Real Estate Market Analysis: Central vs. Suburban London

Author: Sharon Lin

Executive Summary

Market research companies assist real estate agents in better understanding the market by collecting real estate data, analyzing it, and generating visual charts to identify key factors and variables influencing housing prices.

This study uses real estate in central London as a sample, analyzing 100 data points to determine that house prices in this area are primarily influenced by property type, house size, number of bathrooms, and sports facilities. In contrast, houses in suburban London are affected not only by property type, number of bedrooms, and size but also by distance from the nearest station.

This research provides valuable insights for real estate agents, enabling them to better understand homebuyers' needs in these two areas and develop tailored operational and sales strategies.

Table of contents

1. Introduction

The W1 postcode is in central London. This area is famous for its luxury residences and convenient facilities. It's one of the most expensive areas of the United Kingdom (Kettle, 2024). The SE25 postcode is in south London. It is mainly residential areas surrounded by markets, parks, and schools. It is one of the cheapest areas in London (Reeves, 2023).

This report uses real estate in the W1 and SE25 regions as a sample and analyzes the following variables: property type, number of bedrooms, number of bathrooms, size, tenure, parking space, sports facilities, distance to the nearest station.

2. Data collection

I used (Rightmove, 2024) to collect data, with the keyword "W1 London", and set it to only display detached, semi-detached, terraced, flat, and bungalow room types, and then collected samples according to price from high to low. A total of 100 data were collected. The same methodology was used to collect data in the SE25 region. In all, 80 items of data were collected.

The reason I chose this approach for data collection is to compare the highest average house prices between the two regions and ensure that both samples consist of the same type of house.

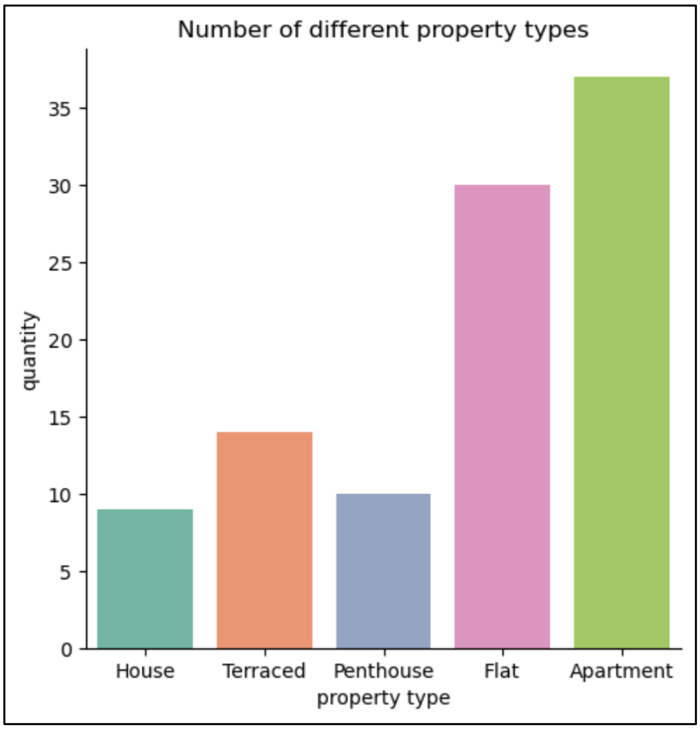
3. Descriptive Statistics (Overall analysis)

	Number_of_bedrooms	Number_of_bathrooms	Size_sqft	Distance from station (miles)	House_price_GBPK
count	100	100	100	100	100
mean	3.8	3.58	3260.46	0.232	10235.294
std	1.95	1.71	3102.85	0.08	8044.99
min	1	1	890	0.1	1100
25%	3	3	1507	0.2	6187.5
50%	3	3	2286	0.2	7950
75%	5	4	3827.25	0.3	11062.5
max	12	10	20987	0.4	65000

According to this descriptive statistics report, the average size of the houses in W1 is 3260.46 square feet, and each house has roughly three bedrooms and three baths. The closest station is typically 0.2 miles away. The average cost comes to about £10,235,000.

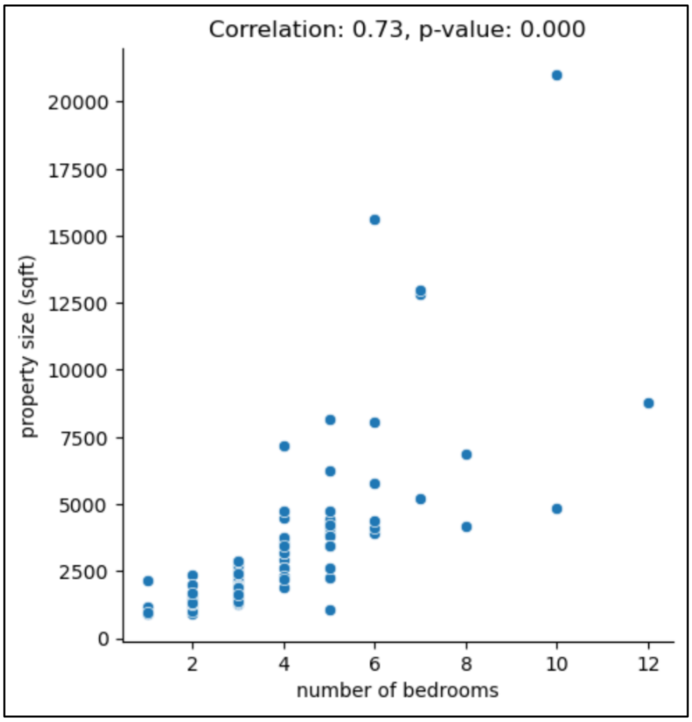
4. Data visualization and Inferential statistics

I. Number of different types of properties



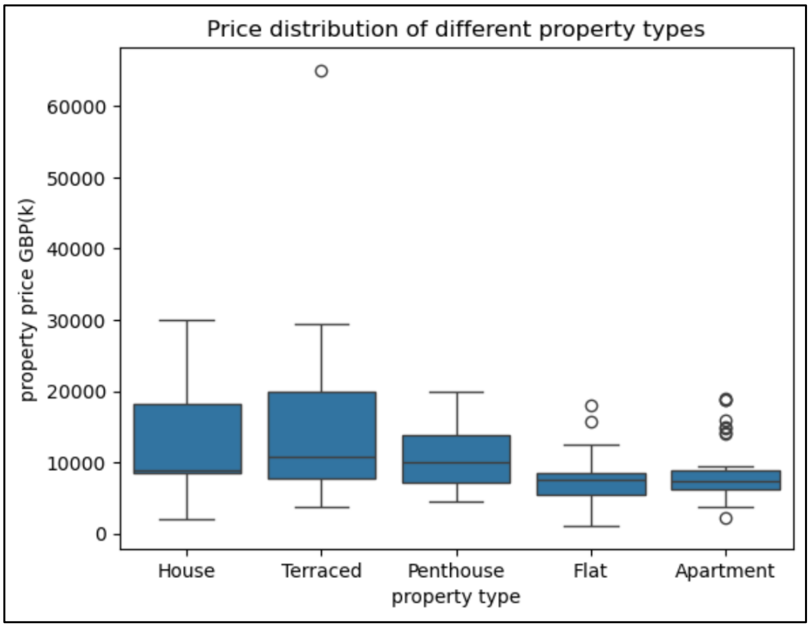
- Insights: As the above figure illustrates, the number of samples of Apartment and Flat far exceeds that of other types. When the sample of house types is unevenly distributed, types with a larger number will have a greater impact on the results of the model, while types with a smaller number may be ignored or not represented correctly in the model.
- Theory: This chart complies with Graphical Integrity, it has a clear title, x-axis, y-axis, and appropriate proportions between graphs to correctly reflect the data. It also conforms to Graphical Excellence with minimal ink application and no excessive decoration. Moreover, It also complies with the requirements in the Munzner hierarchy, using different colors to distinguish different property types, making it clear at a glance.

II. Correlation between number of bedrooms and property size



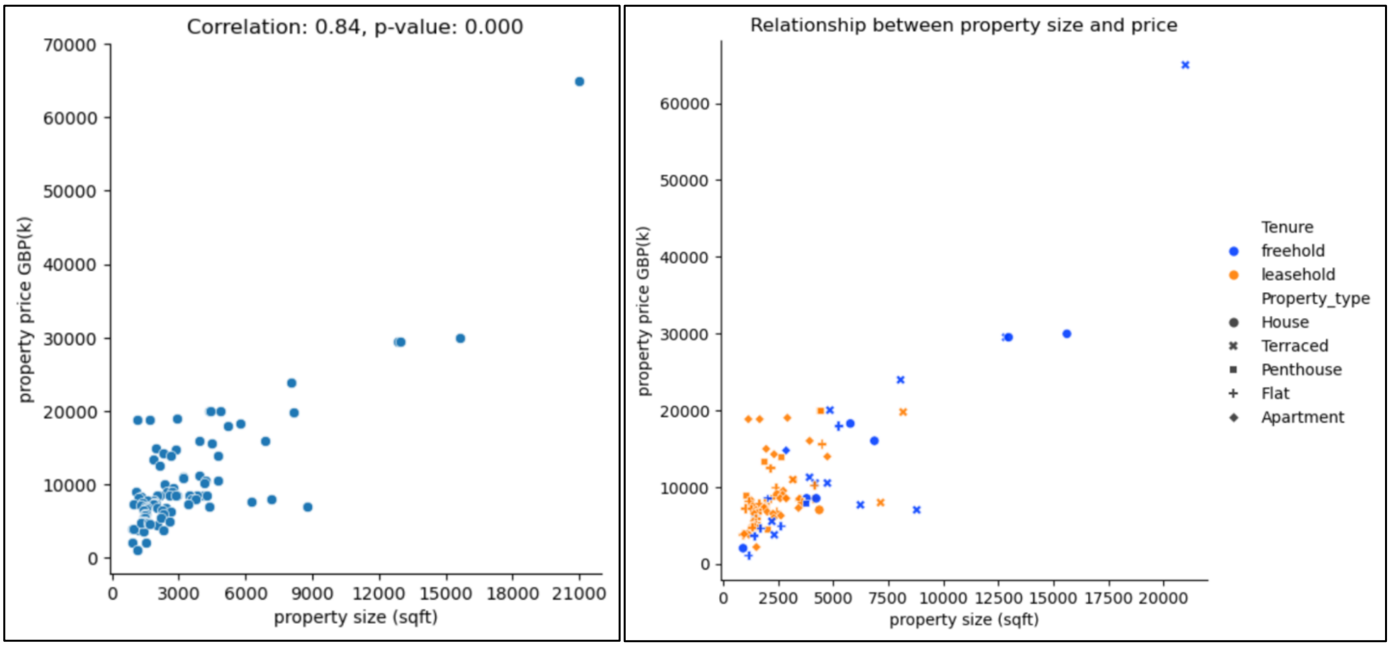
- Insights: It can be seen from the chart that the two variables are highly positively correlated and reach a statistically significant level.
- Theory: In addition to complying with Graphical Integrity and Graphical Excellence, this chart also complies with the MECE principle, showing all data on the number of bedrooms and property size, mutually exclusive and fully covered, so that the audience can fully understand the relationship between these two variables.

III. Price distribution of different types of properties



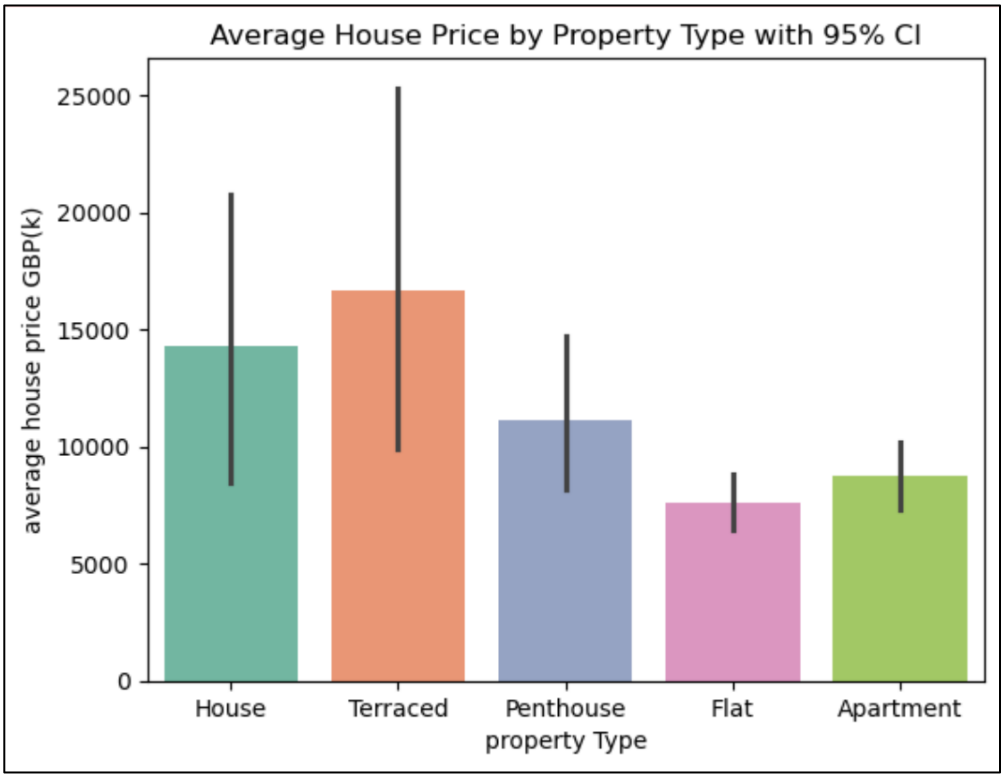
From this chart, we can see that Terraced's prices are relatively high and widely dispersed, with extreme values, while Flat and Apartment's prices are lower, more concentrated, and have extreme values.

IV. Correlation between property size and price



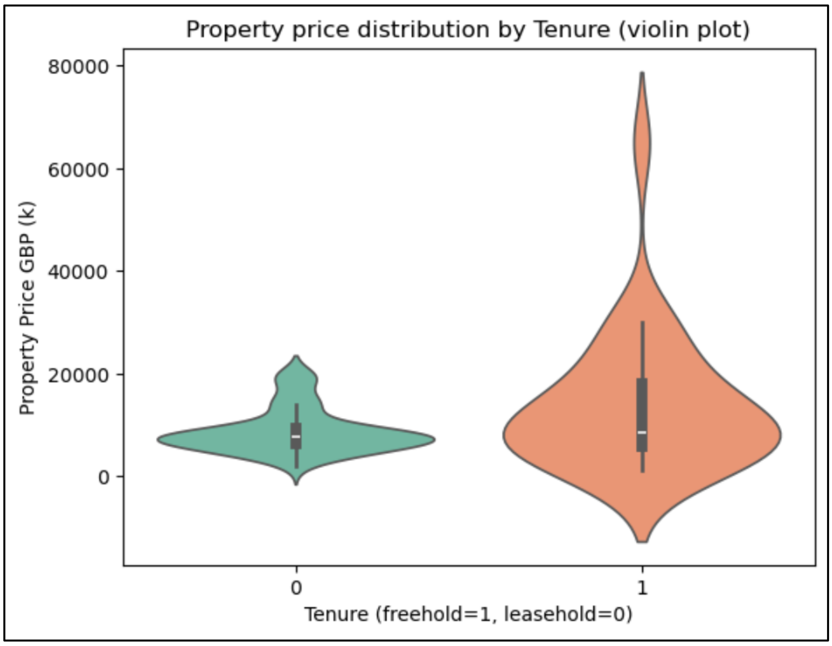
- Insights: From this figure, we can see that the two variables are highly positively correlated. Properties belonging to the freehold type are generally larger in size and more likely to have higher prices. Most freehold properties are Houses and Terraced homes, whereas Apartments, Flats, and Penthouses are mostly leasehold.
- Theory: This chart illustrates the Gestalt principle of similarity by using color and shape to distinguish different categories, making it easier for viewers to differentiate between various data groups. It also adheres to the principle of Graphical Excellence by presenting data comprehensively with multiple variables.

V. Average house price by property type with 95% CI



Terraced has the highest average house price, while Flat and Apartment have the lowest. The confidence interval of Terraced is very wide, which means that there are large differences between its prices. The confidence intervals of Flat and Apartment are relatively narrow, it means they have relatively stable prices. It may also be because the larger the sample size, the smaller the confidence interval.

VI. **Correlation between Tenure and property price**



Point Biserial Correlation: 0.28 P-value:0.004

Group Statistics

	Tenure	N	Mean	Std. Deviation	Std. Error Mean
House_price_GBPK	1	27	13957.407	13281.5495	2556.0354
	0	73	8858.622	4263.3250	498.9845

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
House_price_GBPK	Equal variances assumed	25.600	<.001	2.919	98	.002	.004	5098.7855	1746.9788	1631.9630	8565.6080
	Equal variances not assumed			1.958	28.005	.030	.060	5098.7855	2604.2854	-235.8101	10433.3811

Two variables are slightly positively correlated, and there is a significant difference in the average house prices of freehold and leasedhold.

VII. Comparison of average price for different properties (One-Way ANOVA)

ANOVA

num_property

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	154.898	62	2.498	2.030	.011
Within Groups	45.542	37	1.231		
Total	200.440	99			

Tests of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
num_property	Based on Mean	4.765	21	37	<.001
	Based on Median	2.919	21	37	.002
	Based on Median and with adjusted df	2.919	21	11.469	.033
	Based on trimmed mean	4.695	21	37	<.001

One-Way ANOVA result shows that the average price of at least one property type is different from the others, but Levene's test result shows that the p-value is <0.05, indicating that the house price variations of the five house types are not equal. The ANOVA assumption is violated, so it is impossible to conclude whether the average prices of the five properties are different.

5. Linear regression

```
                            OLS Regression Results
==============================================================================
Dep. Variable:       House_price_GBPK   R-squared:                       0.824
Model:                            OLS   Adj. R-squared:                  0.802
Method:                 Least Squares   F-statistic:                     37.46
Date:                Wed, 04 Dec 2024   Prob (F-statistic):           1.68e-28
Time:                        19:20:27   Log-Likelihood:                -953.80
No. Observations:                 100   AIC:                             1932.
Df Residuals:                      88   BIC:                             1963.
Df Model:                          11
Covariance Type:            nonrobust
=================================================================================================
                                    coef    std err          t      P>|t|      [0.025      0.975]
-------------------------------------------------------------------------------------------------
Intercept                       -41.3531   2000.627     -0.021      0.984   -4017.179    3934.473
C(Property_type)[T.Flat]      -1090.8006    918.349     -1.188      0.238   -2915.826     734.225
C(Property_type)[T.House]     -5061.8764   1660.584     -3.048      0.003   -8361.939   -1761.814
C(Property_type)[T.Penthouse]  2044.0159   1340.685      1.525      0.131    -620.314    4708.346
C(Property_type)[T.Terraced]  -5023.3314   1569.253     -3.201      0.002   -8141.893   -1904.770
Tenure[T.leasehold]              84.0048   1099.386      0.076      0.939   -2100.794    2268.804
Parking[T.yes]                  -83.0740    844.737     -0.098      0.922   -1761.811    1595.663
Sports_facilities[T.yes]       4164.6927    944.289      4.410      0.000    2288.116    6041.269
Number_of_bedrooms             -301.6068    342.595     -0.880      0.381    -982.442     379.228
Number_of_bathrooms            1195.3305    360.773      3.313      0.001     478.369    1912.292
Size_sqft                         2.1475      0.193     11.108      0.000       1.763       2.532
Distance_from_station_miles    1012.5859   4871.279      0.208      0.836   -8668.057    1.07e+04
==============================================================================
Omnibus:                        9.156   Durbin-Watson:                   1.479
Prob(Omnibus):                  0.010   Jarque-Bera (JB):               10.014
Skew:                           0.547   Prob(JB):                      0.00669
Kurtosis:                       4.098   Cond. No.                     6.29e+04
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 6.29e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```

I. **Adj. R-squared (0.802)**: R-squared is an indicator to measure the performance of the regression model. In this case, the adjusted R-squared is strong (80.2%), suggesting that these factors are meaningful in predicting house prices.

II. **F-statistic (37.46)**: The p-value is much less than 0.05, indicating that this regression model is statistically significant overall.

III. **Model**: Establish a regression model by statistically significant variables.

House_price_GBPK=−41.3531−5061.8764*House−5023.3314*Terraced+4164.6927*Sports_facilit

ies+1195.3305*Number_of_bathrooms+2.1475*Size_sqft

IV. **Optimization model**: We can further optimize the model by adjusting the variable combination, such as removing insignificant variables and observing the change in adjusted R-squared.

6. Residual analysis:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:        House_price_GBPK   R-squared:                       0.822
Model:                             OLS   Adj. R-squared:                  0.809
Method:                  Least Squares   F-statistic:                     60.83
Date:                 Wed, 04 Dec 2024   Prob (F-statistic):           8.93e-32
Time:                         19:35:25   Log-Likelihood:                -954.28
No. Observations:                  100   AIC:                             1925.
Df Residuals:                       92   BIC:                             1945.
Df Model:                            7
Covariance Type:             nonrobust
===================================================================================================
                                    coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------------------------
Intercept                      -121.8509   1020.641     -0.119      0.905   -2148.931    1905.230
C(Property_type)[T.Flat]      -1159.5775    869.645     -1.333      0.186   -2886.768     567.613
C(Property_type)[T.House]     -5183.9345   1471.160     -3.524      0.001   -8105.785   -2262.084
C(Property_type)[T.Penthouse]  1852.6755   1264.775      1.465      0.146    -659.277    4364.627
C(Property_type)[T.Terraced]  -5414.7928   1343.535     -4.030      0.000   -8083.169   -2746.417
Sports_facilities[T.yes]       4342.2068    805.868      5.388      0.000    2741.683    5942.731
Number_of_bathrooms            1047.0567    310.951      3.367      0.001     429.480    1664.633
Size_sqft                         2.0795      0.174     11.934      0.000       1.733       2.426
==============================================================================
Omnibus:                       10.511   Durbin-Watson:                   1.494
Prob(Omnibus):                  0.005   Jarque-Bera (JB):               12.830
Skew:                           0.561   Prob(JB):                      0.00164
Kurtosis:                       4.350   Cond. No.                     2.28e+04
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 2.28e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```

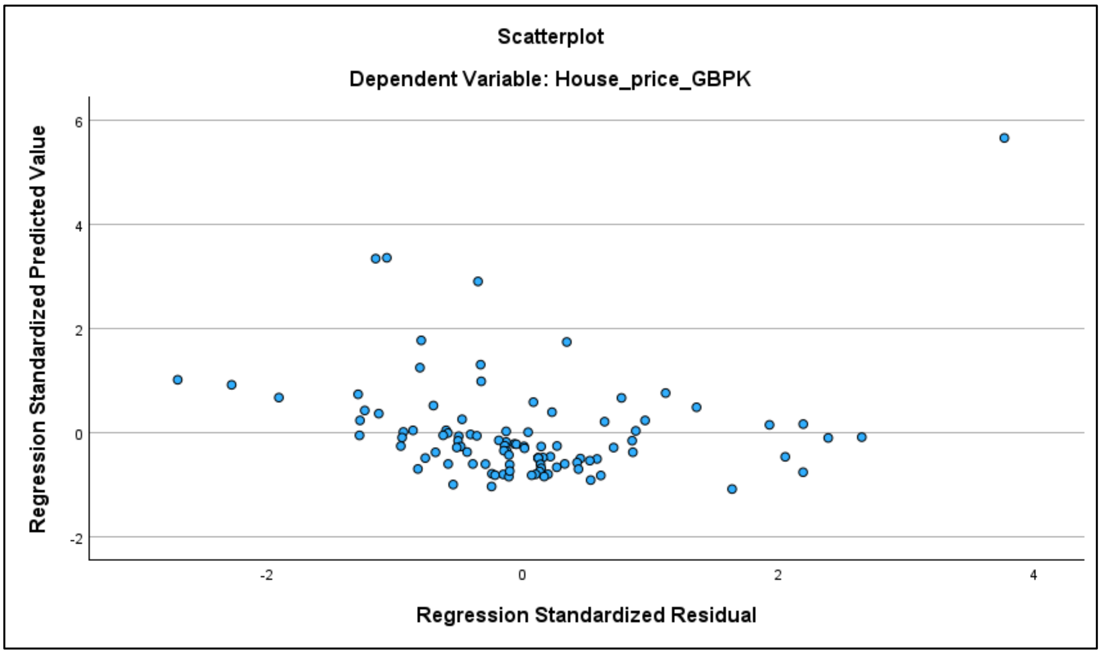
I. **Optimized model**:

House_price_GBPK=−121.8509−5183.9345*House−5414.7928*Terraced+4342.2068*Sports_facilities+1047.0567*Number_of_bathrooms+2.0795*Size_sqft

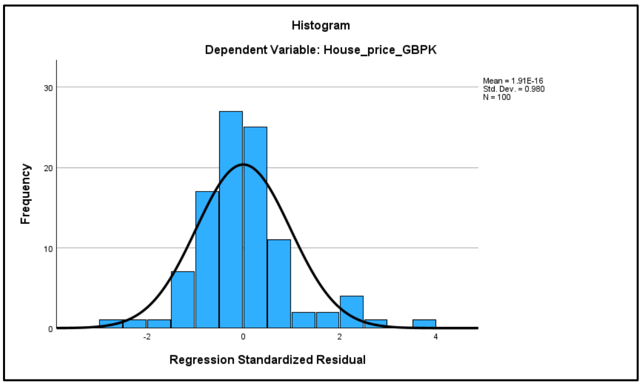
II. **Adj. R-squared (0.809)**: Adj. R-squared increased from 0.802 to 0.809, indicating that the improved model is better at explaining house price changes.

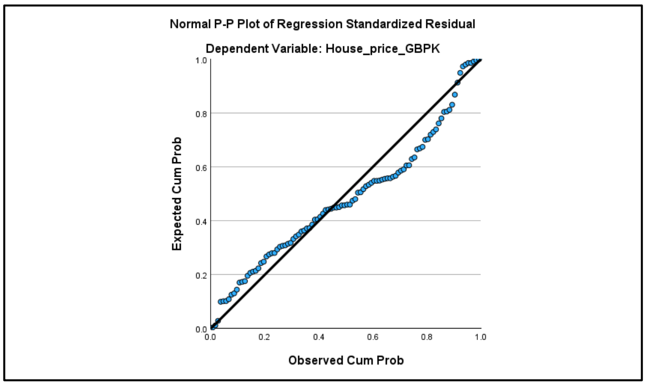
III. **The five assumptions of linear regression-Residual analysis**

- Linearity: As the below figure shows, most points are randomly distributed around zero, which roughly conforms to the linear relationship assumption. However, several extreme values may still affect the results, indicating that the model does not fully explain some data.
- Independence: The Durbin-Watson value is close to 2 means the residuals are independent. The 1.494 here indicates some correlation issues.
- Homoscedasticity: As the chart below illustrates, most points are randomly distributed near zero, but the overall distribution is uneven and has extreme values, which does not fully meet this assumption.



- Normality: Jarque-Bera test result shows that p-value<0.05, indicating that the residual does not meet the normality assumption. As shown in the figure below, the residual is close to the normal distribution, but there is still a slight deviation.





- No multicollinearity: Condition Number 2.28e+04 is large, which means that this model may have multicollinearity issues.

IV. **How to solve the problem**: Based on the above analysis results, this model should remove extreme values and variables that are highly correlated with each other to improve the stability of the model and increase its predictive ability.

7. Regression model usage example

I. Example 1:

Property type: House, with sport facilities, 4 bathrooms, 1400 sq ft.

House_price_GBPK=−121.8509−5183.9345*1−5414.7928*0+4342.2068*1+1047.0567*4+2.0795*1400= 6135.95(k)=£6,135,950

II. Example 2

Property type: Terraced, without sport facilities, 2 bathrooms, 3000 sq ft.

House_price_GBPK=−121.8509−5183.9345*0−5414.7928*1+4342.2068*0+1047.0567*2+2.0795*3000= 2795.97(k)=£2,795,970

8. Compare with other areas of London

I. Average price comparison:

	cat	N	Mean	Std. Deviation	Std. Error Mean
Group Statistics					
House_price	1	80	445.09994	201.341679	22.510684
	2	100	10235.29400	8044.992915	804.499291

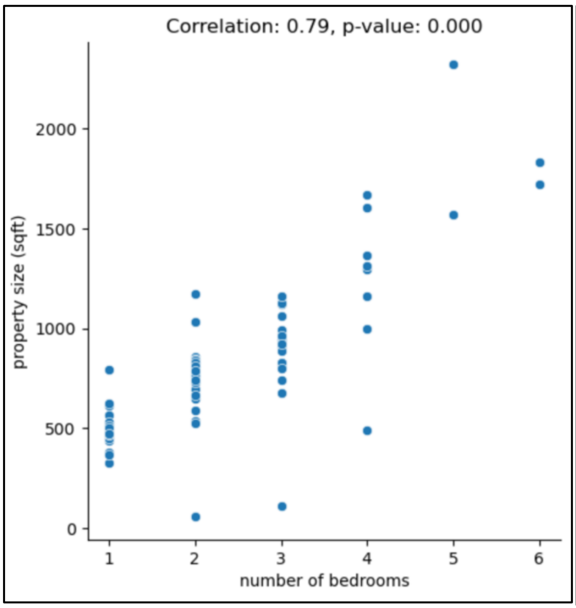
Independent Samples Test

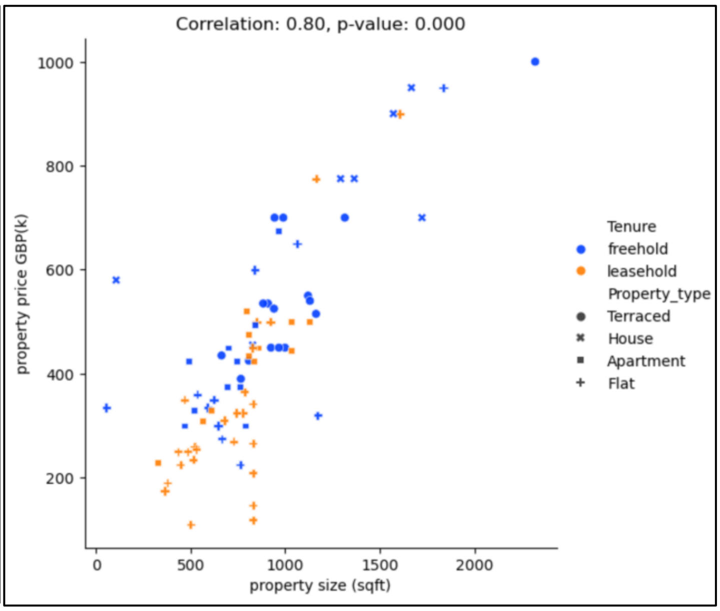
		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
House_price	Equal variances assumed	49.637	<.001	-10.876	178	<.001	<.001	-9790.194063	900.187602	-11566.607072	-8013.781053
	Equal variances not assumed			-12.165	99.155	<.001	<.001	-9790.194063	804.814165	-11387.089090	-8193.299035

According to the above figure, there is a significant difference in the average house price between W1 and SE25 areas.

The following charts are created using data from the SE25 area.

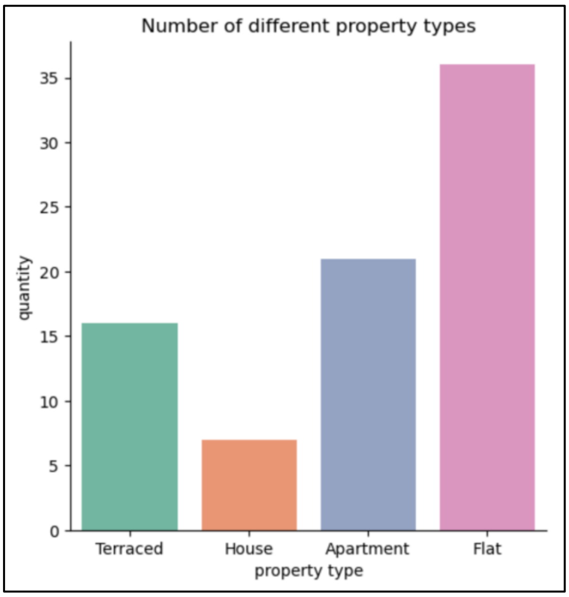
II. Similarities: Property size in SE25 areas is also highly positively correlated with the number of bedrooms and prices, and the distribution in Tenure is like that in the W1 area.

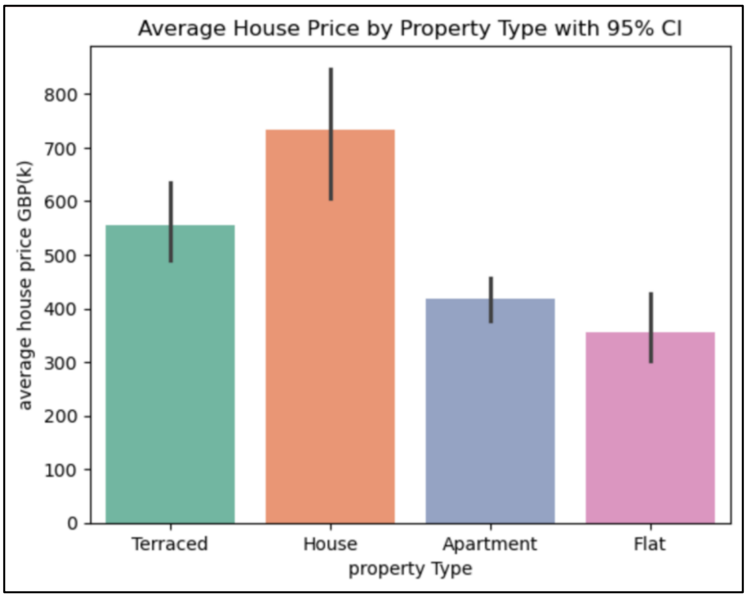




III. Differences: No Penthouse samples were collected in the SE25 area. The regression model is highly significant, indicating it effectively explains housing prices. Significant variables include Property type (Flat), Number of bedrooms, Size in square feet, and Distance to the nearest station, suggesting these factors help predict housing prices in this area. While the model generally meets statistical assumptions, multicollinearity still needs improvement.

	Number_of_bedrooms	Number_of_bathrooms	Size_sqft	station_distance_miles	House_price_GBPK
count	80	80	80	80	80
mean	2.48	1.30	847.55	0.35	445.10
std	1.15	0.62	374.25	0.19	201.34
min	1	1	55	0.1	110
25%	2	1	619	0.2	307.5
50%	2	1	817.5	0.3	425
75%	3	1.25	972.75	0.4	527.5
max	6	5	2321	0.9	1000





```
                          OLS Regression Results
==============================================================================
Dep. Variable:       House_price_GBPK   R-squared:                       0.776
Model:                            OLS   Adj. R-squared:                  0.747
Method:                 Least Squares   F-statistic:                     26.88
Date:                Fri, 06 Dec 2024   Prob (F-statistic):           2.22e-19
Time:                        22:39:08   Log-Likelihood:                -477.64
No. Observations:                  80   AIC:                             975.3
Df Residuals:                      70   BIC:                             999.1
Df Model:                           9
Covariance Type:            nonrobust
================================================================================================
                                  coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------------------------
Intercept                      93.2889     42.603      2.190      0.032       8.319     178.259
C(Property_type)[T.Flat]      -73.1412     30.001     -2.438      0.017    -132.975     -13.307
C(Property_type)[T.House]      53.3621     54.075      0.987      0.327     -54.487     161.211
C(Property_type)[T.Terraced]  -23.9121     39.883     -0.600      0.551    -103.457      55.632
Tenure[T.leasehold]           -39.0437     29.742     -1.313      0.194     -98.361      20.274
Parking[T.yes]                 45.2739     25.966      1.744      0.086      -6.513      97.061
Number_of_bedrooms             39.7106     18.182      2.184      0.032       3.448      75.973
Number_of_bathrooms            -5.3373     26.572     -0.201      0.841     -58.334      47.660
Size_sqft                       0.2733      0.055      5.001      0.000       0.164       0.382
station_distance_miles        170.8492     65.358      2.614      0.011      40.496     301.202
==============================================================================
Omnibus:                        1.337   Durbin-Watson:                   1.565
Prob(Omnibus):                  0.512   Jarque-Bera (JB):                0.775
Skew:                          -0.188   Prob(JB):                        0.679
Kurtosis:                       3.303   Cond. No.                     5.62e+03
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 5.62e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
```

9. Problems and limitations

I. The number of samples collected is too small, resulting in insufficient representativeness, and the prediction model may not be truly applicable to these areas.

II. Some variables may be insignificant in the regression model because they are highly correlated with other variables. For example, "property size" and "number of bedrooms" may have collinearity, causing "number of bedrooms" to be insignificant, but this does not mean that it has no impact on the model

III. Although significance is an important factor for variable selection, significance cannot be the only criterion because the external environmental background must also be considered. For example, although the correlation between "Distance to station" and “house prices” does not reach statistical significance, the reason may be that the city center of London is very convenient, and almost all houses are close to the station.

IV. Statistical assumptions need to be considered, and conclusions should not be drawn just because the results are statistically significant. This part cannot be presented due to a lack of space and statistical knowledge.

10. Conclusion and recommendations

According to the findings, in urban London, the size and type of houses, the number of bathrooms, and whether the house is equipped with sports facilities are important indicators for predicting housing prices. This could be attributed to the high socioeconomic position of residents in this neighborhood, who usually pursue larger homes, living space, and quality (ONS, 2024). And they may also be under greater work pressure, so they may be more care about sports. Furthermore, residents in the SE25 area might be more attentive to the proximity of the nearest station. This may be because they need to take public transport to work in the city center. When developing a prediction model, real estate companies must undertake more in-depth research on the characteristics of residents to choose the most appropriate predictor.

11. Reference list

Kettle, E. (2024). *London’s 20 cheapest postcodes to buy a home revealed*. [online] Ilford Recorder. Available at: https://www.ilfordrecorder.co.uk/news/24294738.londons-20-cheapest-postcodes-buy-home-revealed/ [Accessed 4 Dec. 2024].

online.stat.psu.edu. (n.d.). *4.1 - Residuals | STAT 462*. [online] Available at:

https://online.stat.psu.edu/stat462/node/116/.

ONS (2024). *Office for National Statistics*. [online] Ons.gov.uk. Available at: https://www.ons.gov.uk/.

Reeves, B. (2023). *The nation's priciest property postcodes revealed where homes sell for £8m on average*. [online] Benham and Reeves. Available at: https://www.benhams.com/press-release/london-property-market/the-nations-priciest-property-postcodes-revealed/ [Accessed 4 Dec. 2024].

Rightmove (2024). *Rightmove.co.uk*. [online] Rightmove.co.uk. Available at: https://www.rightmove.co.uk/.

12. Appendix

London W1 data, Collection date: December 2, 2024

Id	Property_type	Number_of_bedrooms	Number_of_bathrooms	Size_sqft	Tenure	Parking	Sports_facilities	Distance_from_station_miles	House_price_GBPK
1	House	2	2	912	freehold	yes	no	0.2	2000
2	Terraced	10	10	20987	freehold	no	yes	0.3	65000
3	Penthouse	3	4	2002	leasehold	yes	no	0.2	4500
4	House	6	6	15608	freehold	yes	no	0.2	29950
5	Terraced	7	5	12831	freehold	yes	yes	0.2	29500
6	House	7	7	12960	freehold	yes	yes	0.2	29500
7	Flat	1	1	890	leasehold	yes	yes	0.3	3860
8	Terraced	6	6	8060	freehold	no	yes	0.2	23950
9	Terraced	10	10	4852	freehold	yes	no	0.2	20000
10	Penthouse	5	3	4351	leasehold	no	no	0.1	19950
11	Penthouse	5	5	4433	leasehold	no	no	0.2	19950
12	Terraced	5	6	8174	leasehold	no	yes	0.2	19750
13	Apartment	4	4	2923	leasehold	yes	yes	0.1	19000
14	Apartment	3	4	1663	leasehold	yes	yes	0.2	18823.2
15	Apartment	2	3	1151	leasehold	yes	yes	0.2	18823.2
16	House	6	6	5780	freehold	yes	yes	0.3	18250
17	Apartment	2	2	1125	leasehold	no	no	0.3	4000
18	Apartment	3	2	1515	leasehold	yes	no	0.3	2150
19	Flat	7	6	5231	freehold	no	no	0.2	18000
20	Apartment	6	6	3931	leasehold	yes	yes	0.4	16000
21	House	8	8	6873	freehold	no	no	0.3	16000
22	Apartment	1	2	1175	leasehold	yes	no	0.3	3750

23	Flat	4	4	4469	leasehold	no	no	0.3	15650
24	Apartment	3	3	1959	leasehold	no	no	0.3	14950
25	Apartment	3	5	2865	freehold	yes	yes	0.3	14750
26	Apartment	3	3	2317	leasehold	yes	yes	0.1	14250
27	Apartment	4	4	4735	leasehold	yes	no	0.1	13950
28	Penthouse	3	3	2641	leasehold	yes	yes	0.3	13950
29	Penthouse	3	3	1835	leasehold	no	no	0.3	13395
30	Flat	1	1	2135	leasehold	yes	yes	0.3	12500
31	Terraced	5	5	3927	freehold	yes	yes	0.3	11250
32	Penthouse	4	4	3176	leasehold	yes	no	0.1	11000
33	Terraced	4	4	3169	leasehold	yes	no	0.3	10950
34	Terraced	8	5	4180	freehold	no	no	0.3	10500
35	Terraced	5	5	4736	freehold	no	yes	0.2	10500
36	Flat	5	5	4123	leasehold	no	no	0.3	10250
37	Flat	2	2	2379	leasehold	yes	yes	0.1	9950
38	Apartment	3	3	2737	leasehold	no	no	0.2	9500
39	House	3	3	2461	leasehold	no	no	0.2	8950
40	Penthouse	5	2	1055	leasehold	no	no	0.1	8950
41	Apartment	3	3	2700	leasehold	no	no	0.2	8950
42	Flat	3	3	2163	leasehold	yes	no	0.2	8500
43	Apartment	4	4	2595	leasehold	no	no	0.2	8500
44	Flat	2	1	1152	freehold	no	no	0.1	1100
45	Flat	3	3	2012	freehold	yes	no	0.3	8500
46	Apartment	3	3	2891	leasehold	yes	no	0.3	8500
47	Flat	6	6	4106	leasehold	yes	no	0.3	8500
48	House	5	3	3794	freehold	no	no	0.3	8500
49	House	5	7	4230	freehold	no	no	0.3	8500
50	Apartment	3	3	2861	leasehold	yes	no	0.3	8500
51	Flat	5	4	3466	leasehold	no	no	0.1	8450
52	Flat	5	2	3466	leasehold	no	no	0.1	8450
53	Apartment	5	3	3466	leasehold	no	no	0.1	8450
54	Apartment	3	4	1272	leasehold	yes	yes	0.2	8268
55	Apartment	1	2	962	leasehold	no	no	0.3	3860
56	Flat	2	2	1164	leasehold	yes	yes	0.2	8200
57	Flat	4	4	3565	leasehold	no	no	0.2	7950
58	Terraced	4	6	7150	leasehold	no	no	0.2	7950
59	Penthouse	4	3	3737	freehold	no	no	0.1	7950
60	Flat	3	4	1873	leasehold	no	no	0.3	7850

61	Flat	3	3	1873	leasehold	no	no	0.2	7850
62	Flat	3	3	1604	leasehold	yes	yes	0.3	7850
63	Terraced	5	6	6236	freehold	no	no	0.1	7650
64	Apartment	3	3	1341	leasehold	yes	yes	0.2	7500
65	Apartment	3	3	1907	leasehold	yes	no	0.3	7500
66	Apartment	3	4	1951	leasehold	no	no	0.3	7295
67	Apartment	3	3	1878	leasehold	no	no	0.3	7295
68	Apartment	4	3	1878	leasehold	yes	no	0.3	7295
69	Apartment	3	2	1878	leasehold	yes	no	0.3	7295
70	Apartment	3	3	1873	leasehold	yes	no	0.3	7295
71	Flat	2	2	992	leasehold	yes	yes	0.2	7250
72	Apartment	4	4	3427	leasehold	no	no	0.1	7250
73	Flat	2	3	1272	leasehold	yes	yes	0.2	7200
74	Terraced	12	3	8793	freehold	no	no	0.3	7000
75	House	6	5	4370	leasehold	no	no	0.3	7000
76	Apartment	2	2	1504	leasehold	no	no	0.1	7000
77	Flat	3	3	2418	leasehold	no	no	0.1	6815
78	Penthouse	3	3	1602	leasehold	no	no	0.1	6815
79	Apartment	2	2	2009	leasehold	no	yes	0.2	6750
80	Flat	3	3	1537	leasehold	no	no	0.1	6595
81	Apartment	5	4	2268	leasehold	no	no	0.4	6500
82	Apartment	3	3	1420	leasehold	no	no	0.2	6500
83	Terraced	4	4	2328	freehold	yes	no	0.2	3750
84	Apartment	4	3	2627	leasehold	no	no	0.2	6250
85	Flat	2	3	1472	freehold	yes	no	0.3	6000
86	Flat	3	3	1508	leasehold	yes	yes	0.3	5950
87	Apartment	4	3	2304	leasehold	no	no	0.2	5950
88	Flat	2	1	1472	leasehold	no	no	0.3	5880
89	Apartment	2	3	1399	leasehold	yes	no	0.3	5525
90	Terraced	4	3	2216	freehold	yes	no	0.3	5500
91	Flat	3	2	1399	freehold	no	no	0.2	3650
92	Flat	2	2	1463	leasehold	yes	no	0.3	5380
93	Flat	2	2	1460	leasehold	no	no	0.3	5000
94	Apartment	2	3	1335	leasehold	yes	yes	0.3	4950
95	Apartment	2	1	1399	leasehold	no	no	0.3	4950
96	Flat	5	3	2609	freehold	no	no	0.3	4950
97	Apartment	2	3	1439	leasehold	yes	no	0.3	4865
98	Flat	2	2	1323	leasehold	yes	yes	0.3	4750

99	Penthouse	3	3	1620	leasehold	no	no	0.2	4750
100	Flat	2	2	1691	freehold	no	no	0.2	4650

London SE25 data, Collection date: December 3, 2024

Id	Property_type	Number_of_bedrooms	Number_of_bathrooms	Size_sqft	Tenure	Parking	Sports_facilities	station_distance_miles	House_price_GBPK
1	Terraced	5	2	2321	freehold	yes	no	0.4	1000
2	House	4	2	1667	freehold	yes	no	0.1	950
3	House	4	2	1365	freehold	yes	no	0.8	775
4	House	4	2	1293	freehold	yes	no	0.6	775
5	House	3	1	107	freehold	yes	no	0.3	580
6	Terraced	3	2	1121	freehold	yes	no	0.4	550
7	Terraced	3	1	1131	freehold	no	no	0.4	540
8	Terraced	3	1	909	freehold	yes	no	0.2	535
9	Terraced	3	1	884	freehold	no	no	0.3	535
10	Terraced	3	1	941	freehold	no	no	0.4	525
11	Terraced	4	2	1163	freehold	no	no	0.1	515
12	Apartment	2	2	1032	leasehold	no	no	0.3	500
13	Apartment	3	2	1125	leasehold	no	no	0.2	500
14	Apartment	2	1	805	leasehold	no	no	0.3	475
15	House	2	1	828	freehold	no	no	0.4	455
16	Terraced	3	1	925	freehold	no	no	0.2	450
17	Terraced	4	1	999	freehold	yes	no	0.3	450
18	Terraced	3	1	967	freehold	no	no	0.1	450
19	Apartment	2	1	855	leasehold	no	no	0.3	450
20	Apartment	2	1	698	freehold	no	no	0.2	449.995
21	Apartment	2	3	1032	leasehold	no	no	0.3	445
22	Terraced	2	1	664	freehold	yes	no	0.6	435
23	Terraced	3	1	808	freehold	yes	no	0.1	425
24	Apartment	2	1	834	leasehold	no	no	0.3	425
25	Terraced	2	1	766	freehold	no	no	0.4	390
26	Apartment	2	1	693	freehold	no	no	0.5	375
27	Apartment	2	1	761	freehold	no	no	0.2	375
28	Flat	2	1	535	freehold	no	no	0.2	360
29	Flat	3	1	830	leasehold	no	no	0.1	342
30	Flat	2	1	55	freehold	yes	no	0.3	335
31	Flat	2	1	590	freehold	yes	no	0.3	335
32	Apartment	1	1	610	leasehold	no	no	0.3	330

33	Apartment	1	1	518	freehold	no	no	0.2	330
34	Flat	2	1	775	leasehold	no	no	0.6	325
35	Flat	3	1	743	leasehold	no	no	0.1	325
36	Flat	2	1	1171	freehold	yes	no	0.3	320
37	Apartment	1	1	565	leasehold	no	no	0.3	310
38	Flat	3	1	678	leasehold	no	no	0.3	310
39	Flat	2	1	646	freehold	yes	no	0.3	300
40	Flat	2	1	665	freehold	yes	no	0.3	275
41	Flat	2	1	727	leasehold	no	no	0.3	270
42	Flat	3	1	830	leasehold	no	no	0.1	266
43	Flat	2	1	522	leasehold	yes	no	0.3	260
44	Flat	1	1	530	leasehold	yes	no	0.2	255
45	Flat	1	1	485	leasehold	yes	no	0.3	250
46	Flat	1	1	434	leasehold	yes	no	0.3	250
47	Flat	1	1	514	leasehold	no	no	0.4	235
48	Apartment	1	1	328	leasehold	no	no	0.3	230
49	Flat	2	1	762	freehold	yes	no	0.6	225
50	Flat	1	1	446	leasehold	no	no	0.3	225
51	Flat	3	1	830	leasehold	no	no	0.1	209
52	Flat	1	1	377	leasehold	no	no	0.5	190
53	Flat	1	1	364	leasehold	yes	no	0.3	175
54	Flat	3	1	830	leasehold	no	no	0.1	147.25
55	Flat	3	1	830	leasehold	no	no	0.2	118.75
56	Flat	1	1	500	leasehold	yes	no	0.3	110
57	Flat	6	2	1835	freehold	yes	no	0.3	950
58	Flat	4	2	1606	leasehold	yes	no	0.6	900
59	House	5	2	1572	freehold	yes	no	0.2	900
60	Flat	3	2	1162	leasehold	yes	no	0.6	775
61	Terraced	3	2	944	freehold	no	no	0.8	700
62	Terraced	3	1	990	freehold	yes	no	0.3	700
63	House	6	5	1723	freehold	no	no	0.6	700
64	Terraced	4	1	1315	freehold	yes	no	0.3	700
65	Apartment	3	1	964	freehold	no	no	0.3	675
66	Flat	3	2	1062	freehold	yes	no	0.6	650
67	Flat	2	2	838	freehold	yes	no	0.6	600
68	Apartment	3	1	797	leasehold	yes	no	0.8	520
69	Flat	3	2	922	leasehold	no	no	0.2	500
70	Flat	2	1	848	leasehold	no	no	0.7	500

71	Apartment	2	1	839	freehold	no	no	0.3	495
72	Flat	2	2	827	leasehold	no	no	0.9	450
73	Apartment	2	2	808	leasehold	yes	no	0.9	435
74	Apartment	2	1	742	freehold	yes	no	0.5	425
75	Apartment	4	1	491	freehold	no	no	0.3	425
76	Flat	2	1	785	leasehold	no	no	0.5	365
77	Flat	1	1	468	leasehold	no	no	0.3	350
78	Flat	1	1	622	freehold	no	no	0.1	350
79	Apartment	1	1	790	freehold	no	no	0.2	300
80	Apartment	1	1	470	freehold	yes	no	0.4	300