Proceedings of CONF-MPCS 2024 Workshop: Quantum Machine Learning: Bridging Quantum Physics and Computational Simulations

DOI: 10.54254/2753-8818/52/2024CH0146

Research on house price prediction in UK based on linear regression model

Mengxuan He

Shenzhen Experimental High School, Shenzhen, 518000, China

A02590504@lawsonstate.edu

Abstract. This paper aims to analyze the factors that affect housing prices and predict the trend of housing prices through a linear regression model. This article will use a linear regression model to analyze data on UK house prices from 1952 to 2022. And selected 12 different factors that affect housing prices, using a linear regression model to predict the fluctuations of housing prices. At the same time, a multiple prior regression model was used to explore the importance of each factor in affecting house prices. And the significance of these variables was also compared. The results show that the location of the house, per capita income, housing environment, immigration rate, actual deposit rate and actual loan rate have a significant linear relationship with house price changes, while per capita productivity and house type after the epidemic do not have a significant linear relationship with house price changes. Therefore, UK house prices can be predicted based on the above factors.

Keywords: House price prediction, influencing factors, multiple linear regression model.

1. Introduction

The increase in housing prices is an issue that every citizen has been concerned about [1]. Most people will understand the fluctuations in housing prices in real time and rent or buy a house at a time they think is appropriate, so that the house will be more cost-effective. There are many factors that can cause house prices to rise and fall, and ordinary citizens may not understand these hidden factors, so they cannot judge the price of a house well. Therefore, they cannot judge the price of a house well, so they may buy or rent a house at the highest point of the price or at an inappropriate time. This article will predict the fluctuations of house prices in the UK by studying these factors.

Given that each location and house size are different, the corresponding prices are also different [2, 3]. From 2004 to 2020, the average salary of British citizens fell by 0.97 times, which also caused the house price in the UK to fall to 0.96 times the original [4]. Incomes vary greatly in different parts of the UK. In order to cater to the economic strength of residents in different regions, the government will make different fine-tuning of housing prices in different regions. Which also leads to differences in house prices. However, the overall house prices do not differ too much, but the ratio of house prices to income varies a lot [5]. The COVID-19 outbreak in the UK in 2020 caused a drop in house prices at the time, which was undoubtedly a huge blow to the UK's economic development, which also led to a drop in house prices at the time. However, after the outbreak, the recovery of productivity prompted the recovery of citizens' personal economy, which in turn promoted, but after the outbreak, the recovery of productivity caused the rate of increase in UK house prices to gradually increase [6]. The

DOI: 10.54254/2753-8818/52/2024CH0146

type of housing in the UK also affects the fluctuations in house prices. For example, apartments are cheaper than single-family homes [7]. Given the surrounding environment, people may choose to move to improve the surrounding environment of their residence or increase the value of their house, which can also cause house prices to rise or fall. People prefer residences with convenient transportation and shopping [8]. The local immigration rate and emigration rate will also affect the housing price changes to a certain extent. At the same time, the actual interest rates of personal deposits and loans also lead to differences in personal economic capabilities [9, 10]. Through the data collected, it is found that the actual amount of money that homeowners can receive from selling their homes is gradually decreasing, which will cause most people to be unwilling to sell their homes, resulting in a shortage of housing [11].

In summary, this article will focus on the impact of these 12 factors on UK house prices, and select corresponding models to study the correlation between these factors and house prices.

2. Methodology

2.1. Data source

This study selected UK house price data from 1952 to 2022. The data in this article are all from Kaggle website (UK house prices), nationwide website and Bank of England website. The dataset from Kaggle is called UK Average House Prices and Salary (1975-2020) which is published in 2021. The images are all come from nationwide website and Bank of England website, which provide many research and interpretation of economic. Through the relevant information of the Economic Help website, which provide data for senior high school students to study, and the data of multiple websites are combined for analysis

2.2. Variable selection

House prices are a reflection of the local GDP. At the same time, house prices are also affected by many different factors, most of which are uncontrollable. And because there are too many factors, house prices fluctuate very frequently. However, the author can determine the approximate fluctuation trend based on certain factors, as the illustrated in Figure 1:

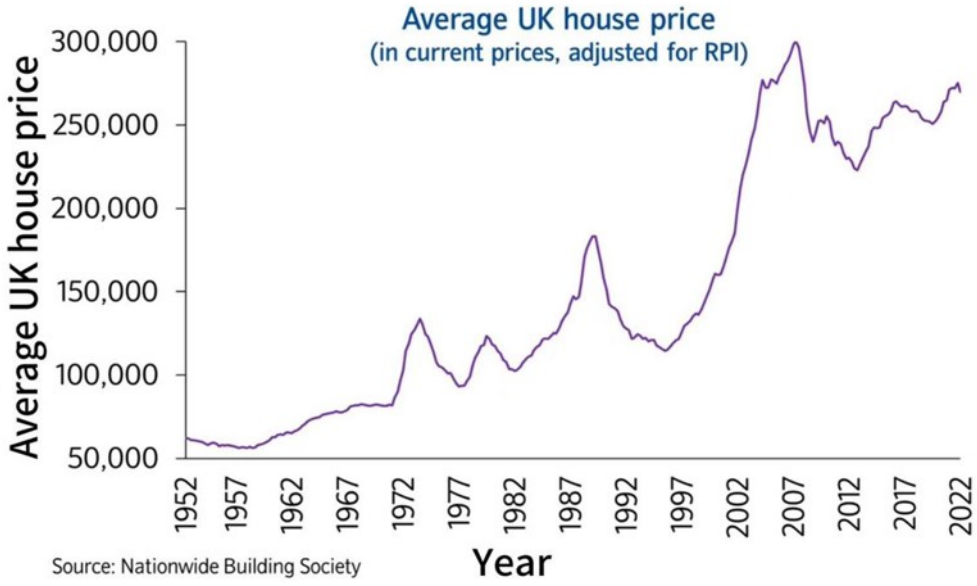

Figure 1. Average UK house price.

DOI: 10.54254/2753-8818/52/2024CH0146

As shown in Figure 1, before 1997, house prices rose steadily, experienced some ups and downs in the middle, and then experienced a significant rise and fall. The house price in 2022 is 1.1 times that of 2004, and the cycle is nearly 5 years. But overall, the average house price in the UK is still on an upward trend. As can be seen from Figure 1, UK house prices reached their highest point in 2007 and then began to plummet. Through the collection of data, the author can understand that the economic recession caused by the financial tsunami on August 9, 2007, investors lost confidence in the value of securities, triggered a liquidity crisis, and the world's developed economies (especially Europe, South America, North America) fell into absolute recession, which led to a sharp drop in house prices starting in 2007. Table 1 shows the variables used in this paper.

Table 1. List of variables.

Variable	Logogram	Meaning
Region	$x1$	Area where the house is located
Salary	$x2$	Average salary
Productivity	$x3$	Post-pandemic productivity per capita
Building Type	$x4$	Flats(1), Terraced(2), Semi(3), Detached(4)
Surroundings	$x5$	Criminal record of the surrounding area
Immigrate	$x6$	Local immigration rate
Deposit	$x7$	Actual deposit rate
Loan	$x8$	Actual loan rate
House Price	Y	Total Housing prices in UK

2.3. Method introduction

This paper uses a linear regression model to compare the impact of different factors on average house prices. Linear regression is a statistical model which estimates the linear relationship between a scalar response and one or more explanatory variables. In linear regression, the relationships are modeled using linear predictor functions whose unknown model parameters are estimated from the data.

Linear regression was the first type of regression analysis to be studied rigorously, and to be used extensively in practical applications. This is because models which depend linearly on their unknown parameters are easier to fit than models which are non-linearly related to their parameters and because the statistical properties of the resulting estimators are easier to determine.

3. Results and discussion

3.1. Data processing

From Figure 2 below, this paper can see that the growth of housing prices and dates is positively correlated, and there is a strong linear correlation. The R-square value of the best fit line in Figure 2 is 0.442. And the p-value is less than 0.05, so this regression model is meaningful.

DOI: 10.54254/2753-8818/52/2024CH0146

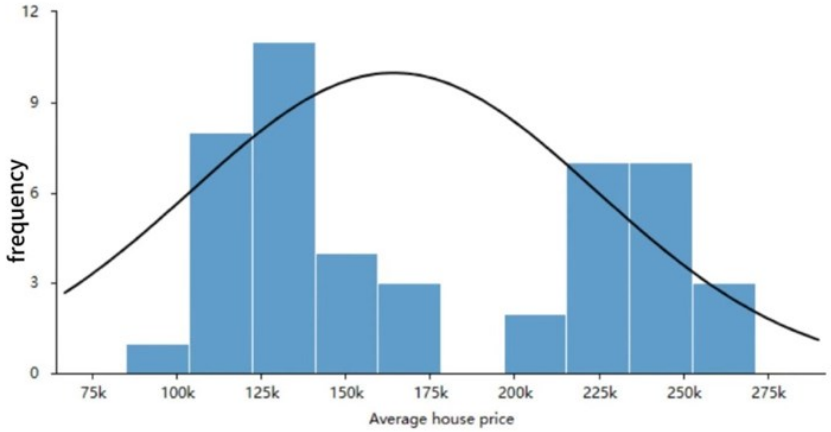

Figure 2. Scatter plot of date and average house price.

The linear fitting formula for scattered data in figure 2 is: $average\ price = -382091.090 + 0.000 * date$, R is 0.442. From table 2, the author can see that when linear regression analysis is performed with date as the independent variable and average price as the dependent variable, the model R-square value is 0.442, which means that date can explain 44.2% of the changes in average price.

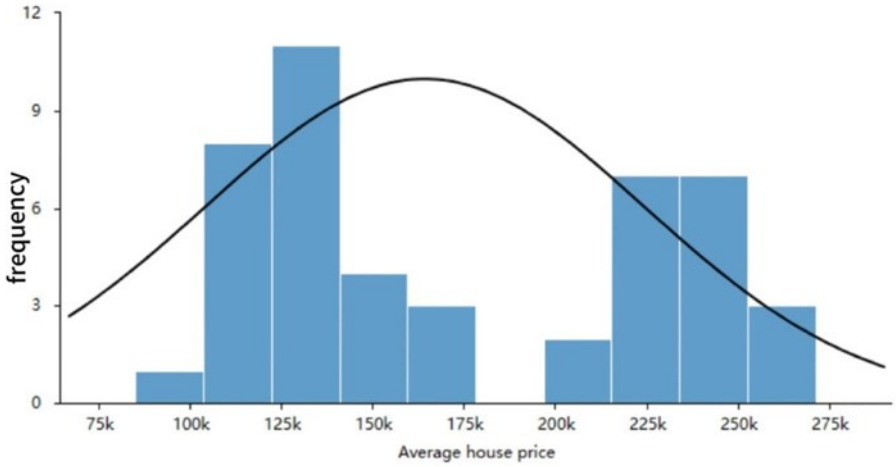

Figure 3. Frequency histogram of average house price.

As can be seen from Figure 3, the majority of UK house prices are between 125k and 137k, and the median of these data is around 162.8k. There are also many houses with prices between 214k and 252k, but most houses are still below 177k.

DOI: 10.54254/2753-8818/52/2024CH0146

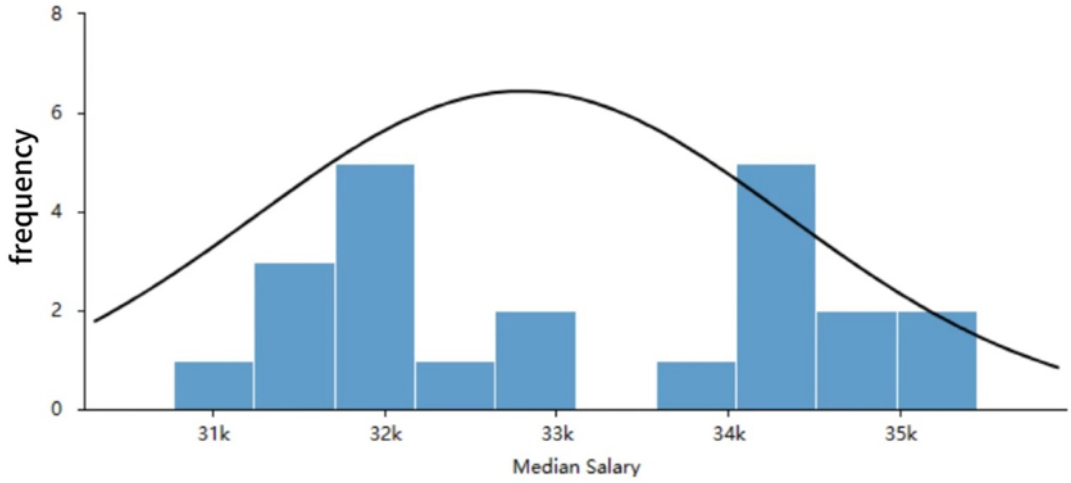

Figure 4. Frequency histogram of median salary.

Figure 4 shows the average salary of local people in the UK. It can be seen that the median per capita income is about 32.8k. In other words, a British citizen can afford an ordinary house with the salary he gets from working locally for 5 years.

3.2. Model results

Figure 5 shows the annual cumulative increase in housing prices. It can be seen that since 2019, housing prices have been on an upward trend, but suddenly fell in June 2020, but then surged. At the end of 2022, housing prices began to decrease, which also confirms the downward trend of housing prices at the end of Figure 1.

Figure 5. Cumulative house price increase by year.

From the data in Table 2, the figures show that x_7 and x_8 have the largest R values and are the main factors affecting housing prices, while x_6 also has a relatively large R value, but the other factors have less impact on housing prices, so loan, deposit and immigration rate are the main influencing factors.

DOI: 10.54254/2753-8818/52/2024CH0146

Table 2. Regression coefficient table.

	B	SE	Beta	T	significance	R
$x1$	-0.152	0.285	-0.152	-0.531	0.605	2.3%
$x2$	4.925	4.715	0.227	1.045	0.309	5.2%
$x5$	23.843	2.413	0.114	9.881	0.000	1.3%
$x6$	1977.371	1027.287	0.563	1.925	0.090	31.7%
$x7$	-7809.962	1785.435	-0.975	-4.374	0.143	95%
$x8$	-3312.253	695.290	-0.979	-4.764	0.132	95.8%

The data in Figure 6 shows that Wales has the largest increase among all regions in the UK, but it is also the region with the most severe decline in 2023. England is the region with the smallest increase in house prices since 2019. Although it has experienced a significant increase between 2021 and 2022, the growth rate is still the lowest among the four regions.

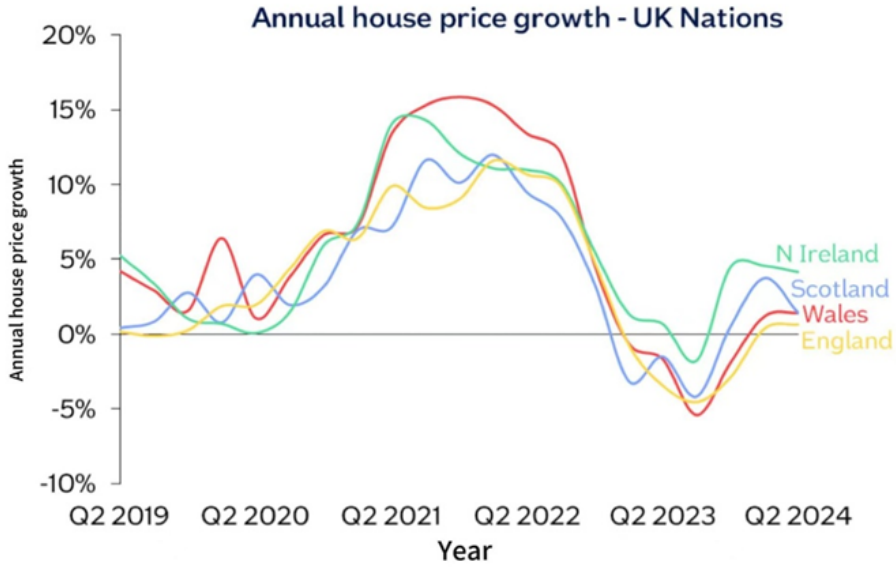

Figure 6. Annual house price growth in UK Nations.

The purpose of a multiple linear regression model is to construct a regression equation that uses multiple independent variables to estimate the dependent variable, thus explaining and predicting the value of the dependent variable. The dependent variables and most of the independent variables in the multiple linear regression model are quantitative values, and some qualitative indicators need to be converted to quantitative values before they can be applied to the regression equation. According to table 2, the author finds that the model can be written as:

$$y = -6451009.042 - 0.152x_1 + 4.952x_2 + 23.843x_5 + 1977.371x_6 - 7809.962x_7 - 3312.253x_8 \quad (1)$$

4. Conclusion

This article studies the UK house price data from 1952 to 2022, and the data contains 12 variables. From this study, it can be concluded that from 1952 to 2022, although the UK house price has fluctuated, it has mainly shown an upward trend in recent years. In addition, due to the impact of the epidemic in recent years, the global economy has been hit hard and is in a recovery period, so the UK house price will be more difficult to predict.

Through this study, people can consider their favorite houses from various angles and make rough house price predictions based on different factors. It is undeniable that given the limited number of samples in this article, the model may have errors, and there may be more different factors affecting

DOI: 10.54254/2753-8818/52/2024CH0146

the changes in house prices, which will affect the accuracy of house prices. But overall, it can still give the public different ideas to predict the development of house prices.

References

[1] Guirguis H, Mueller G, Dutra V, et al. 2024 Advances in Forecasting Home Prices. Comput. Econ.
[2] Chris H, John H and Bruce M 2018 Differing house price linkages across UK regions. Sage Publications, Ltd.
[3] Kuzlu M, Pipattanasomporn M and Rahman S 2015 Impacts of House Sizes, Appliance Ratings and Usage Patterns on Demand Response Applications: A Case-Based Study. Intelligent Industrial Systems.
[4] Luca E 2023 Corrigendum to Renewable energy consumption and per capita income: An empirical analysis in Finland. University of Salerno, 209, 558-568.
[5] Elias O, Steven C. Bourassa M H and Janne E 2023 Revisiting metropolitan house price-income relationships. Journal of Housing Economics.
[6] Robert C M, Beyer RC, Florian M, Claire L, Ezgi O O, and Lev R 2024 Monetary Policy Pass-Through to Interest Rates: Stylized Facts from 30 European Countries. IMF Working Papers.
[7] Wang Z Y, et al. 2023 Housing rental suggestion based on e-commerce data. Knowledge-Based Systems, 268.
[8] Inoue Y and Hashimoto M 2023 Significance of face-to-face service quality in last mile delivery for e-commerce platforms. Transportation Research Interdisciplinary Perspectives.
[9] Alam I A, Mansor H I and Ramzan Y 2023 Impact of House Price on Economic Stability: Some Lessons from OECD Countries. The Journal of Real Estate Finance and Economics.
[10] Mehmet B, Gizem U, Festus Victor B and Mark E W 2023 Housing price uncertainty and housing prices in the UK in a time-varying environment. Empirica.
[11] Omar H M N 2020 An Intra-City Analysis of House Price Convergence and Spatial Dependence. The Journal of Real Estate Finance and Economics.