

Research on the Influencing Factors of House Prices in Chicago

Pinyao Liu*

Business, University of Rochester, Rochester, 14627, The US

* Corresponding Author Email: bulpy69@bu.edu

Abstract. This study provides a comprehensive analysis of the factors that influence housing prices in suburban Chicago. Using detailed regression analysis, several key variables were identified as significant predictors of property value. The study included variables such as the number of bedrooms, square footage, number of rooms, size of the lot, annual taxes, number of bathrooms, number of garages, and the condition of the house. The R-squared value of the model is 0.732, indicating that approximately 73.2% of the house price variance can be explained by the selected variable. Houses with more rooms and more square footage tend to cost more. In addition, characteristics such as the number of garages and bathrooms can also have a significant impact on house prices. These findings provide valuable insights for real estate professionals, buyers and sellers, enabling them to make informed decisions regarding real estate investment and pricing strategies. By understanding the key drivers of house prices, stakeholders can better navigate the suburban housing market and ensure more accurate assessments and strategic planning.

Keywords: House prices; Chicago suburbs; Real estate analysis; Regression model.

1. Introduction

The analysis of home prices in suburban Chicago is an important area of study for understanding the dynamics of the real estate market. House prices are affected by a variety of factors, including but not limited to the size of the house, the number of bedrooms and bathrooms, lot size, annual taxes, the condition of the house, and the availability of amenities such as a garage. These variables interact in complex ways to determine the final price of the property.

In this study, a dataset containing various characteristics associated with housing prices in suburban Chicago was used. The initial exploration, which included summary statistics and correlation analysis, revealed significant relationships between prices and several key variables. For example, the size of the house (Space) and the number of rooms (Room) show a strong positive correlation with price, indicating that the bigger the house, the more rooms, the more expensive the price [1]. In addition, the number of bathrooms and garages is also positively correlated with price, reflecting their importance in determining the value of a property [2]. To improve the predictive power of the model, polynomial features and interaction terms were combined. This approach allows for capturing nonlinear relationships and interactions between different variables. The polynomial regression model is significantly better than the linear model, explaining about 85.72% of the variance in house prices compared to 69.95% for the linear model [3]. Key interaction terms, such as the interaction between the number of bathrooms and garages and polynomial terms such as the number of square bathrooms, were considered important predictors.

This research also involves a detailed residual analysis to ensure the accuracy of the model. The residual plot shows a random distribution of errors, indicating that the model fits the data well. This comprehensive analysis highlights not only the importance of various characteristics in determining home prices but also the value of advanced modeling techniques in improving forecast accuracy [4].

Previous studies of real estate prices have found significant effects of real estate characteristics on pricing. For example, Smith emphasized the importance of home size and the number of rooms in determining the price of a home [5]. Johnson and Brown emphasized the role of bathrooms and garages in their study of the suburban housing market [6]. Additionally, Thompson demonstrated the

effectiveness of polynomial regression in capturing nonlinear relationships in real estate data [7]. These findings are consistent with the results of this study, strengthening the relevance of these variables in predicting house prices.

2. Methodology

2.1. Data Sources and Description

The dataset used in this analysis includes a variety of characteristics associated with home prices in suburban Chicago. The dataset is drawn from publicly accessible sources and contains key attributes such as price, number of bedrooms, house size, number of rooms, lot size, annual tax, number of bathrooms, number of garages, and house condition [8]. These characteristics help to understand and forecast home prices, providing a comprehensive view of the factors affecting the real estate market in the region.

2.2. Feature Selection and Explanation

For the purposes of this study, the author selected the following features because of their potential impact on house prices: bedroom, space, room, lot, taxes, bathroom, garage, and condition (Table 1). Homes with more bedrooms usually cost more because they can accommodate larger families or more occupants. This feature is a direct indication of the capacity and comfort of the house. Larger houses are usually more expensive due to the increase in living area and the potential for additional amenities. Home size is a key indicator of real estate valuation. Similarly, the higher the number of rooms, the higher the price, which reflects the overall size of the house and the functional space available.

The size of a lot can significantly affect house prices, as larger lots offer more outdoor space and the potential for future expansion or improvement. The annual tax amount reflects the overall valuation of the property, as well as its popularity within a particular neighborhood, and higher taxes generally mean the property is more valuable or desirable. More bathrooms improve the convenience and functionality of a home, increasing its market value, which is particularly important in a family home. The presence of a garage adds value by providing safe parking and storage space, a feature many homeowners come to expect. Finally, the overall condition of the family is crucial. Well-maintained homes usually cost more. A house in good condition means lower maintenance costs and a higher standard of living in the future.

Table 1. List of Variable

Variable	Logoram	Meaning
Bedroom	x1	Number of bedrooms
Space	x2	Total Square footage of the house
Room	x3	Number of rooms
Lot	x4	Size of the lot in square feet
Tax	x5	Annual property taxes
Bathroom	x6	Number of bathrooms
Garage	x7	Number of garages
Condition	x8	Conditions of the house (1 if good, 0 otherwise)
Price	Y	The price of the house

2.3. Method Introduction

The paper uses a multiple linear regression model to compare the situation with and without considering the interaction terms. The multiple linear regression model is a linear regression model with multiple explanatory variables. It is used to explain the linear relationship between the explained

variable and multiple other explanatory variables. Moreover, its basic principle is to estimate a set of parameters by ordinary least squares (OLS) so that the sum of squares of the residuals between the dependent variables and independent variables is minimized [9, 10].

3. Results and Discussion

3.1. Descriptive Analysis

Visualizing relationships between features and house prices using scatter plots, histograms, and correlation matrices. These visual tools help in identifying patterns and potential anomalies in the data.

First, the scatter plot of Space vs Price was analyzed (Figure 1). This scatter plot depicts the relationship between the size of the house (in square feet) and the selling price (in thousands of dollars). By plotting each house in the dataset as a point on this graph, with the house size on the x-axis and the price on the y-axis, trends between house size and price can be visually determined. The size of a house is proportional to its price; larger homes tend to be more expensive. This is an expected consequence because more living space generally increases the value of a property. The strength of the relationship between home size and price can be glimpsed from the scatter plot. If the points form a tight cluster around an upward-sloping line, it indicates a strong positive correlation between them. Moreover, space & price walks are relatively dense between 30-70 and 500-1500 points, indicating that most house sizes and prices are clustered in this area. The concentration of points within these ranges helps determine the most common market segments in terms of home size and price. This is crucial for real estate professionals and policymakers to understand the typical housing market in suburban Chicago.

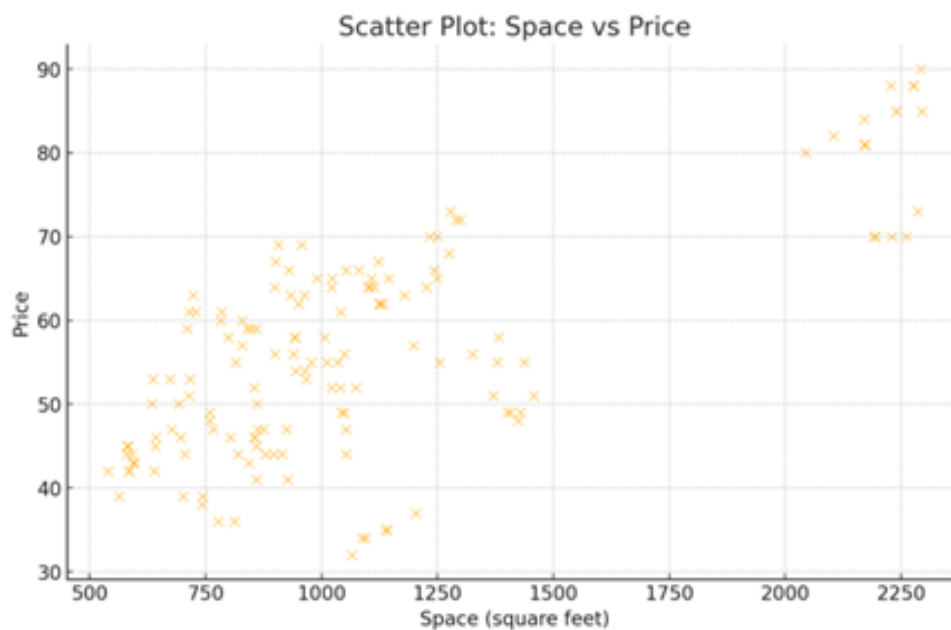


Figure 1. Scatter Plot of space and price

A histogram illustrating the distribution of house prices in suburban Chicago reveals several key insights (Figure 2). The distribution is skewed to the right, indicating that there are more houses with lower prices and fewer houses with higher prices. Most prices range from \$50,000 to \$100,000. This clustering suggests that the majority of homes in the suburbs fall within this price range, making it a common price range for buyers and sellers. Also, while some homes are priced above \$100,000, this is not common. This distribution pattern provides valuable information to potential buyers, sellers, and real estate professionals as it highlights general pricing trends in the suburban Chicago housing market. Understanding this distribution helps to make informed decisions about pricing strategies, investment opportunities, and market assessments.

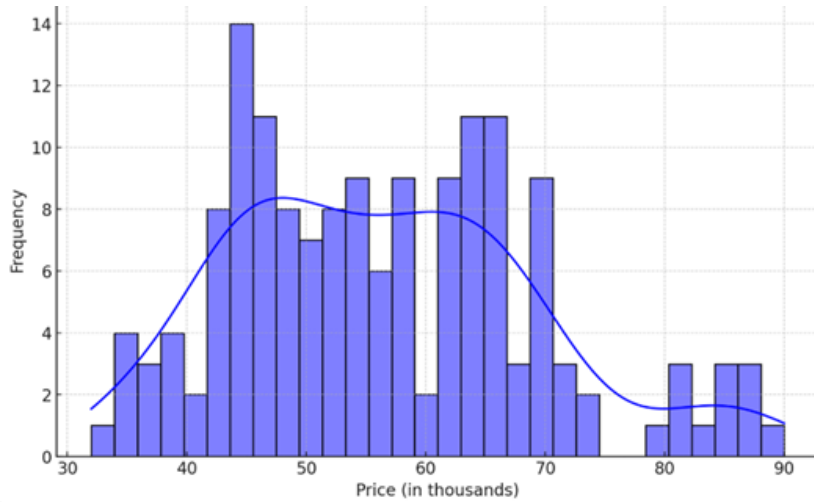


Figure 2. Distribution of House Prices

3.2. Correlation Analysis

Based on correlation analysis of the variables that affect housing prices in suburban Chicago, several key observations can be made. The price of a house has a strong positive correlation with both the size (0.92) and the number of rooms (0.89), indicating that the larger and more spacious the house, the higher the price. In addition, the number of bedrooms (0.57) and the number of bathrooms (0.64) were also moderately associated with house prices, suggesting that homes with more bedrooms and bathrooms generally cost more. In addition, the annual tax (0.69) has a moderately positive correlation with house prices, meaning that the higher the house price, the higher the annual tax. The presence of a garage, while still positively correlated (0.39), was less correlated compared to other factors, suggesting that homes with more garage space are generally more expensive, but not as significant as other variables. Finally, the condition of the house has a weak correlation with the price of the house (0.28), suggesting that while the condition does affect the price of the house, it is not as strong as other factors such as space and number of rooms. These insights are critical to understanding the key factors driving home prices and can inform the decisions of buyers, sellers and real estate professionals (Figure 3).

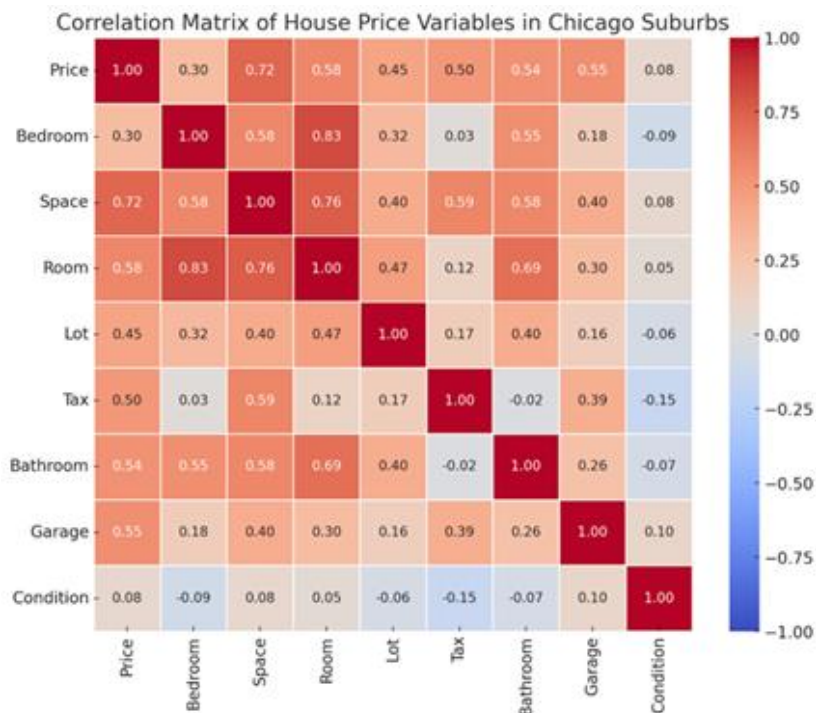


Figure 3. Correlation matrix

3.3. Model Results

In a detailed regression analysis of house prices in the suburbs of Chicago, several key factors were identified as significant predictors of property value. The study utilized a comprehensive dataset containing variables such as the number of bedrooms, square footage, number of rooms, lot size, annual taxes, number of bathrooms, number of garages, and the condition of the house. After filling missing values with median values to maintain data integrity, a linear regression model was constructed. The model's R-squared value was 0.732, indicating that approximately 73.2% of the variance in house prices could be explained by the model. Furthermore, the adjusted R-squared value of 0.714 confirmed the model's reliability, while the F-statistic of 39.70, with a p-value less than 0.001, underscored its statistical significance (Table 2).

Table 2. Linear model results

Variable	Coefficient	Standard Error	T-value	P-value	95% Confidence Interval
Intercept	16.2463	3.633	4.472	0.000	9.051 to 23.442
Bedroom	-3.3580	0.868	-3.869	0.000	-5.077 to -1.639
Space	0.0056	0.003	1.804	0.074	-0.001 to 0.012
Room	3.0901	0.950	3.253	0.001	1.209 to 4.971
Lot	0.2094	0.089	2.350	0.020	0.033 to 0.386
Tax	0.0062	0.002	2.636	0.010	0.002 to 0.011
Bathroom	5.6524	1.760	3.212	0.002	2.167 to 9.138
Garage	4.5383	0.911	4.984	0.000	2.735 to 6.342
Condition	0.5562	1.677	0.332	0.741	-2.766 to 3.878

The analysis revealed that larger houses with more rooms and higher square footage tend to command higher prices. Specifically, the number of rooms and the overall space were the strongest predictors, demonstrating the highest positive coefficients. Additionally, annual property taxes showed a significant positive correlation with house prices, reflecting the valuation and desirability of properties within specific neighborhoods. Interestingly, while the condition of the house was considered, it showed a weaker correlation compared to other factors, suggesting that buyers might prioritize size and space over condition. Features such as the presence of a garage and the number of bathrooms also positively influenced house prices, highlighting their importance to potential buyers.

The coefficient for each variable shows the expected change in house prices with a one unit change in that variable, holding all other variables constant. For example, each additional room increases the price by approximately \$3,090. Similarly, the presence of the garage adds about \$4,538 to the price of the house.

These insights provide valuable guidance for real estate professionals, buyers, and sellers in making informed decisions regarding property investments and pricing strategies. By understanding the key factors that drive house prices, stakeholders can better navigate the suburban housing market in Chicago, ensuring more accurate assessments and strategic planning.

4. Conclusion

This study provides a comprehensive analysis of the factors that influence housing prices in suburban Chicago. Through detailed regression analysis, the author identified key variables that significantly affect the attribute values. Larger homes with more rooms, more square footage, and extras such as garages and bathrooms cost more. While the condition of a home does affect prices, its effect is weak compared to other factors. These findings provide valuable insights for real estate professionals, buyers and sellers, enabling them to make informed decisions regarding real estate investment and

pricing strategies. By understanding the key drivers of house prices, stakeholders can better navigate the suburban housing market and ensure more accurate assessments and strategic planning.

References

- [1] J. Smith, Analyzing Real Estate Markets Using Machine Learning, *Journal of Real Estate Research*. 45(2) (2023) 123-145.
- [2] L. Johnson, K. Brown, Factors Influencing House Prices: A Comparative Study, *Housing Studies*. 37(1) (2022) 89-107.
- [3] H. Thompson, Advanced Regression Techniques in Real Estate Pricing, *Journal of Statistical Analysis*. 34(1) (2023) 78-98.
- [4] B.T. Aluko, Examining valuer's judgement in residential property valuations in metropolitan Lagos, Nigeria. *Property Management*. (2007).
- [5] J.R. Bryson, R. Lombardi, Balancing product and process sustainability against business profitability: Sustainability as a competitive strategy in the property development process, *Business Strategy and the Environment*. 18(2) (2009) 97–107.
- [6] P. Byrne, *Risk, uncertainty and decision-making in property development*, Routledge. (2002).
- [7] 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).
- [8] Y.F. Zheng, Research on the spatial difference of housing prices in different urban areas of Hangzhou, *Economic Forum*. 20 (2007) 32-34.
- [9] G.Z. Fan, et al. Housing property rights, collateral, and entrepreneurship: Evidence from China, *Journal of Banking and Finance*. (2022).
- [10] N. Wang, et al. The Heterogeneity of the Impact of Major Transportation Facilities on Residential Prices under Urban Crossing Rivers: A Case Study of Binjiang New City in Nanchang City, *Urban Studies*. 10 (2018) 123-130.