

# Research on the influencing factors of housing prices in Beijing

Tunan Bai\*

Beijing Huiwen middle school, Beijing, China

\*Corresponding author: 1812231120@mail.sit.edu.cn

**Abstract.** With the development of the times, people's living standards have improved, and housing prices are constantly rising with the increase of price levels. The overall increase in people's demand for houses and the consideration of more environmental and cultural factors have brought huge challenges to the real estate industry. This paper aims to explore various factors that affect housing prices, including floors, construction time, and the presence of elevators. By comprehensively analyzing factors at multiple levels such as policy, economy, market, and society, we can better understand the fluctuations and trends of housing prices, providing decision-making references for the government, homebuyers, and real estate market participants. This study used various methods, including data analysis, literature review, quantitative and correlation analysis, as well as multiple linear regression analysis, to reveal the main influencing factors on housing prices. The main analysis method is linear regression analysis, which specifically analyzes the linear overlap rate of various factors through previous years' data.

**Keywords:** Housing price; influencing factors; Beijing.

## 1. Introduction

Beijing is the capital of China and one of China's economic and cultural centers, so the level of housing prices has always been a concern. The research on the influencing factors of housing prices in Beijing can serve as a reference for industry professionals to predict the development trend of related housing prices, and can also have reference value for evaluating the development of the Beijing real estate market [1].

House price is an important indicator of a country or region's economy, which profoundly impacts individuals and society. Understanding the factors that affect housing prices is of great significance for the government to formulate appropriate real estate policies, homebuyers to make correct decisions, and real estate market participants to make investment decisions. With the development of economic globalization and the rapid changes in the real estate market, studying the factors affecting housing prices has become more necessary. Studying the influencing factors of housing prices in Beijing can provide industry professionals with reference and predict the development trend of related housing prices, and can also provide a reference for evaluating the development of the Beijing real estate market [2].

Gu and Chen used the Granger causality test and ADF test to determine stability and analyzed and empirically concluded that Beijing's housing prices and inflation are Granger reasons for each other; Conducting path analysis reveals the impact of rising loan interest rates on the demand for home purchases and indirectly affects prices [3]. Wang Xiaowei estimated the dynamic impact of regional economic variables on housing price fluctuations in Beijing by constructing an SBAR matrix and conducting a series of test analyses [4]. Through analysis, it has been found that the macro economy profoundly impacts housing prices. Yang and Li analyzed the importance of various indicators of housing prices in China from a macro perspective, connected some economic indicators, and considered regional factors in China [5]. Using the square of R, tested the stable impact of factors such as interest rates and stock prices on housing prices. Gregory C. Chow and Niu have conducted a correlation study was conducted on the impact of supply and demand on Chinese housing prices in the microeconomic context [6]. By calculating the elasticity coefficient and comparing it with the income base, it was concluded that the impact of income on the demand curve is greater than that of

the supply curve [7]. This article uses correlation analysis and the construction of a multiple linear regression model to analyze the micro-influencing factors of housing prices using a relatively simple analysis model and draws conclusions.

## 2. Methods

### 2.1. Data Source and Description

The data source for selecting the Kaggle website is Beijing housing prices from LinkedIn from 2011 to 2017. This article randomly selects 5000 samples from 200000 samples for research. The indicators are numerous and complex, and this article selects indicators that are more relevant to housing prices for data analysis. Including floor, building construction time, whether there is an elevator, whether there is a subway nearby, unit price of 1 square meter, area size, and average community price [8].

### 2.2. Correlation Analysis

From the below table 1, it can be seen that using correlation analysis to study the correlation between total price and 7 items including floor, construction time, elevation, subway, price, square, and community average, the Pearson correlation coefficient is used to represent the strength of the correlation [9, 10]. Specific analysis shows that:

The correlation coefficient value between total price and floor is -0.056, and shows significance at the 0.01 level, indicating a significant negative correlation between total price and floor. The correlation coefficient between Total price and construction time is 0.132 and shows a significant level of 0.01, indicating a significant positive correlation between Total price and construction time. The correlation coefficient between the Total price and the elevator is 0.204 and shows a significant level of 0.01, indicating a significant positive correlation between the Total price and the elevator. The correlation coefficient between the total price and the subway is 0.110 and shows a significant level of 0.01, indicating a significant positive correlation between the total price and the subway. The correlation coefficient between the Total price and price is 0.500 and shows a significant level of 0.01, indicating a significant positive correlation between Total price and price. The correlation coefficient between the Total price and square is 0.693 and shows a significant level of 0.01, indicating a significant positive correlation between the Total price and square. The correlation coefficient value between Total price and community average is 0.521, and shows significance at the 0.01 level, indicating a significant positive correlation between Total price and community average (Figure 1).

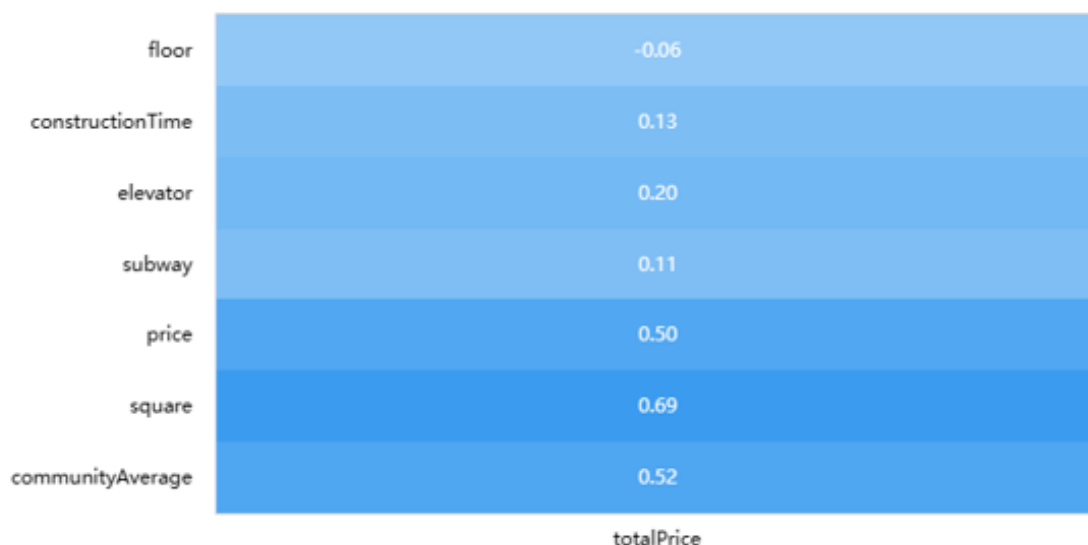


Fig. 1 Correlation results

**Table 1.** Pearson Correlation - Triangle Format

	Total price	floor	Cons time	elevator	subway	price	square	Community
Total price	1							
floor	-0.056**	1						
Cons time	0.132**	-0.08**	1					
elevator	0.204**	-0.26**	0.484**	1				
subway	0.110**	-0.06**	-0.166**	0.141**	1			
price	0.500**	-0.08**	-0.287**	0.062**	0.313**	1		
square	0.693**	0.012	0.313**	0.124**	-0.13**	-0.16**	1	
community	0.521**	-0.06**	-0.271**	0.075**	0.313**	0.933**	-0.09**	1

### 3. Results and Discussion

#### 3.1. Multiple Linear Regression Model Analysis

From table 2, the regression coefficient value of the price is 0.008 ( $t=40.535$ ,  $p=0.000<0.01$ ), indicating that price will have a significant positive impact on total price. The regression coefficient value of the square is 4.871 ( $t=144.379$ ,  $p=0.000<0.01$ ), indicating that the square will have a significant positive impact on the total price. The regression coefficient value of the floor is 0.023 ( $t=0.711$ ,  $p=0.477>0.05$ ), indicating that the floor does not have an impact on the total price. The regression coefficient value of construction time is 2.153 ( $t=9.108$ ,  $p=0.000<0.01$ ), indicating that construction time will have a significant positive impact on the total price. The regression coefficient value of the elevator is 23.765 ( $t=5.790$ ,  $p=0.000<0.01$ ), indicating that the elevator will have a significant positive impact on the total price. The regression coefficient value of the subway is 13.709 ( $t=3.779$ ,  $p=0.000<0.01$ ), indicating that the subway will have a significant positive impact on the total price. The regression coefficient value of the community average is 0.001 ( $t=4.280$ ,  $p=0.000<0.01$ ), indicating that community average will have a significant positive impact on total price.

**Table 2.** Linear regression analysis results (n=4849)

	Non standardized coefficient		Standardized Coefficient	t	p	collinearity diagnosis	
	B	standard error	Beta			VIF	tolerance
constant	-4798.520	472.388	-	-10.158	0.000**	-	-
price	0.008	0.000	0.578	40.535	0.000**	8.145	0.123
square	4.871	0.034	0.778	144.379	0.000**	1.163	0.860
floor	0.023	0.033	0.004	0.711	0.477	1.085	0.922
Construction time	2.153	0.236	0.058	9.108	0.000**	1.642	0.609
elevator	23.765	4.104	0.036	5.790	0.000**	1.521	0.658
subway	13.709	3.628	0.020	3.779	0.000**	1.170	0.854
community average	0.001	0.000	0.061	4.280	0.000**	8.051	0.124
R 2				0.879			
Adjusting R2				0.879			
F				F (7,4841)=5031.743,p=0.000			
D-W value				1.963			

dependent variable: total price  
 \* p<0.05 \*\* p<0.01

From the below table 3, it can be seen that using price, square, floor, construction time, elevator, subway, and community average as the independent variable and total price as the dependent variable for linear regression analysis, it can be seen that the R-squared value of the model is 0.879, which means that price, square, floor, construction time, elevator, subway, and community average can

explain the 87.9% change in total price. That shows the high reliability of trend lines for subways, elevators, and community averages. This further demonstrates that these three indicators can be directly related to the total price.

**Table 3.** Model Summary (Intermediate Process)

R	R 2	Adjusting R2	Model error RMSE	DW value	AIC value
0.938	0.879	0.879	113.088	1.963	59630.638

### 3.2. Anova Test Results

From the below table 4, it can be seen that during the F-test on the model, it was found that the model passed the F-test ( $F=5031.743$ ,  $p=0.000<0.05$ ), indicating that the model construction is meaningful. The significant difference between different independent variables can be observed through the F value in the Anova table.

From the B-value, it can be seen that elevators have a significant impact on overall housing prices, with some elevators being more expensive; The distribution of subways also has a deeper impact on housing prices, and those with subways are more expensive; direct impact; In addition, factors such as unit price, community average price, and floor have a small but positive impact on housing prices; Through the standardization coefficient, also known as Beta, it can be seen that the response of house area to housing prices is significant and has a significant impact; In addition, the R-squared value of 0.879 is closer to 1, so the fitting degree is good and can better explain the variation of the dependent variable. Therefore, the output of this model has a certain contribution.

**Table 4.** ANOVA table (intermediate process)

	Sum of squares	df	mean square
regression	451198897.639	7	64456985.377
residual	62013557.310	4841	12810.072
grand total	513212454.949	4848	

## 4. Conclusion

Overall, all of the above indicators have a certain degree of impact on the total housing price, with elevators and subways having a greater overall impact; The choice of location has a significant impact on housing prices. This indicates the important impact of transportation on housing prices, and elevators and subways, as part of convenient transportation, have a significant positive correlation with housing prices. For the real estate industry, it is important to focus on the close connection between the location of the house and transportation; There are many loopholes in this article, such as outdated data, relatively simple and single analysis methods, unclear explanation of the problem, and few indicators, resulting in less significant explanation of differences. However, it also indicates the correlation between some main factors and total housing prices. As the data is not too new, this article can be compared and analyzed with current data, but people's dependence on transportation continues to exist and the correlation has always been high.

## References

- [1] Gu Liubao, Chen Bofei. Comparative Analysis of Factors Influencing House Price Changes in First and Second tier Cities: Based on Research Data from Beijing and Shijiazhuang. *Statistics and Management*, 2012, (04): 8-10.
- [2] Wang Xiaowei, Cha Jiaqi, Zhang Hanqiao, et al. The impact of regional economic factors on housing price fluctuations in Beijing: a study based on the SVAR model. *Economic Research Guide*, 2016, (01): 58-60.
- [3] Yang Hui, Li Chao. Research on the influencing factors and contribution of urban housing prices in China: A relative importance decomposition based on R2. *Exploration of Economic Issues*, 2019, (11): 49-62.

- [4] Sun Dao. Selection of the Linear Regression Model Recording to the Parameter Estimation. Wuhan University Journal of Natural Sciences, 2000, 5(4): 400-405.
- [5] William H. Econometric Analysis. Beijing: China Social Science Press, 1998.
- [6] Zhang Houcan, Xu Jianping. Modern Psychology and Educational Statistics. Beijing Normal University Press, 2009.
- [7] Zhao Lili, Jiao Jiwen. Grey correlation analysis of factors affecting housing prices. Statistics and Decision, 2007, (23): 2.
- [8] Luo Yubo. Analysis of Factors Influencing House Prices: Quantile Regression Method. Statistics and Decision Making, 2011, (6): 2.
- [9] Cui Zheng, Zhou Mengqian, Kong Lingzhu. A Study on the Heterogeneity of Factors Influencing Urban Housing Prices in China. Taxation and Economics, 2022 (6): 10.
- [10] Tang Hongtao, Liu Yipeng, Wu Zhongcai. Spatial Heterogeneity Analysis of Factors Influencing House Prices Based on POI Data: A Case Study of Changsha City. Urban Issues, 2021, (2): 9.