WEP Warwick Evans Publishing

Transactions on Economics, Business and Management Research
ISSN: 2960-1789, eISSN: 2960-2254 | Volume 10
MEEA 2024

Investigation Upon Factors Contributing to Housing Price Variations-Specifically Sydney for Reference

Haowen Pang *

High School affiliated to China Renmin University, Beijing, 100086, China

* Corresponding Author Email: cnhpang2@babson.edu

Abstract. The purpose of the article is to determine the small aspects that can affect the overall housing prices around the globe. It will be established by looking into one specific area and finding the relationship between them. This research will be using a multiple linear regression model to find the connection. The study used 300 samples and 9 variables each, all are possible variables that can influence overall house price. The data are from Sydney Urban areas collected during the 2020-2021 house market. For the outcome to be valuable, this research will be comparing the significance and VIF value for those variables to find out the biggest connection. Through this process, the null data was deleted as much as possible. As a result, this article discovered that Safety and Median house rent are the biggest influencing factors to the house price. The two factors are the most considered during a house purchase. And the price of Sydney urban houses can be determined by those two factors.

Keywords: Housing prices; Urban area housing; Multiple linear regression model; Influencing factors.

1. Introduction

Housing is a long-passed tradition and way of living from the beginning of the human race. It is crucial for living in society, but in the last two decades house price increased dramatically causing many people unable to purchase their own. Housing prices around the globe have increased along with the overall inevitable inflation, and even more [1]. In fact, house prices from the 19th century to the 20th century were quite stable compared to now. It started to bloom during the latter half of the 20th century. All this house price rise can be attributed to one reason-rising land prices [2]. More than that, population growth can be seen through the reach of 8 billion people. With more people in total, buying a house became even more expensive [3]. Though such huge changes have happened around the globe, many people still have no idea what factor might have caused the increase in price in their surroundings. To understand how small factors influence overall house prices, those factors would need to be separated and evaluated. Therefore, this passage will provide data and evaluation focusing on which factors make the greatest impact on rental prices, and to help people see them for their selves. This paper is here to evaluate the factors that influence house price increases specifically in Sydney's urban areas as an example.

Houses in recent culture has become the symbol of one's wealth. The overpriced real estate market makes people work extremely hard to get their own house. Many people pay their house dept for their entire life. Every month they have to take out large part of their income to pay the dept, leaving them little amount to live and enjoy life with family [4]. However, the cause of overprice is still under investigation and research. Some Chinese scholars such as Pan Jia and Wang Xiaojuan claimed year of building construction is a factor that can make influence to the overall price of houses, so in this research that will be included in this research data to validate it [5, 6]. Safety is also a key aspect people care greatly, so having it considered in the factors is crucial for house price. The analysis of housing market data from 1991 to 2017 for four cities reveals that the main financial hub has the highest house prices but the lowest risk, providing a possibility of a great influence factor [7]. Overall, this research uses 9 variables (Median House Rent(Per Week), Avg. Year Held, Traffic, Public Transport, Nature Environment, Noise, Family Friendliness, Pet Friendliness, and Safety) and one

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.

dependent variable (Median House Price) to discuss the relationship from house price, and study the relevance of their change toward the house price. All factors are chosen from different possible perspectives in order to find the best related variable [8, 9]. For the reasons mentioned above, this passage will focus on using a multiple linear regression model to discuss the effect factors have on house prices [10].

2. Methods

2.1. Data Source

The dataset is from the Kaggle website search, specifically on housing price variables and conditions in Sydney urban areas. All data within this set was collected by KARLTSE during 2020-2021. It includes 30 different features over 421 suburbs in Sydney. This research picked 9 of them for sample. All data remained in .csv format during this research.

2.2. Variable Selection

In the original data set, it contained a great amount of information, including a great amount of unnecessary information for this research. Some of the data consists of lots of nulls such as construction time length and length of time the house has been idle. That information has no benefit to the overall passage. Therefore, the useless data from the data set was excluded from the analysis. After eliminating redundant variables, the data in this research contains 9 variables (Median House Rent (Per Week), Avg. Year Held, Traffic, Public Transport, Nature Environment, Noise, Family Friendliness, Pet Friendliness, and Safety) and one dependent variable (Median House Price). All of the specific details of the data sets of the variables are listed in Table 1 below:

Table 1. List of variables

Variable	Logogram	Meaning
Median House Rent(Per week)	x_1	Median rental price for a week
Avg. Year Held	x_2	Length of time the house was built
Traffic	x_3	Traffic jam possibility near houses
Public Transport	x_4	Whether there is nearby public transportation
Nature Environment	x_5	Surrounding greenery
Noise	x_6	Whether there is a noisy environment in the vicinity
Family Friendliness	x_7	Availability of children's facilities
Pet Friendliness	x_8	Acceptance of pets by neighborhood residents
Safety	x_9	Community crime rate
Median House Price	Y	Median House prices in Sydney

2.3. Method Introduction

In this data set a multiple linear regression model was used to find the interaction between different factors and the influence it can provide. The aim of the model is to compare all situations and seek the few factors making the greatest impact on housing prices, and able to pridict possible housing price changes with different factors the house has. The model used here – Multiple Linear Regression Model, is best to explain the relationship of multiple variables and conclude the mutual effect within. It has the ability to minimize the sum of the squares of the differences between the dependent and independent variables, ensuring the residuals are reduced to their minimum. In this way the model can ensure the data result to be accurate and enough for anylyzing use.

3. Results and Discussion

3.1. Correlation Analysis

The graph below shows the relevance of different factors contributing to the overall price of a house and rank order of relevance.

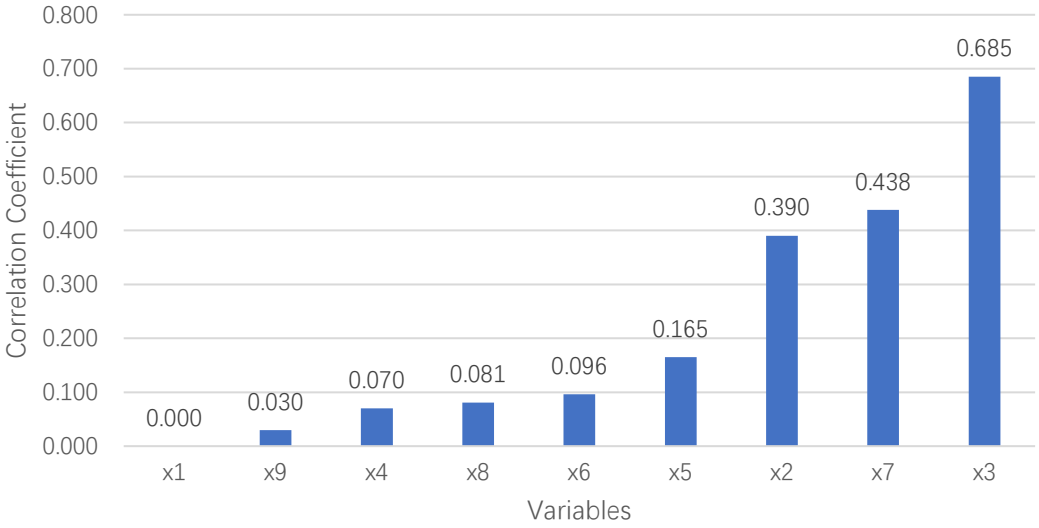

Figure 1. Correlation ranking of variables affecting housing prices

The data in Figure 1 shows the Pearson Correlation between variables in Table 1 and house price. The smaller the number is, the more crucial to house price differences. Median House Rent and Safety, from all of the variables, are the most influential factors. That means people would first consider those factors while purchasing a house, and use it as a tool to consider whether the house is worth the price. As for the rest, they still are being considered but they won't be making a huge influence on overall price and decision. One thing that surprises people is that transportation is negatively correlated with the price of a house, which is not one of the main considerations when buying a house. After all, though there are significant factors and insignificant factors, people still have to think about a lot of different aspects before deciding on purchasing maybe their highest-priced item in life. The decision has to be made wisely to not regret it in later days.

The Pearson Correlation is not the only model that can be used for analysis. Other than that, multiple regression analysis was also great to use for analyzing. The mathematical model is as below:

$$E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{13} x_{12} + e \tag{1}$$

In this formula $\beta_0$is a constant term and e is a residual term.

3.2. Model Results

The Table 2 shows the regression coefficient for multiple linear equation model. Only two of the independent variables x1(Median House Rent (Per week)) and x9(Safety) have a P value (significance value) below 0.03. That indicates them to be having significant effect on the dependent variable Y (Median House Price). And based on this result, the relevant multiple linear regression equation would be:

$$E(Y) = -936039.200 + 2827.938 x_1 + 14090.438 x_2 + \cdots + 61962.171 x_9 \tag{2}$$

Table 2. Regression coefficient table

	B	S.E.	Beta	T	significance	VIF
Constant	-936039.200	377413.855		-2.480	0.000	
X1	2827.938	147.397	0.901	19.186	0.000	1.280
X2	14090.483	16372.360	0.038	0.861	0.392	1.142
X3	11733.213	27402.432	0.022	0.428	0.670	1.587
X4	40492.296	19812.065	0.094	2.044	0.044	1.223
X5	47097.655	30603.334	0.093	1.539	0.128	2.101
X6	-45282.789	25636.386	-0.108	-1.766	0.081	2.180
X7	30444.195	38947.510	0.044	0.782	0.437	1.865
X8	-71507.776	38696.169	-0.103	-1.848	0.068	1.808
X9	61962.171	27131.132	0.130	2.284	0.025	1.882

4. Conclusion

This experiment included 300 sample houses from all over Sydney 2020-2021 urban areas and managed to use 9 different variables for analysis. Multiple linear regression analysis was chosen for its accuracy and efficiency. It provided enough data analysis to determine the variable that has the most impact on overall house price. Which is supported by the Pearson correlation coefficients.

Moreover, analyzing the regression coefficient for multiple linear equation models has the ability to point out relationships between variables and housing prices that are difficult to seek. As for the overall conclusion, safety and house rental price became the most important factors that can influence the house price. And shockingly to know, traffic barely relates to the dependent variable which is the overall house price. This means people don't consider it as the key factor when buying a house.

By investigating this topic, the research provides different aspects and different perspectives for people to understand how house prices can vary. Therefore, giving them a chance to take reference when they are purchasing their own home in the future. Surely this would help people by allowing them to know how house prices are influenced and avoid negative factors. With all the positives, however, there are still some inadequate parts that this research can still be improved. For example, the relatively small sample number, covered areas, and there are still some null information after the elimination. To improve, more data and information will have to be collected in the future in order to find more possible relationships between house prices and different factors. With more information it will be able to fill in the gaps in this research.

References

[1] K. Tsatsaronis, H. Zhu, What Drives Housing Price Dynamics: Cross-Country Evidence, Journal of Monetary Economics. (2004).

[2] K. Knoll, S. Moritz, S. Thomas, No Price Like Home: Global House Prices, 1870-2012, American Economic Review. 107(2) (2017) 331–353.

[3] J.Y. Campbell, J.F. Cocco, how do house prices affect consumption? Evidence from micro data, Journal of Monetary Economics. 54(3) (2007) 591–621.

[4] J. Pan, Y.Y. Luan, X.Y. Hong, M. Li, Analysis and prediction of second-hand housing prices in Qingdao based on integrated algorithms, Advances in Applied Mathematics. 12(4) (2023) 1671-1682.

[5] X.J. Wang, Research on the impact of second-hand housing prices in Chongqing. Journal of Langfang Normal University (Natural Science Edition), 19(3) (2019).

[6] F.N. Chu, I.C. Tsai, do higher house prices indicate higher safety? Price volatility risk in major cities in Taiwan, International Journal of Strategic Property Management. 24(3) (2020) 165-181.

[7] C.G. Yang, H.B. Li, Population Migration, Changes in Residential Supply and Demand, and Regional Economic Development: An Economic Analysis of the Current "Man Snatching War" in Domestic Cities, Theoretical Investigation. 3 (2019) 93-98.

[8] V.N. Stijn, P.O. Weill, Why Has House Price Dispersion Gone Up, Review of Economic Studies. 77(4) (2010) 1567–1606.

[9] A. Gustafsson, S. Wogenius, Modelling Apartment Prices with the Multiple Linear Regression Model, DIVA. (2014).

[10] Z. Peng, Q. Huang, Y.C. Han, Model Research on Forecast of Second-Hand House Price in Chengdu Based on XGboost Algorithm. 2019 IEEE 11th International Conference on Advanced Infocom Technology (ICAIT). IEEE, (2019).