

Real Estate Market Data Analysis: Correlation Between Housing Prices, Living Space, Building Age and Location

Renyun Huang^{1,a,*}

¹*Qian Weichang College, Shanghai University, Shanghai, 200444, China*

a. 2801148635@qq.com

**corresponding author*

Abstract: The real estate market plays a significant role in economic activity, with housing price fluctuations affected by a variety of factors. Understanding these factors is crucial for investors, developers, and policymakers. This study analyzes the relationship between housing prices and variables such as living space, age of construction, age of renovation, and geographic location using data visualization techniques. The results indicate a positive correlation between living area and housing prices, and show that the distribution of building and renovation ages highlights peak periods of construction and renovation activity. In addition, location has a significant impact on housing prices, which may be related to factors such as land prices, transportation convenience, etc. This study collects data and performs statistical analysis to present and interpret the data through data visualization techniques. The findings suggest that house prices are positively correlated with living area, the distribution of building and renovation ages follows a specific periodic pattern, and location has a significant impact on house prices. This study offers a new perspective on the dynamics of the real estate market and provides valuable insights for guiding investment decision and shaping policy.

Keywords: real estate market, house price, living area, age of construction, geographical location.

1. Introduction

The real estate market, a pivotal component of economic activity, is subject to the whims of numerous factors that cause its prices to fluctuate. These fluctuations are more than mere market dynamics; they reflect the complex interplay of economic, social, and environmental factors. For investors, developers, and policymakers, having a nuanced understanding of the factors influencing these price movements is indispensable. It informs investment decisions, shapes development strategies, and guides the creation of effective policies.

Despite the increasing attention the real estate market receives from researchers, a significant gap persists, particularly in the thorough analysis of basic housing data. There is a dearth of comprehensive studies that scrutinize the intricate relationship between the physical attributes of housing, such as living area and the age of construction and renovation, and the geographic context in which they are situated. This study takes on the challenge of bridging this gap by conducting a detailed examination of these factors and their correlation with housing prices.

Employing a robust methodology that includes meticulous data collection, rigorous statistical analysis, and the use of modern data visualization techniques, this research endeavors to shed light on the multifaceted nature of housing prices. The aim is to provide stakeholders with clear, actionable guidance that can be directly applied to their decision-making processes.

By identifying and analyzing the key determinants of housing prices, this study goes beyond surface-level observations to offer profound insights. It peels back the layers to reveal the underlying mechanisms that drive price variations in the real estate market. These insights are invaluable for decision-makers who must navigate the complexities of the market with confidence and precision.

Moreover, this study's contribution extends to enhancing the collective understanding of real estate market dynamics. It offers a foundation for predicting future trends, identifying potential risks, and crafting policies that promote market stability and sustainable growth. As the real estate market continues to evolve, this research serves as a stepping stone towards a more informed and strategic approach to housing development and investment. Ultimately, it is through such detailed and dedicated research efforts that we can hope to achieve a more balanced and resilient real estate market that meets the needs of all stakeholders.

2. Data analysis

2.1. Data Acquisition

Housing data for Washington, D.C. was selected as the basis for analysis in this study. As the capital of the United States, Washington, D.C. has a unique political, economic, and social position. Its real estate market may represent certain specific market trends and patterns that may not be common in other cities or regions.

As a center for policymaking, Washington's real estate market may be directly affected by federal policies, such as housing subsidies and tax incentives. Examining the impact of these policies on housing prices and market supply and demand can provide valuable insights. In addition, government agencies and the private sector routinely release data related to the real estate market, allowing for a more comprehensive data collection and reporting system in the Washington region, resulting in easier and more accurate access to housing data, including rich details such as the size of the home, the age of the building, the year of renovation, and the geographic location.

2.2. Data visualization and analysis

After obtaining the housing information data in the state of Washington, USA through web crawling techniques, the obtained data was visualized and analyzed. The combination of web crawler technology and data visualization and analysis can provide a more in-depth understanding of the dynamics of the real estate market, and provide strong support for investment decisions and policy making. Scatter plots can visualize the relationship between living area and house price, bar charts show the distribution of housing construction years, and pie charts show the proportion of house price in different geographical locations [1-5].

3. The relationship between house prices and living space

This paper analyzes the existence of a significant positive relationship between residential area and house price using scatterplot and linear fitting methods. This relationship is clearly demonstrated through data collection and visualization techniques. The scatterplot in Figure 1 shows the correspondence between the living area and house price for a large number of houses, while the linear fit line further reveals the positive correlation between the two.

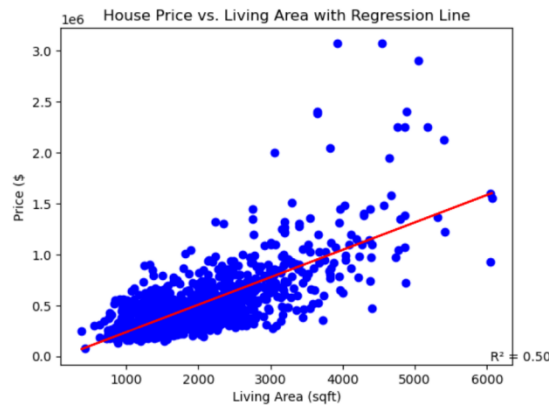


Figure 1: Scatter plot and linear fit of living area and house price

It is evident from the graph that housing prices are also showing an upward trend as the living area increases. This trend is especially pronounced during times of economic prosperity, as people are more inclined to buy or invest in larger living space when economic conditions allow. This not only reflects people's pursuit of a higher quality of life, but also their need for more living space. In this case, the supply of the real estate market often cannot meet the growing demand, which leads to an increase in housing prices.

However, this positive correlation is not true in all cases. During a recession, people's income and purchasing power can be affected, reducing the need for a larger living area. At this time, house prices may fall due to reduced demand. In addition, different types of homes have different impacts on the demand for living space and housing prices.

Single-family villas are often favored for their larger living space and greater privacy, and their prices are naturally relatively high. In contrast, apartments and townhouses may have relatively low rates due to space constraints. In an analysis of the influencing factors of real estate prices, it is mentioned that the housing problem is related to the vital interests of the people and social harmony, and the activity of the real estate market and the increase in housing prices, as well as the macroeconomic control policies introduced by the state, all have a significant impact on housing prices[6]. In addition, another study pointed out that the main driving force for the rise in housing prices came from the level of household consumption, the cost of land acquisition and the area of housing sales [7]. These factors may indirectly explain why single-family villas are more expensive due to their larger space and privacy. The relatively low prices for apartments and townhouses may be due to space constraints. In a study on the analysis and prediction of factors influencing real estate prices based on variable selection, the authors found that house size is an important factor affecting housing prices, and larger dwellings tend to be more expensive [8]. This may explain why apartments and townhouses are relatively inexpensive due to less space.

The increase in the cost of building materials and labor is also an important factor affecting housing prices. As the cost of construction rises, so does the price of the house. Larger living area means higher construction costs, which is directly reflected in housing prices. Therefore, changes in the cost of building materials and labor have a direct impact on housing prices.

In addition to the above factors, the type of home, the level of renovation, and the market supply and demand situation are also key factors that affect the price of the house. These factors are analyzed in detail in an empirical study of the influencing factors of real estate prices in China based on STATA software [7]. These factors may indirectly explain the higher prices of single-family houses and the relatively lower prices of condominiums and townhouses. In a study on the analysis and forecasting of factors influencing real estate prices based on variable selection, the size of the house was an important factor influencing house prices, with larger homes tending to be more expensive.

In addition to differences in living space, different types of houses vary in terms of their architectural structure, design and level of decoration, all of which affect the attractiveness and final price of a house. Homes that are beautifully furnished and uniquely designed tend to attract more buyers, thus increasing their market value.

Furthermore, the supply and demand situation in the market is also an important factor in determining home prices. When there is an adequate supply of homes on the market, home prices may remain stable or decline slightly. And when the supply of homes exceeds the demand, home prices are likely to rise. And market expectations and investor sentiment can also have an impact on home prices. If the market generally expects home prices to continue to rise, investors may increase their home purchases, thus pushing up home prices

4. Distribution of the age of construction and the age of renovation

Building age distribution is an important indicator for understanding the history and trends of the real estate market. By analyzing the number of buildings in different years, it reveals the peaks of construction activities in a particular year, thus providing insight into the cyclical pattern of the construction industry. As shown in Figure 2, the building age distribution chart clearly shows the number of buildings from different periods, and these data provide us with important information about the age distribution of the housing stock. During the half-century from 1950 to 2000, construction activity was more frequent and formed a significant peak. Especially in the years other than 1970, the number of buildings reached more than 120. This pattern usually shows a bimodal distribution, with construction alternating between booms and busts during specific periods.

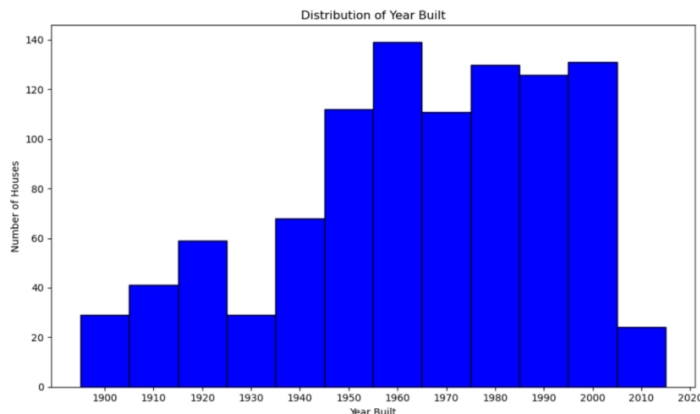


Figure 2: Building age distribution for the bimodal model

The cyclical nature of construction and renovation activities may be related to a number of factors. First, economic development cycles have a direct impact on the construction industry. During periods of economic growth, construction activity usually increases as people's incomes rise and interest in investment grows. Conversely, in periods of economic recession, construction activity may decrease due to financial constraints and reduced willingness to invest.

Second, supply and demand in the real estate market also have a direct impact on construction and renovation activities. When demand exceeds supply in the market, construction and renovation activities may increase. Fluctuations in housing prices can also affect construction and renovation decisions. High house prices may discourage the construction of new buildings, but may increase the demand for renovation of existing houses, as the relatively low cost of renovation can increase the market value of a house.

In addition, government housing and land policies can influence construction and renovation activities. Changes in building codes and safety standards also increase the need to renovate existing

buildings. Existing laws and regulations provide for regular maintenance of buildings based on housing inspections and assessments. The importance of building renewal and renovation efforts has also been emphasized in order to promote green and high-quality urban development. Environmental regulations encourage the use of energy-efficient materials and green building techniques to minimize the environmental impact of buildings and promote a green transition in the construction industry.

As well, socio-cultural changes, such as changes in family structure, may also affect the demand for different housing types and sizes. The growth of single-parent and small families may increase the demand for small housing, while the growth of large families may require more living space. In addition, yearly changes in birth policies, such as changes in family planning policies, may also affect family structure and housing needs.

Finally, climate change and natural disasters can also affect construction and renovation activities. With global warming, extreme weather events will become more frequent, which may require more insulation and waterproofing to cope.

5. The relationship between location and house prices

Location plays a crucial role in the real estate market, and it directly affects the level of housing prices. Through the house price distribution map, it can be seen that the difference in house prices in different regions, and then analyze the specific impact of geographical location on house prices. Figure 3 shows the distribution of house prices by geography on the map, highlighting the impact of location on house prices.

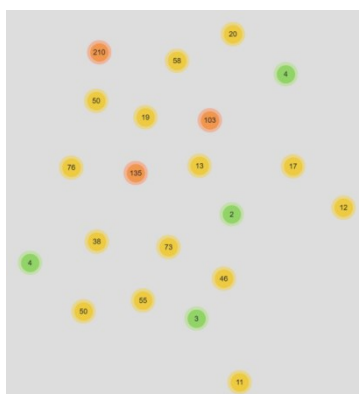


Figure 3: Geography of house price distribution

Generally speaking, land resources are scarce in the city center, land prices are higher and so are property prices. Suburban areas are relatively rich in land resources, with lower land prices and relatively lower house prices. Conveniently located areas, such as those close to major transportation arteries, subway stations and bus stops, usually have higher housing prices. In areas of historical and cultural value, prices increase as properties have investment and collector value. Government urban planning and land policies can also affect house prices. For example, Governments can control house prices by restricting the supply of land or by providing housing subsidies to attract people to live in certain areas. Uneven distribution of educational resources also affects housing prices. In areas with quality schools and educational resources, house prices tend to be higher. Properties near commercial centers and shopping malls tend to have higher prices due to improved convenience and quality of life due to easy access to shopping and entertainment. In areas lacking commercial amenities, home prices are relatively low.

Additionally, home prices tend to be higher in areas with better and safer living environments. Areas with poorer neighborhoods may have lower home prices, but as neighborhoods improve, home prices may rise. Housing prices may be lower in areas with poor environments, such as near industrial

areas or areas with high levels of pollution. Environmental problems in these areas can affect the health and quality of life of residents, so housing prices are relatively low. According to Chen Jie et al, air pollution in neighboring cities has a significant impact on housing prices. The study found that doubling the total amount of air pollutants in neighboring cities within a 100-kilometre radius on average had the potential to reduce local housing prices by about 6%. This suggests that the cross-regional nature of air pollution has an important impact on housing choices and housing prices [9]. With the increase in the level of urbanization, the increase in environmental pollution may reduce the value of urban real estate, which in turn has a negative impact on housing prices. This further confirms the importance of environmental factors in house price decisions [10].

However, the price of housing is not absolute, and with the passage of time and changes in the market, the price of housing will change accordingly. Therefore, investors and home buyers need to consider a variety of factors when choosing a home to make an informed decision.

6. Conclusion

The study concludes with robust findings that there is a significant positive correlation between the size of the residential area and housing prices. Larger living spaces are associated with higher housing costs, providing valuable insights for investors and developers to make strategic investment decisions. Additionally, the identified cyclical patterns in building and renovation ages suggest a link to broader economic cycles, policy changes, and market demands. Geographical location, as posited in the introduction, is confirmed to be a pivotal factor affecting housing prices. The conclusions drawn effectively respond to the research questions introduced in the study's introduction, particularly regarding the influence of residential area and location on housing prices.

While the study offers valuable insights, it acknowledges limitations that may impact the depth and breadth of its conclusions. The primary limitation is the restricted sample size, which may hinder the accuracy and generalizability of the results. Additionally, the geographical confinement of data collection could limit the comprehensive understanding of the real estate market's trends.

To enhance the study's validity and applicability, future research should aim to expand the sample size, encompassing a wider range of regions and covering more time periods. This approach will provide a more reliable and generalizable dataset.

References

- [1] Tian Xueli, Guo Zhibin, Liu Mengxian. *Web page data crawling and visual analysis based on Python[J]. Computer Knowledge and Technology*, 2022, 18(06): 24-26.)
- [2] Liu Xinpeng, Gao Bin. *Using Python and Pandas for student grade processing[J]. Information and Computer(Theoretical Edition)*, 2020, 32(07): 41-43.)
- [3] Liu Xiaozhi. *Information crawling and data analysis of recruitment website based on Python[J]. Electronic Testing*, 2020(12): 75-76+110.
- [4] Que Jinhua. *Python data analysis and visualization based on Ana-conda environment[J]. Information Technology and Informatization*, 2021, 253(4): 215-218..
- [5] Hao Haiyan, Pan Ping. *Electronic Technology and Software Engineering*, 2020, 253(12): 179-181.
- [6] Wang Ruili. *Wuhan University of Science and Technology*, 2011. DOI:10.7666/d.y2041732.
- [7] Suo Menghui. *Analysis of Influencing Factors of Housing Prices: An Empirical Study of China's Panel Data Based on STATA[J]. Progress in Applied Mathematics*, 2022, 11(4): 1503-1511.
- [8] Lai Xinlin. *Analysis and prediction of influencing factors of real estate price based on variable selection[J]. Statistics and Applications*, 2022, 11(3): 12. DOI:10.12677/SA.2022.113062.
- [9] Haiyong Z ,Jie C ,Zhen W . *Spatial heterogeneity in spillover effect of air pollution on housing prices: Evidence from China[J]. Cities*, 2021, 113
- [10] Chen Meng. *Urbanization, Environmental Pollution and China's Housing Prices: An Empirical Analysis Based on a Dynamic General Equilibrium Model[D]. Guangdong: South China Agricultural University*, 2020(in Chinese).