Research on the Influencing Factors of Housing Price

Yushan Chen[(✉)]

University of Manchester, Manchester, Greater Manchester, UK
yushan.chen@student.manchester.ac.uk

Abstract. The change of commodity prices affects people's lives. House prices have increased over the last 10 years, and more people are prepared to participate in the real estate market. This paper focuses on the influencing factors of housing price. And Examine the elements that influence the cost of the house. The paper mainly uses correlation analysis and linear regression analysis. Finally, under the verification of the two methods, it is concluded that the price of the house is significantly related to the distance from the house to the nearest MRT station, the age of the house, the number of convenience stores around the house and the dimension of the house. In particular, The closest MRT station's proximity to the apartment has the biggest impact on the price of the apartment. Therefore, when investors choose to invest in real estate, they can choose suitable houses for investment according to these influencing factors.

Keywords: House Price · Correlation · Linear Regression

1 Introduction

With the development of economy, more and more people tend to invest in financial products with stable income or low risk. With the growth of population, people's demand for houses is increasing, so that house prices are getting higher and higher. Therefore, more and more people pay attention to the real estate industry. Therefore, more and more people pay attention to the real estate market and want to invest in it. As real estate is the largest asset class in the world, timely house price data becomes easier to find and use [1].

From 1997 to 2006, the real estate price in the United States rose by 124% [2]. From 1997 to 2006, the average inflation index in the United States was only 2.55% [3]. This shows the confidence of many investors in the real estate market. However, the price bubble in the housing market triggered a global financial crisis in 2008 [2]. Excessive real estate prices can't effectively reflect market information, and cannot play a positive role in resource allocation, which leads to the financial crisis [4]. The global financial crisis (GFC) reveals the strategic role of real estate valuation in mortgage loans. The crisis originated in the American subprime mortgage market [5]. Usually, the price can well reflect the relationship between supply and demand. For example, according to the

principle of supply and demand in economics, demand increases and prices rise. On the contrary, prices fall [6].

There is a strong correlation between the cost of a house and its geographical location [1]. The price of a house is related to the distance from important landmarks such as schools, subway stations and hospitals [7]. The distance to these important landmarks is positively related to the house price. The closer the distance, the higher the price [7]. Houses close to these important landmarks are more convenient for people's daily life. So, more people will choose this kind of house. Because of the increase in demand, prices have gone up [6].

Housing-related facilities also have a positive impact on its price [8]. For example, heating facilities, parking spaces, etc. [8]. But for high-end houses, the number of hospitals and toilets will not increase their external value [8]. Therefore, many people will choose houses close to important landmarks in order to bring convenience to their lives. Because of the large demand for this kind of house, according to the principle of economics, the demand increases and the price rises [6].

Urbanization also has some effect on house costs [9]. In fact, urbanization means an increase in population density. The more developed the city, the more people there are, and the higher the demand for housing, which leads to the rise of housing prices [9]. In addition, the rise in housing prices has also brought some pressure to the housing rental market [9]. In addition to the situation of the house itself and its surroundings, the price of the house is related to the laws promulgated by the government. The survey results show that the housing ownership allowance has a distorted effect on the real estate price by capitalizing the house [10].

On this basis, this paper selects several different cities in China and the United States for statistical analysis. The purpose is to find out the most important key factors that affect housing prices through these data.

2   Method

2.1   Data Source and Basic Data Processing

The data in this study include the selling price per unit of house and several factors that may affect the price. Each influencing factor includes 400 data and its corresponding selling price per unit of house. All these data are from 2012 and 2013.

In terms of data processing, this study analyzes each house price and its influencing factors, and then selects the data of different influencing factors on house prices from more than 400 data. For example, Trading dates of houses, the age of the house, how many convenience stores are close to the house, etc. According to 400 data, this article draws a conclusion and find out which influencing factors may affect the unit price of the house. In the following illustration, X1 represents the unit price of the house, The house's transaction time is represented by X2, X3 represents the distance between the house and the closest MRT station, X4 represents t The abundance of convenience retailers in the neighbourhood, X5 represents the age of the house, X6 indicates the longitude and X7 represents the latitude.

2.2   Research Method

In data analysis, this study mainly adopts correlation studies. Correlation analysis refers to the correlation two or even more elements analysis to measure the degree of correlation between the two variables. First of all, we must first see whether there is a significant relationship between variables. Then we can see whether the correlation is positive or negative and explain the tightness of the relationship between variables through the size of the correlation coefficient. In order to establish the link between the dependent variables and other dependent factors in this study, linear regression is also used. Firstly, the fitting of the model is analyzed by R-square. Then combined with regression coefficient, the influence degree of independent variables on dependent variables is compared and analyzed.

3   Results and Discussion

As a way to ascertain what factors will affect the house price, this article selects six possible factors, namely, the transaction date, the age of the property, the quantity of convenience stores nearby, the distance to the closest Metro station, longitude, and latitude.

Table 1 is the correlation analysis between several variables. The strength of the link is indicated by the Pearson correlation coefficient. The correlation coefficients between the house price per square foot and the number of stores, the closest MRT station the age of the house and the latitude and longitude are all significant. There is no significant home price per square foot and transaction date are correlated.

Specifically, the length between the nearest Metro station and the cost per unit area have a -0.674 connection coefficient. Which shows that the cost of a unit's area and the length from the closest Metro station have a very strong negative association. To put it bluntly, the shorter length from the closest MRT station, the more expensive the house. Obviously, this is because the house closer to the MRT station saves people's commuting time and is more convenient. The correlation coefficient between house price of unit area and number of comfort stores is 0.57, which shows that The number of comfort stores and house price per square foot have a strong positive link. Greater number of convenience

Table 1. Correlation results.

X1		X2	X3	X4	X5	X6	X7
X1	1						
X2	0.09	1					
X3	-0.67	0.06	1				
X4	0.57	0.01	-0.6	1			
X5	-0.21	0.02	0.03	0.05	1		
X6	0.52	-0.04	-0.81	0.45	-0.05	1	
X7	0.55	0.04	-0.59	0.44	0.05	0.41	1

shops around the property, the more expensive the house. The correlation coefficient between house price of unit area and house age is -0.21, which shows that age of the home and housing price per square foot are negatively correlated. The older the house, the lower the unit price. The older the house, the less perfect the facilities in the house, so the unit price of the house will be reduced. The correlation coefficients between house price of unit area with longitude and latitude are 0.52 and 0.54, respectively. In addition, there is no correlation between the purchase date and the variable dwelling price of unit area.

Table 2 is the output result with house price of per square foot as dependent variable and length to the closest Metro station, the date of the transaction, number of coincidence stories, house age, longitude and latitude as independent variables. The model formula is: house price of unit area = -14437.103 + 5.146 * transaction date - 0.004 * distance to the nearest MRT station + 1.133 * number of coincidence stories - 0.270 * house age - 12.424 * longitude + 225.473 * latitude. The model's r-squared value is 0.582. This indicates that the 58.2% change in the housing costs per square foot can be explained by the transaction date, the distance to the closest metro station, the number of consignment shops, the age of the house, the longitude, and the latitude. The model passed the F-test ($F = 94.591$, $p = 0.000 < 0.05$), which means that the house price of the unit area will depend on at least one of the following factors: transaction date, proximity to the closest

Table 2. Regression results.

	Denormalization coef.		Standardization coef.	t	p	VIF
	B	S.E.	Beta			
Constant	-14437.103	6775.672	-	-2.131	0.034*	-
transaction date	5.146	1.557	0.107	3.305	0.001*	1.015
distance to the nearest MRT station	-0.004	0.001	-0.416	-6.250	0.000**	4.323
number of convenience stores	1.133	0.188	0.245	6.023	0.000**	1.617
house age	-0.270	0.039	-0.226	-7.000	0.000**	1.014
longitude	-12.424	48.582	-0.014	-0.256	0.798	2.926
latitude	225.473	44.567	0.206	5.059	0.000**	1.610
R2			0.582			
Adjusted R2			0.576			
F			F(6.407) = 94.591. p = 0.000			
D-W Value			2.153			

Metro station, the quantity of convenience stores, the age of the property, longitude, and latitude. In addition, when the model's multiple collinearity is checked, it is discovered that all of the VIF values are less than 5. This indicates that the collinearity issue is solved. Moreover, the model has no autocorrelation because the D-W value is close to 2, and that the sample data are not correlated, proving the validity of the model.

The correlation coefficient for the length to the closest Metro station is -0.004(t = -6.250, p = 0.000 < 0.01). It indicates that the housing price per unit area will be significantly impacted negatively by the distance to the closest MRT station. The correlation factor of number of retail shops is 1.133(t = 6.023, p = 0.000 < 0.01). This indicates that the quantity of retail establishments will significantly increase the area's average home price per square foot. The correlation factor of room age is -0.270(t = -7.000, p = 0.000 < 0.01), which demonstrates how the age of the room will significantly lower the price of a home per square foot. The correlation factor of longitude is -12.424(t = -0.256, p = 0.798 > 0.05), Because p > 0.05, The cost of housing per unit area is not influenced by longitude. The correlation factor of latitude is 225.473(t = 5.059, p = 0.000 < 0.01), which shows that Latitude will significantly increase the area's average home price per square foot. In conclusion, the housing price of the unit area will be significantly influenced positively by the transaction date, the quantity of consignment shops, and the latitude. The age of the home and the distance to the closest Metro station will have a considerable negative impact on a home's pricing. Nevertheless, longitude has little impact on the cost of housing in a given area.

4   Conclusion

This paper analyzes the relationship between housing price and factors affecting housing price from 2012 to 2013. The correlation test and regression test used in this paper prove that The price of a house and the length between that home and the closest Metro station are significantly correlated, the quantity of convenience stores around the house, the age of the house, and the latitude and longitude, which can explain the 58.2% change of the house price. After a comprehensive analysis of seven variables, it is found that The most significant factor affecting the price of a home is its proximity to the closest Metro station. The time of the house's transaction has no effect on the house price. It may be because people generally live a fast-paced life now, so that people need to hope to reduce commuting time to improve their happiness. Therefore, more and more people are willing to choose a house near the MRT station to get a more convenient life. The age of the house represents that people are more inclined to live in a house with better comfort. Because the newer the house, the surrounding supporting facilities and the infrastructure of the community are more perfect, which can bring people a sense of happiness. Because the data is very old, perhaps for the current housing market, the most important factor that may affect the housing price is not necessarily the distance from the house to the closest Metro station.

References

1. Brown, R. J.: Private real estate investment: data analysis and decision making. Amsterdam, Elsevier Academic Press, (2005).

2. Mooya, M. M.: Real estate valuation theory: a critical appraisal. Berlin: Springer, (2016).
3. Bureau of Labor Statistics. Annual inflation rate in the United States from 1990 to 2022. Statista. Statista Inc., (2023). https://www.statista.com/statistics/191077/inflation-rate-in-the-usa-since-1990/, last accessed 2023/04/01.
4. Dec, P., et al.: Price Bubbles in the Real Estate Markets-Analysis and Prediction. Wseas Transctions on Business and Economics19, 292–303 (2022).
5. Amato, M., Coskun, Y.: Property valuation and market cycle. Cham, Switzerland: Springer, (2022).
6. Perloff, J.: Microeconomics. Pearson Education UK, (2014).
7. Abhyankar, A. A., Singla, H. K.: Comparing predictive performance of general regression neural network (GRNN) and hedonic regression model for factors affecting housing prices in 'Pune-India'. International journal of housing markets and analysis 15(2), 451–477 (2022).
8. Özsoy, O., Şahin, H.: Factors affecting housing prices in Izmir, Turkey: a quantile regression approach. International journal of housing markets and analysis 15(1), 145–164. (2022).
9. Nijskens, R., et al.: Hot property: the housing market in major cities. Cham, Switzerland: Springer, (2019).
10. Oertel, C. Y.: Impact of Public Policy Measures on the German Real Estate Market. Wiesbaden: Springer Fachmedien Wiesbaden, (2016).