger areas might have slightly lower median salaries.

Population Size

- **Number of Houses (0.79)**: There is a strong positive correlation between population size and the number of houses, indicating that larger populations naturally have more houses.

4. Analysis and Insights

4.1 Monthly Data Analysis
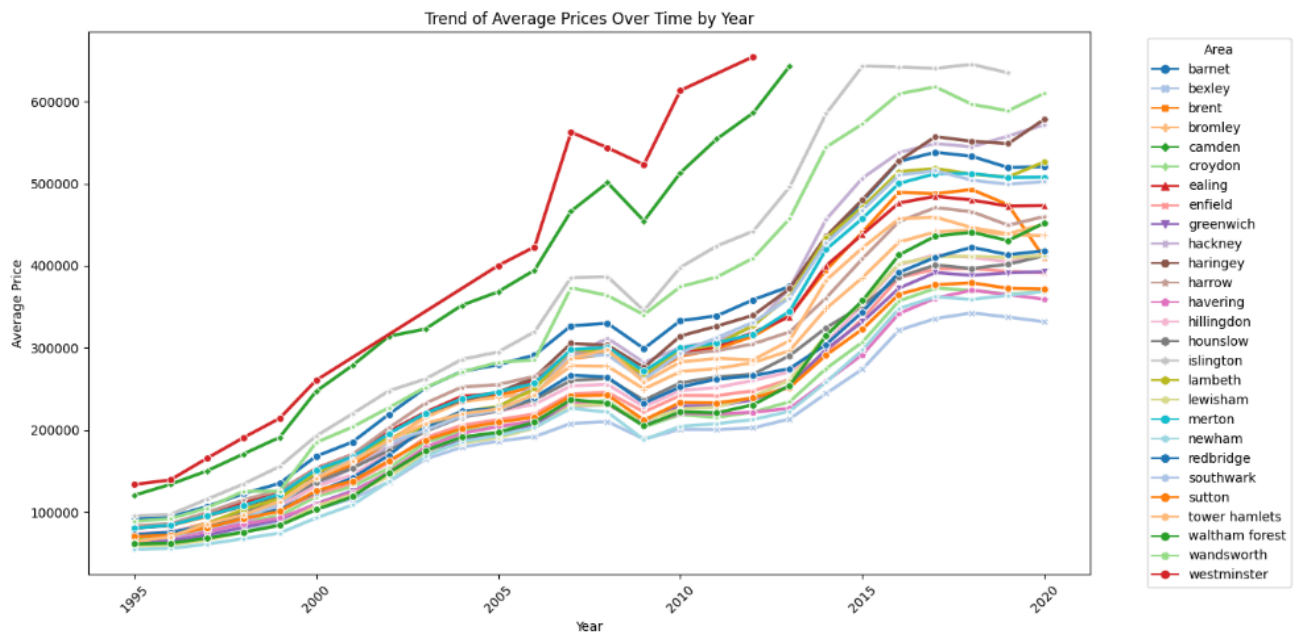
4.1.1 Housing Prices

Image 10: Trend of Average Prices Over Time by Year

4.1.1.1 Key Observations:

Overall Trends:

- The general trend shows an increase in average housing prices over the years across all areas.
- Some areas exhibit sharper increases compared to others, indicating varying levels of market dynamics and demand.

Top Areas with Highest Price Increases:

- **Westminster**: Shows significant growth, especially notable around 2007 and 2010, with prices peaking at £654,620 in 2012.
- **Wandsworth**: Steady increase with significant jumps around 2007 and 2010, peaking at £610,256 in 2020.
- **Camden**: Sharp increase, especially around 2007 and 2010, with prices reaching £643,130 in 2013.

Moderate Price Increases:

- **Barnet**: Consistent growth with some fluctuations, peaking at £538,280 in 2017.
- **Brent**: Steady increase with significant fluctuations, peaking at £492,845 in 2018.

Areas with Lower Increases:

- **Lewisham**: Steady growth with moderate increases, peaking at £412,522 in 2020.
- **Croydon**: Consistent growth with smaller fluctuations, peaking at £370,298 in 2018.

Notable Observations:

- **Economic Impact**: The financial crisis around 2008-2009 is visible with dips or slowed growth in many areas.
- **Recovery Period**: Post-2010, many areas show recovery with significant increases in prices.

Overall Growth: All areas show a general upward trend in housing prices, reflecting broader market dynamics and economic conditions.

Significant Increases: Areas like Westminster, Wandsworth, and Camden show the highest increases, indicating high demand and possibly better economic conditions.

Moderate Increases: Areas like Barnet, Brent, and Croydon show steady but moderate increases, indicating stable growth.

Economic Impacts: The 2008-2009 financial crisis impact is visible with slowed growth or dips in many areas, followed by recovery post-2010.
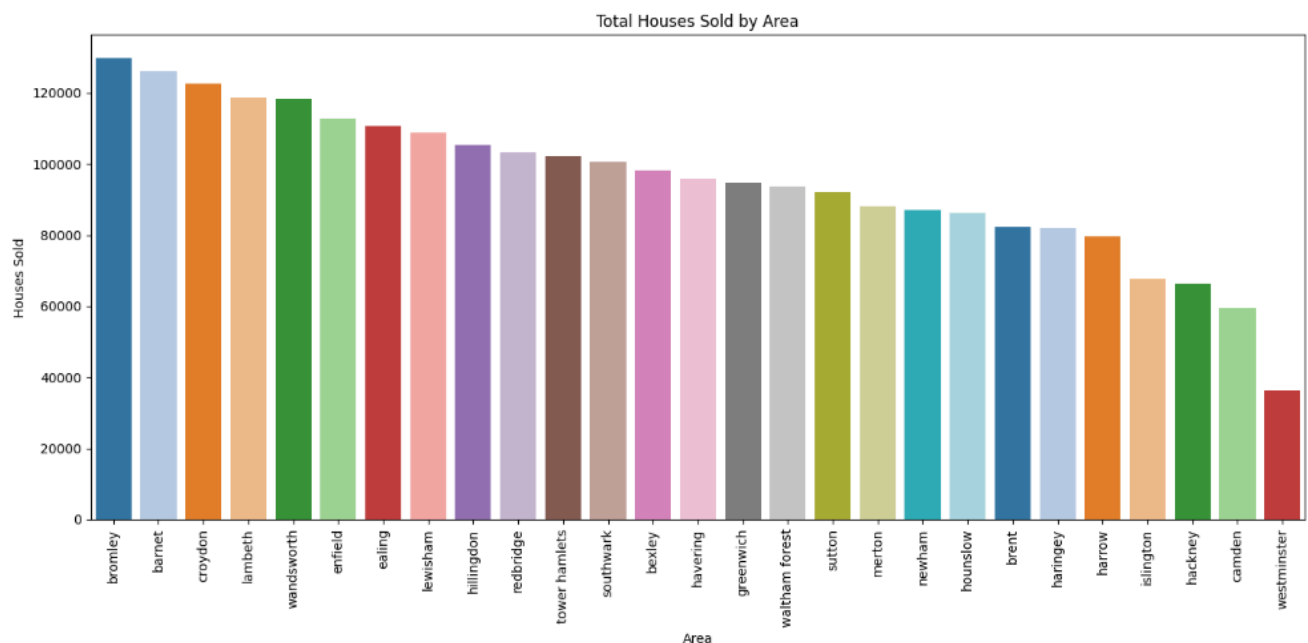
4.1.2 Sales Volume

Image 11: Total Houses Sold by Area

4.1.2.1 Key Observations:

1. **Top Areas by Number of Houses Sold:**
   - Bromley: 129,829 houses sold
   - Barnet: 126,056 houses sold
   - Croydon: 122,490 houses sold
   - Lambeth: 118,695 houses sold
   - Wandsworth: 118,434 houses sold

These areas have the highest number of houses sold, indicating significant market activity.

2. **Moderate Areas by Number of Houses Sold:**
   - Enfield: 112,664 houses sold
   - Ealing: 110,784 houses sold
   - Lewisham: 108,865 houses sold
   - Hillingdon: 105,463 houses sold
   - Redbridge: 103,253 houses sold

These areas also have a substantial number of houses sold, showing they are important residential markets.

3. **Areas with Lower Numbers of Houses Sold:**
   - Sutton: 92,141 houses sold
   - Newham: 86,978 houses sold
   - Hounslow: 86,214 houses sold
   - Brent: 82,355 houses sold
   - Westminster: 36,409 houses sold

These areas have the lowest number of houses sold, indicating they are less active residential markets.

4.1.2.2 Detailed Observations:

1. **Top Residential Areas:**
   - Bromley has the highest number of houses sold (129,829), making it the most significant residential area in this dataset.
   - Barnet and Croydon follow closely, with 126,056 and 122,490 houses sold, respectively. These areas are also major residential zones with high market activity.
2. **Mid-Range Residential Areas:**
   - Areas like Enfield, Ealing, and Lewisham have around 100,000 houses sold each, indicating substantial but slightly lower market activity compared to the top areas.
   - The distribution in this mid-range is relatively close, indicating similar levels of residential activity.
3. **Lower Residential Areas:**

○ Sutton has 92,141 houses sold, which is significantly lower than the top areas.
○ Newham and Hounslow have around 86,000 houses sold each, indicating lower levels of market activity.

4. **Least Residential Areas:**
   ○ Westminster has the lowest number of houses sold (36,409), making it the least significant residential area in this dataset. This could be due to various factors such as higher property prices, commercial zoning, or smaller geographic areas.

The top residential areas like Bromley, Barnet, and Croydon have the highest sales counts, indicating significant market activity.

Mid-range areas like Enfield, Ealing, and Lewisham have substantial but slightly lower sales numbers.

Lower residential areas like Sutton, Newham, and Westminster have significantly fewer sales, indicating less active markets.

4.1.3 No. of Crimes
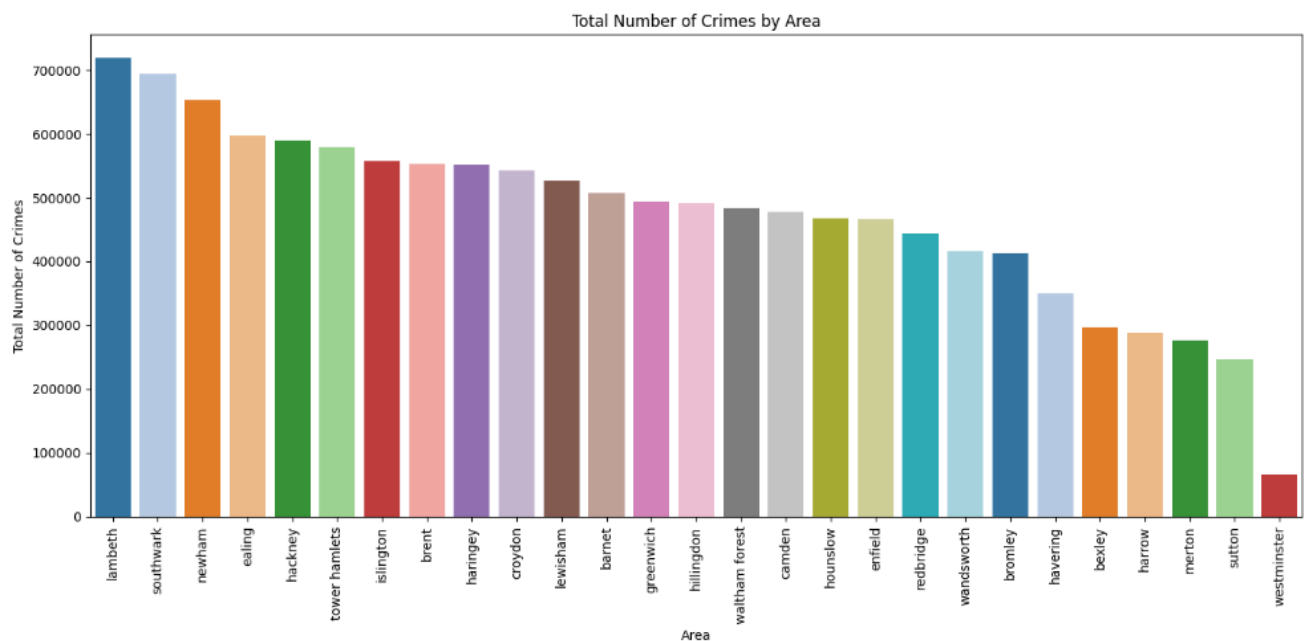
Image 12: Total Number of Crimes by Area

4.1.3.1 Key Observations:

1. **Top Areas by Number of Crimes:**
   ○ Lambeth: 719,454 crimes
   ○ Southwark: 693,541 crimes
   ○ Newham: 652,978 crimes
   ○ Ealing: 597,017 crimes

○ Hackney: 589,847 crimes

These areas have the highest number of crimes, indicating significant criminal activity.

2. **Moderate Areas by Number of Crimes:**
    ○ Tower Hamlets: 579,110 crimes
    ○ Islington: 556,991 crimes
    ○ Brent: 553,173 crimes
    ○ Haringey: 551,072 crimes
    ○ Croydon: 542,187 crimes

These areas also have a substantial number of crimes, showing they are areas with moderate criminal activity.

3. **Areas with Lower Numbers of Crimes:**
    ○ Redbridge: 444,370 crimes
    ○ Wandsworth: 416,173 crimes
    ○ Bromley: 413,809 crimes
    ○ Havering: 395,452 crimes
    ○ Westminster: 66,403 crimes

These areas have the lowest number of crimes, indicating they are areas with relatively low criminal activity.

4.1.3.2 Detailed Observations:

1. **Top Crime Areas:**
    ○ Lambeth has the highest number of crimes (719,454), making it the area with the most significant criminal activity in this dataset.
    ○ Southwark and Newham follow closely, with 693,541 and 652,978 crimes, respectively. These areas are also major hotspots for criminal activity.
2. **Mid-Range Crime Areas**:
    ○ Areas like Tower Hamlets, Islington, and Brent have between 550,000 and 580,000 crimes. These areas are important but have slightly lower criminal activity compared to the top areas.
    ○ The distribution in this mid-range is relatively close, indicating similar levels of criminal activity.
3. **Lower Crime Areas:**
    ○ Redbridge has 444,370 crimes, which is significantly lower than the top areas.
    ○ Wandsworth and Bromley have around 400,000 crimes each, indicating lower levels of criminal activity.
4. **Least Crime Areas:**
    ○ Westminster has the lowest number of crimes (66,403), making it the least significant area for criminal activity in this dataset. This could be due to various factors such as better policing, smaller geographic area, or lower population density.

The top areas like Lambeth, Southwark, and Newham have the highest crime counts, indicating significant criminal activity.

Mid-range areas like Tower Hamlets, Islington, and Brent have substantial but slightly lower crime numbers.

Lower crime areas like Redbridge, Wandsworth, and Westminster have significantly fewer crimes, indicating relatively safer regions.

4.2 Yearly Data Analysis
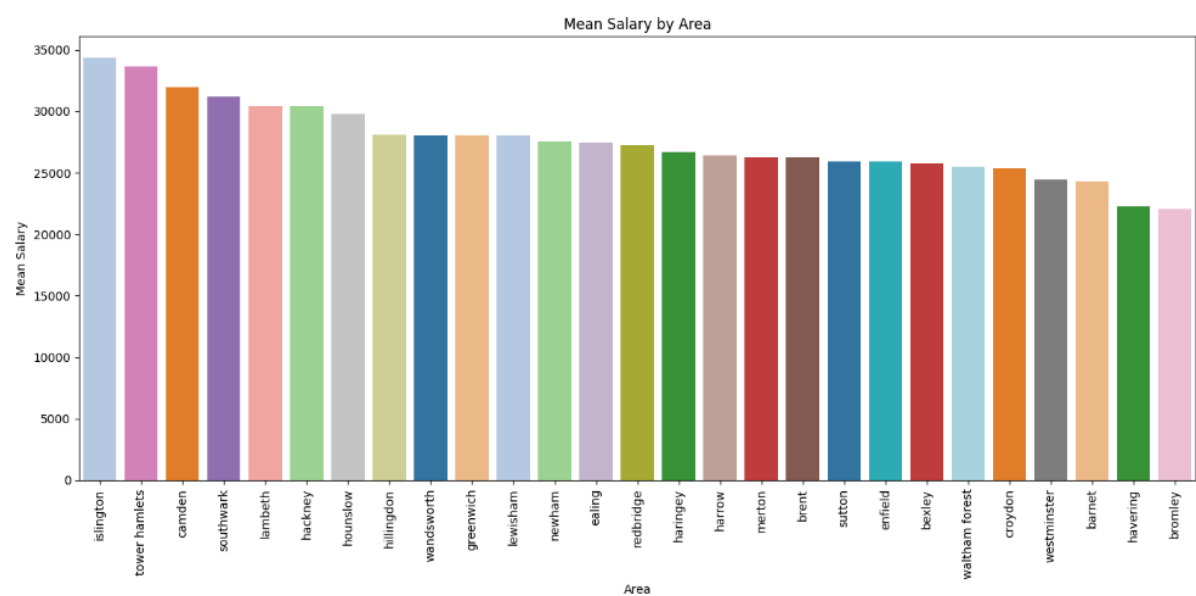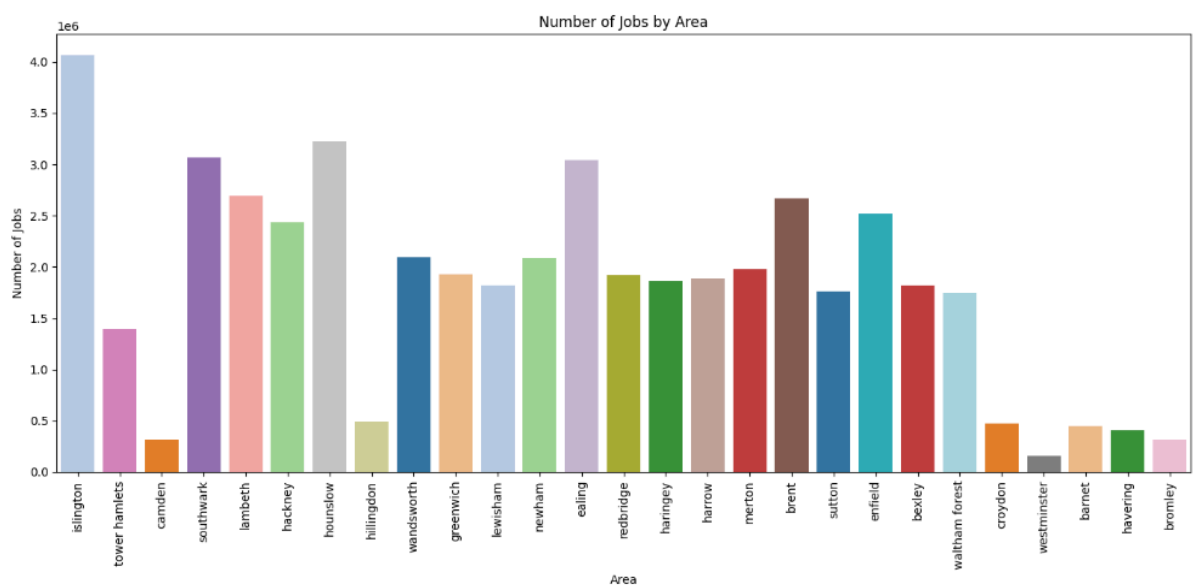
4.2.1 Income & Employment

Image 13: Mean Salary by Area

Image 14: Number of Jobs by Area

4.2.1.1 Key Observations:

1. **Top Areas by Mean Salary:**
   - Islington: £34,366.40
   - Tower Hamlets: £33,637.25
   - Camden: £31,936.50
   - Southwark: £31,200.60
   - Lambeth: £30,444.39

These areas have the highest mean salaries, indicating higher earning potential in these regions.

2. **Moderate Areas by Mean Salary:**
   - Hackney: £30,837.43
   - Hounslow: £29,801.67
   - Hillingdon: £28,809.33
   - Wandsworth: £28,440.44
   - Greenwich: £28,010.05

These areas have moderate mean salaries, showing they are areas with good earning potential.

3. **Areas with Lower Mean Salaries:**
   - Westminster: £24,447.00
   - Barnet: £24,317.67
   - Havering: £23,801.33
   - Bromley: £22,080.50

These areas have the lowest mean salaries, indicating lower earning potential in these regions.

4. **Top Areas by Number of Jobs:**
   - Islington: 4,066,000 jobs
   - Tower Hamlets: 1,393,000 jobs
   - Camden: 314,000 jobs
   - Southwark: 3,070,000 jobs
   - Lambeth: 2,694,000 jobs

These areas have the highest number of jobs, indicating significant employment opportunities.

5. **Moderate Areas by Number of Jobs:**
   - Hackney: 2,439,000 jobs
   - Hounslow: 3,225,000 jobs
   - Hillingdon: 4,910,000 jobs
   - Wandsworth: 2,095,000 jobs

- ○ Greenwich: 1,926,000 jobs

These areas have a moderate number of jobs, indicating good employment opportunities.

6. **Areas with Lower Number of Jobs:**
   - ○ Barnet: 452,000 jobs
   - ○ Havering: 403,000 jobs
   - ○ Bromley: 314,000 jobs

These areas have the lowest number of jobs, indicating fewer employment opportunities.

4.2.1.2 Detailed Observations:

Mean Salaries by Area:

1. **Top Salary Areas:**
   - ○ Islington has the highest mean salary (£34,366.40), making it the area with the highest earning potential.
   - ○ Tower Hamlets and Camden follow with mean salaries of £33,637.25 and £31,936.50, respectively, indicating high earning potential.
2. **Mid-Range Salary Areas:**
   - ○ Areas like Hackney, Hounslow, and Hillingdon have mean salaries between £28,000 and £31,000. These areas offer good earning potential but are slightly lower than the top salary areas.
3. **Lower Salary Areas:**
   - ○ Westminster (£24,447.00), Barnet (£24,317.67), and Havering (£23,801.33) have mean salaries below £25,000, indicating lower earning potential in these regions.

Number of Jobs by Area:

1. **Top Employment Areas:**
   - ○ Islington has the highest number of jobs (4,066,000), making it a significant employment hub.
   - ○ Hillingdon (4,910,000) and Hounslow (3,225,000) also have a high number of jobs, indicating significant employment opportunities.
2. **Mid-Range Employment Areas:**
   - ○ Areas like Southwark (3,070,000), Lambeth (2,694,000), and Hackney (2,439,000) have a substantial number of jobs, indicating good employment opportunities.
3. **Lower Employment Areas:**
   - ○ Barnet (452,000), Havering (403,000), and Bromley (314,000) have the lowest number of jobs, indicating fewer employment opportunities.

Comparative Analysis:

1. **High Salary and Employment Areas:**

- ○ Islington stands out with the highest mean salary and the highest number of jobs, making it a top area for both earning potential and employment opportunities.
  - ○ Tower Hamlets and Camden also rank high in both salary and employment, indicating strong job markets with good earning potential.
2. **High Salary but Lower Employment Areas:**
   - ○ Southwark and Lambeth have high mean salaries but do not rank as high in terms of employment opportunities compared to areas like Hillingdon and Hounslow.
3. **Lower Salary and Employment Areas:**
   - ○ Barnet, Havering, and Bromley rank lower in both mean salary and number of jobs, indicating areas with lower earning potential and fewer employment opportunities.
4. **Discrepancies:**
   - ○ Hillingdon has a high number of jobs (4,910,000) but a relatively moderate mean salary (£28,809.33), indicating a large job market with moderate earning potential.

Overall Trends: Areas like Islington, Tower Hamlets, and Camden are top regions for both earning potential and employment opportunities, making them attractive for job seekers and professionals.

Employment Hubs: Hillingdon and Hounslow offer substantial employment opportunities, although their mean salaries are not the highest.

Lower Potential Areas: Barnet, Havering, and Bromley rank lower in both salary and job numbers, indicating regions with fewer opportunities and lower earning potential.
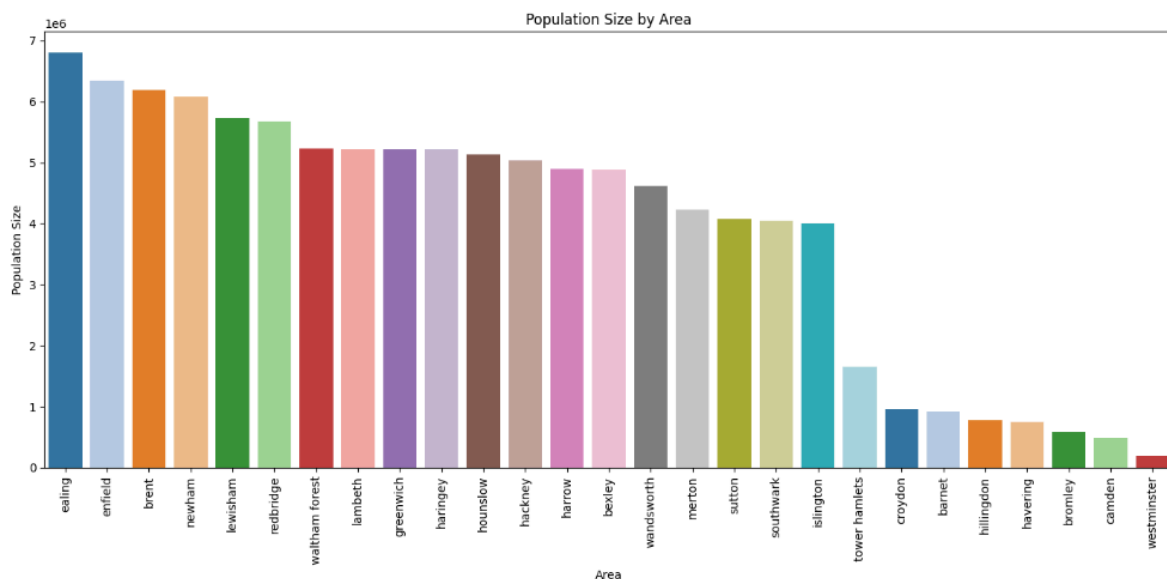
4.2.2 Population & Housing

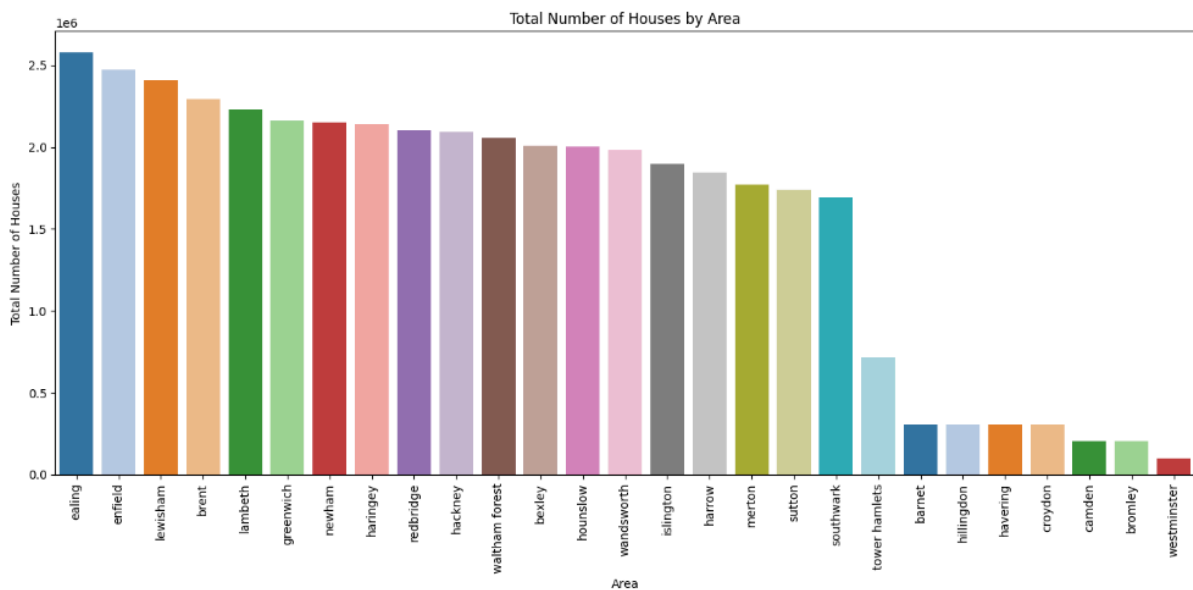

Image 15: Population Size by Area

Image 16: Total Number of Houses by Area

4.2.2.1 Key Observations:

1. **Top Areas by Number of Houses:**
   - Ealing: 2,579,029 houses
   - Enfield: 2,471,109 houses
   - Lewisham: 2,408,357 houses
   - Brent: 2,291,324 houses
   - Lambeth: 2,227,395 houses

These areas have the highest number of houses, indicating significant residential capacity.

2. **Moderate Areas by Number of Houses:**
   - Barnet: 307,206 houses
   - Hillingdon: 307,206 houses
   - Havering: 307,206 houses
   - Croydon: 307,206 houses
   - Camden: 204,804 houses

These areas have a substantial number of houses, showing they are important residential markets.

3. **Areas with Lower Numbers of Houses:**
   - Westminster: 102,402 houses
   - Bromley: 204,804 houses
   - Camden: 204,804 houses

These areas have the lowest number of houses, indicating they are less active residential markets.

4.2.2.2 Key Observations for Population Size:

1. **Top Areas by Population Size:**
   - Ealing: 6,806,329.5
   - Enfield: 6,337,991.5
   - Brent: 6,197,628.5
   - Newham: 6,076,996.5
   - Lewisham: 5,733,595.5

These areas have the highest population sizes, indicating they are highly populated regions.

2. **Moderate Areas by Population Size:**
   - Barnet: 923,856.5
   - Hillingdon: 785,567.5
   - Havering: 745,456.5
   - Bromley: 589,505.5
   - Camden: 484,606.5

These areas have moderate population sizes, showing they are significant but less populated regions compared to the top areas.

3. **Areas with Lower Population Sizes:**
   - Westminster: 189,233

These areas have the lowest population sizes, indicating they are less populated regions.

4.2.2.3 Comparative Analysis:

1. **High Residential Capacity and Population:**
   - Ealing stands out as both the area with the highest number of houses and the largest population size, making it a significant residential and populated area.
   - Enfield, Brent, and Lewisham also rank high in both categories, indicating strong residential capacity and high population.
2. **Discrepancies Between Houses and Population:**
   - Barnet and Hillingdon have a substantial number of houses (307,206 each) but lower population sizes (923,856.5 and 785,567.5, respectively). This could indicate larger house sizes or less dense populations.
   - Westminster has the lowest number of houses (102,402) and also one of the lowest population sizes (189,233), reflecting a smaller residential and population capacity.
3. **Areas with High Population but Moderate Housing:**
   - Newham and Lambeth have high population sizes (6,076,996.5 and 5,219,822.5, respectively) but rank slightly lower in the number of houses, indicating higher population density.

Overall Trends: Areas like Ealing, Enfield, and Brent are top regions for both housing and population, making them significant residential and populated zones.

Discrepancies: Barnet and Hillingdon show a notable discrepancy between the number of houses and population sizes, suggesting larger house sizes or less dense populations.

Lower Capacity Areas: Westminster has the lowest residential and population capacities, indicating it is less significant in both categories.
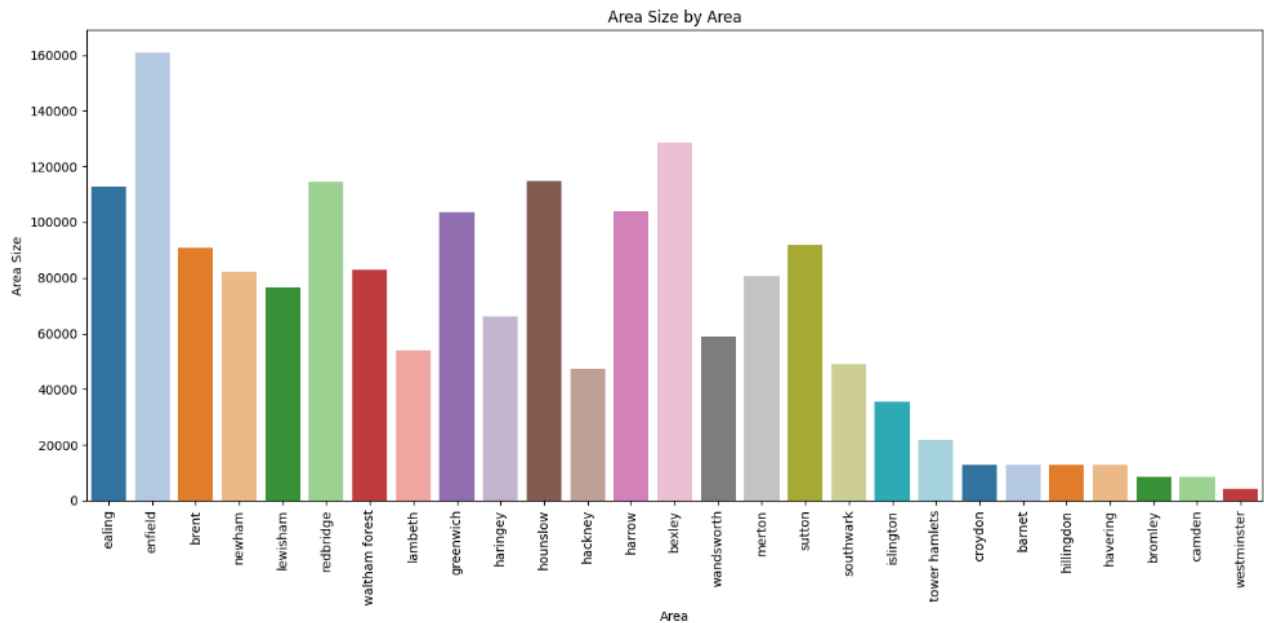
4.2.3 Area Size

Image 17: Area Size by Area

4.2.3.1 Key Observations:

1. **Top Areas by Size:**
   - Enfield: 160,929.0
   - Ealing: 112,941.0
   - Redbridge: 114,561.0
   - Hounslow: 114,831.0
   - Bexley: 128,691.0

These areas are the largest in size, indicating they cover significant geographical regions.

2. **Moderate Areas by Size:**
   - Brent: 98,783.0
   - Newham: 82,395.0
   - Lewisham: 76,545.0
   - Greenwich: 103,761.0
   - Haringey: 66,249.0

These areas have moderate sizes, showing they cover fair but not extensive geographical regions.

    3. **Areas with Smaller Sizes:**
- Westminster: 4,323.0
- Camden: 8,646.0
- Bromley: 86,460.0
- Havering: 12,969.0
- Barnet: 12,969.0

These areas are the smallest in size, indicating they cover limited geographical regions.

4.2.3.2 Detailed Observations:

    1. **Top Size Areas:**
- Enfield stands out with the largest area size (160,929.0), making it the region with the most extensive geographical coverage.
- Ealing and Redbridge follow with area sizes of 112,941.0 and 114,561.0, respectively, indicating significant geographical coverage.

    2. **Mid-Range Size Areas:**
- Areas like Brent, Newham, and Lewisham have area sizes ranging between 76,545.0 and 98,783.0. These areas cover fair but not extensive regions.
- The distribution in this mid-range is relatively close, indicating similar geographical coverage.

    3. **Smaller Size Areas:**
- Westminster has the smallest area size (4,323.0), making it the region with the least geographical coverage.
- Camden (8,646.0) and Bromley (86,460.0) also have smaller area sizes, indicating limited geographical regions.

The top areas like Enfield, Ealing, and Redbridge have the largest sizes, indicating extensive geographical coverage.

Mid-range areas like Brent, Newham, and Lewisham have moderate sizes, indicating fair geographical coverage.

Smaller areas like Westminster, Camden, and Bromley have the smallest sizes, indicating limited geographical coverage.
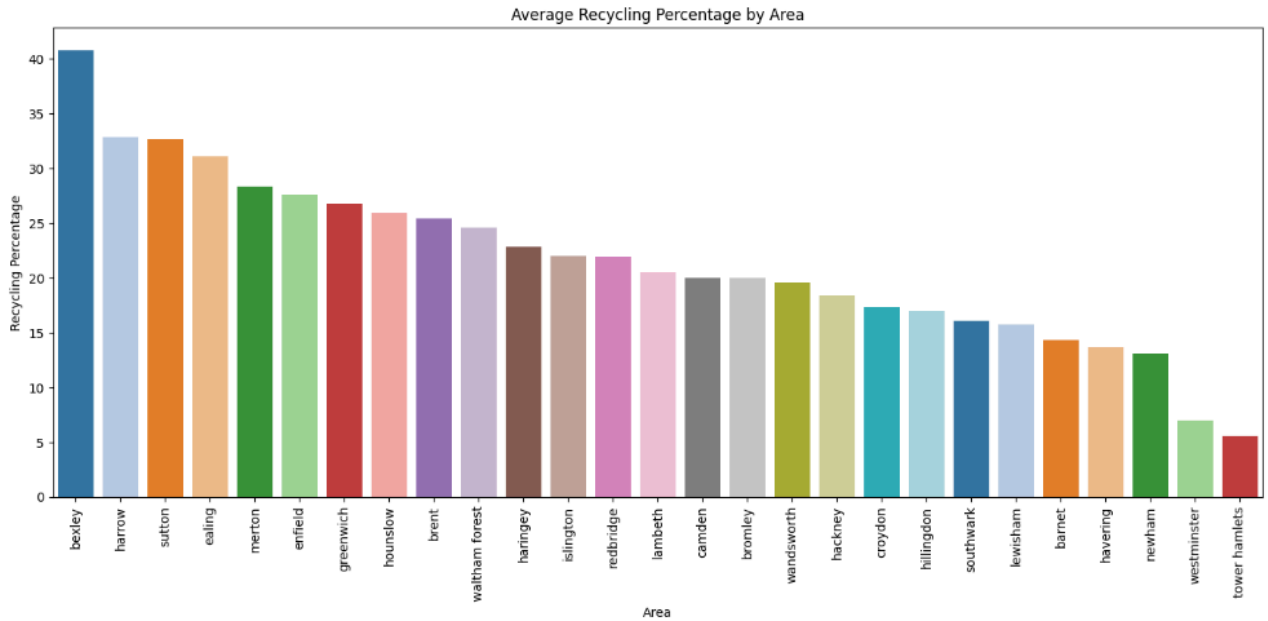
4.2.4 Environmental

Image 18: Average Recycling Percentage by Area

4.2.4.1 Key Observations:

1. **Top Areas by Recycling Percentage**:
   - **Bexley**: 40.81%
   - **Harrow**: 32.90%
   - **Sutton**: 32.67%
   - **Ealing**: 31.10%
   - **Merton**: 28.38%
   - These areas have the highest recycling percentages, indicating strong recycling programs and community participation.
2. **Moderate Areas by Recycling Percentage**:
   - **Enfield**: 27.62%
   - **Greenwich**: 26.76%
   - **Hounslow**: 25.95%
   - **Brent**: 25.43%
   - **Waltham Forest**: 24.62%
   - These areas have moderate recycling percentages, showing fair but not exceptional recycling rates.
3. **Areas with Lower Recycling Percentages**:
   - **Westminster**: 7.00%
   - **Newham**: 13.10%
   - **Havering**: 13.67%
   - **Barnet**: 14.33%
   - **Lewisham**: 15.76%
   - These areas have the lowest recycling percentages, indicating potential challenges in recycling programs or lower community engagement in recycling.

4.2.4.2 Detailed Observations:

1. **Top Recycling Areas**:
   - **Bexley** stands out with the highest recycling percentage (40.81%), making it the area with the most successful recycling initiatives.
   - **Harrow** and **Sutton** follow with recycling percentages of 32.90% and 32.67%, respectively, indicating strong recycling efforts.
2. **Mid-Range Recycling Areas**:
   - Areas like **Enfield**, **Greenwich**, and **Hounslow** have recycling percentages between 25% and 28%. These areas show good recycling rates but have room for improvement to reach the top tier.
3. **Lower Recycling Areas**:
   - **Westminster** (7.00%), **Newham** (13.10%), and **Havering** (13.67%) have the lowest recycling percentages, indicating they might face significant challenges in their recycling efforts or have less community engagement in recycling activities.

The top areas like Bexley, Harrow, and Sutton have the highest recycling percentages, indicating successful recycling programs and community engagement.

Mid-range areas like Enfield, Greenwich, and Hounslow show moderate recycling rates, suggesting fair but not exceptional recycling efforts.

Lower recycling areas like Westminster, Newham, and Havering have significantly lower recycling percentages, indicating potential challenges in recycling initiatives or lower community participation.

5. Comparative Analysis

5.1 Comparing Boroughs

1. **Ealing:**
   - Number of Houses: Highest with 2,579,029 houses.
   - Population Size: Highest with 6,806,329.5 people.
   - Recycling Percentage: High at 31.10%.
   - Area Size: Large at 112,941.0.

Purpose Relation: Ealing's large housing stock and population indicate high demand and supply. Its high recycling percentage enhances its attractiveness, potentially leading to higher prices.

2. **Enfield:**
   - Number of Houses: Second highest with 2,471,109 houses.
   - Population Size: Second highest with 6,337,991.5 people.
   - Recycling Percentage: Moderate at 27.62%.
   - Area Size: Largest at 160,929.0.

Purpose Relation: Enfield's large area and housing stock suggest availability, but its high population could drive demand and prices. Its moderate recycling rate indicates a balance between demand and environmental quality.

3. **Lewisham:**
    - Number of Houses: High at 2,408,357 houses.
    - Population Size: High at 5,733,595.5 people.
    - Recycling Percentage: Low at 15.76%.
    - Area Size: Moderate at 76,545.0.

Purpose Relation: Lewisham's high housing and population suggest strong demand, but its low recycling rate may affect attractiveness, impacting prices.

4. **Brent:**
    - Number of Houses: High at 2,291,324 houses.
    - Population Size: High at 6,197,628.5 people.
    - Recycling Percentage: Moderate at 25.43%.
    - Area Size: Moderate at 98,783.0.

Purpose Relation: Brent's high demand and moderate recycling efforts suggest potential for increasing prices, balanced by adequate housing supply.

5. **Westminster:**
    - Number of Houses: Lowest at 102,402 houses.
    - Population Size: Lowest at 189,233 people.
    - Recycling Percentage: Low at 7.00%.
    - Area Size: Smallest at 4,323.0.

Purpose Relation: Westminster's limited housing and population with poor recycling make it less attractive, but scarcity might drive higher prices.

Ealing and Enfield are standout boroughs with high residential capacity, large populations, and significant geographical coverage, suggesting potential for balanced prices due to availability and demand.

Lewisham and Brent have high demand and moderate recycling, indicating areas with potential for rising prices but room for improvement in environmental attractiveness.

Westminster is less significant in terms of residential capacity and attractiveness, but scarcity of housing might lead to higher prices.

6. Conclusion

- **Summary of Findings**: Summarise the main insights from the analysis.
- **Recommendations**: Provide any actionable recommendations based on the findings.
- **Future Work**: Suggest areas for further analysis or data collection.

6.1 Summary of Findings

Comparative Insights:

1. **Ealing:**
   ○ Strengths: High number of houses and population, good recycling practices, and large area size.
   ○ Challenges: Needs to maintain a balance between demand and supply to manage housing prices.
2. **Enfield:**
   ○ Strengths: Large housing stock and population, largest area size.
   ○ Challenges: Moderate recycling rate; managing high demand to prevent excessive price increases.
3. **Lewisham:**
   ○ Strengths: Significant residential capacity and high population.
   ○ Challenges: Low recycling rate; improving environmental practices to enhance attractiveness.
4. **Brent:**
   ○ Strengths: High demand with moderate recycling efforts.
   ○ Challenges: Balancing housing supply to manage prices effectively.
5. **Westminster:**
   ○ Strengths: Prime location but low residential capacity.
   ○ Challenges: Limited housing and population with poor recycling practices, potentially leading to high prices due to scarcity.

The analysis reveals that the dynamics of housing prices in London are influenced by a combination of factors including housing supply, population density, environmental practices, and geographical size. Understanding these metrics on both yearly and monthly bases provides valuable insights for predicting future housing costs and identifying the best and cheapest places to live. Boroughs like Ealing and Enfield stand out for their balance between housing supply and demand, while areas like Westminster may face challenges due to limited availability and lower environmental practices.

This comprehensive view aids buyers, sellers, policymakers, and urban planners in making informed decisions to navigate the complex housing market in London.

6.2 Future

6.2.1 Why Are House Prices Increasing?

1. **High Demand and Limited Supply:**
   ○ Population Growth: The graphs for population sizes indicate that boroughs like Ealing, Enfield, and Brent have large and growing populations. High population growth increases the demand for housing, driving up prices.
   ○ Limited New Developments: Boroughs with limited area sizes, such as Westminster and Camden, have restricted space for new housing developments, leading to higher prices due to limited supply.
2. **Economic Factors:**

- ○ Low-Interest Rates: Historically low mortgage rates have made borrowing cheaper, encouraging more people to buy homes and increasing demand.
- ○ Inflation: The costs of building materials and labour have increased due to inflation, which is passed on to homebuyers in the form of higher prices.

3. **Desirability and Amenities:**
   - ○ Environmental Practices: Boroughs with higher recycling rates, such as Bexley and Harrow, are perceived as more desirable places to live, increasing demand and prices.
   - ○ Safety and Crime Rates: Boroughs with lower crime rates are more attractive, further driving up housing prices. For example, boroughs like Camden and Westminster, despite having lower numbers of houses, have high prices due to their desirability and amenities.

4. **Investment Potential:**
   - ○ Real Estate Investment: Properties in London are seen as safe investments, attracting both domestic and international buyers, thus increasing demand and prices.

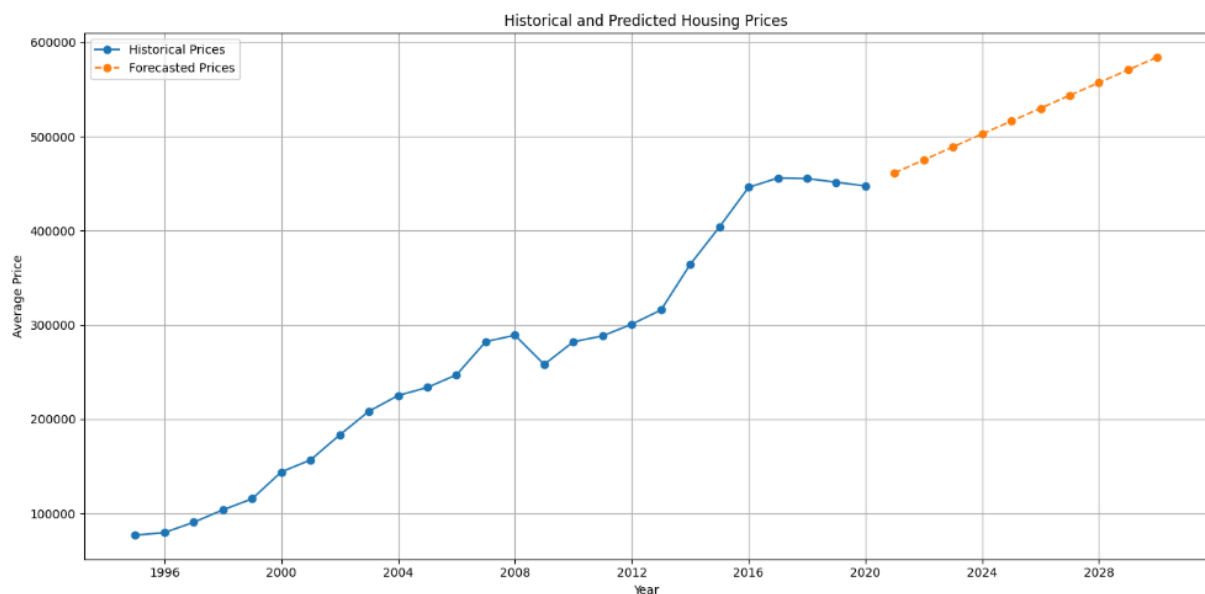### 6.2.2 Will House Prices Continue to Increase?

Image 19: Predicted Housing Prices

Predicted Trends (2021 - 2030):

- **Steady Increase**: The forecasted prices from 2021 to 2030 indicate a steady and continuous increase. This suggests that the underlying factors driving housing prices upwards, such as high demand, limited supply, and economic conditions, are expected to persist.
- **Rate of Increase**: The slope of the forecasted trend line is similar to the historical trend pre-2008, indicating robust growth expectations.

Factors Influencing Predictions:

1. **Sustained Demand**: Continued population growth and urbanisation in London are likely to keep demand for housing high.
2. **Economic Conditions**: Assuming stable economic growth and low-interest rates, the conditions are favourable for continued price increases.
3. **Supply Constraints**: Limited availability of land and regulatory constraints on new developments contribute to a persistent supply-demand imbalance.
4. **Desirability**: London's status as a global city with high-quality amenities and services makes it an attractive place to live and invest in, further driving up prices.

Considerations for Future Predictions:

1. **Economic Shocks**: Unexpected economic downturns or shocks could alter the forecasted trend, as seen in the 2008 financial crisis.
2. **Policy Changes**: Government policies aimed at increasing housing supply or changing interest rates could moderate price increases.
3. **Global Factors**: International factors, such as foreign investment flows and geopolitical events, could also impact the housing market.

The graph indicates a strong upward trend in housing prices, both historically and in future predictions. The key drivers of this trend include sustained demand, economic conditions, supply constraints, and the desirability of London as a place to live and invest. While this suggests potential for appreciation, it also underscores challenges related to affordability. Future predictions remain contingent on various economic, policy, and global factors that could influence the market dynamics.

6.3 Recommendations

Best and Cheapest Areas to Buy a House

1. **Cheapest Areas:**
   ○ Westminster: Despite its desirability, it has a lower number of houses, leading to higher prices. However, areas like:
   ○ Bromley: Offers lower prices with moderate population and area size.
   ○ Havering: Has lower house prices and decent environmental practices.
2. **Best Value Areas:**
   ○ Ealing: With a high number of houses, population, good recycling practices, and large area size, Ealing offers a balanced choice.
   ○ Enfield: Its large housing stock and population, coupled with the largest area size, suggest good potential for appreciation.
   ○ Lewisham: High residential capacity and population size make it a potential growth area despite its low recycling rate.

Implications for Buyers and Investors:

1. **Potential for Appreciation**: The predicted steady increase in housing prices suggests good potential for property value appreciation, making it a favourable time for long-term investment.

2. **Market Timing**: Given the continuous upward trend, earlier investment could yield better returns compared to waiting for a potential dip.
3. **Affordability**: With prices expected to rise, affordability may become a more significant issue, particularly for first-time buyers.