A Comprehensive Study of Factors Influencing Rental Prices in Shanghai

Jiayin Li*

School of Arts and Sciences, Rutgers University, New Brunswick, NJ 08901, United States

*Corresponding author: jl2996@scarletmail.rutgers.edu

Abstract. Shanghai, as a city with a high demand for rental housing, rents and the factors affecting rents are a concern for many individuals. This paper analyzes Shanghai Lane house rental in 2021. Using data from Kaggle, key variables such as square meters, location, latitude and longitude, bedrooms, living rooms, bathrooms, lofts, outdoor spaces, and underfloor heating were examined to understand their impact on rents. Using multiple linear regression and stepwise regression methods, the study identified the important factors that influence rental prices. Correlation analysis and multiple regression models revealed a strong relationship between rent and variables such as square meters, bedrooms, living rooms, dining rooms and bathrooms. The R-square values of the models were 0.689 and 0.360, respectively, indicating that these variables collectively explain a significant portion of the variation in rents. It highlights the positive impact of neighborhoods, living-dining rooms, and outdoor spaces on rents, while lofts negatively affect rental prices. Providing valuable insights for policymakers, real estate professionals, and urban planners, this comprehensive analysis greatly contributes to the understanding of urban housing markets, especially in fast-growing cities such as Shanghai.

Keywords: Shanghai; housing rental prices; influencing factors; apartments.

1. Introduction

As one of the most dynamic and fastest growing cities in the world, Shanghai presents a unique and complex housing rental market. the residential market in Shanghai using price-rent ratios adjusted for parity quality and finds that this adjustment The ratio of price to rent being decreased by 14%. Nonetheless, Shanghai's price-rent ratio was still at a high level internationally in 2017, and its variation across market segments suggests that China's hukou system has an asymmetric impact on owner-occupiers and renters [1]. Housing rental prices in Shanghai are influenced by a variety of factors, including rapid economic growth, population, and an extensive public transportation system, such as the Shanghai subway. In addition, a study by Zhou et al. investigated that the new Metro line's improved access to employment centers resulted the average house price appreciation is 3.75%, with the greatest appreciation in the furthest residential areas. appreciation was most significant for homes with poor initial access to employment centers and attractive job opportunities [2]. Li et al. Analyzed the patterns and factors that influence residential rents in Shanghai, revealing the concentration of high rents in the inner city places a burden on low-income households and non-local individuals. Using publicly available data, the study found that factors such as salary levels, as well as transportation and services, had a significant impact on rents, with variations across regions [3]. A study by Zou et al. examines the housing market in Shanghai and suggests that long-term sustainable housing prices are determined by fundamental factors. The study found that in addition to the impact of housing completions and sales, the consumer price index (CPI) is an important short- and long-term determinant of housing prices. The study concludes that a reduction in the rate of inflation could lead to more stable and sustainable housing prices in the future [4]. Furthermore, Zhou et al. used a fuzzy linear regression model to examine the affordability level of housing prices in Shanghai from a consumer's point of view. The study examines both policy and non-policy factors, such as mortgage interest rates, real estate taxes, and household size. Household size has a negative impact on the affordability level of house prices [5].

House price may also influence the housing rental prices. Complex interactions and price spillovers between submarkets are derived using extensive transaction data from 2004 to 2018 and exploring the diffusion of house prices across 25 identified word markets. A more equitable distribution of urban amenities could improve housing affordability [6]. The value of houses is influenced by the structural attributes and accessibility, and the impact of public and private services is significant. The results vary in different urban areas. In city centers, prices are boosted by amenities such as parks and schools [7]. Therefore, the price of renting a house may be related to the house price acquisition and the surrounding environment. By studying the influences on housing prices and rents in 30 cities in China from 2008 to 2013, it is obtained that there is an intrinsic the correlation between housing prices and rents [8]. Shen et al. address the limitations of previous approaches by utilizing the models FCNN-GWR and GWR to capture rental prices and complex nonlinear relationships and spatial heterogeneity. Quantity-based location and neighborhood variables, such as the number and type of POIs, were also introduced by them. The resulting model shows higher accuracy and stability in predicting rental prices. The results provide a more valuable tool for, among other transactions, rental housing [9].

Wang et al. analyzed the effect of Shanghai metro stations on rental prices by using spatially quantified hedonic regression methods. The average asking plus rent of apartments has a positive correlation with the proximity of the Shanghai subway station, indicating that apartments located close to the metro station result in a significant rental premium [10]. Lastly, a relationship exists between apartment orientation and the Shanghai housing market's prices. South-facing apartments have been found by studies to have a significant premium when it comes to property value. South-facing apartments increase in value by an average of 14%, 6% if they have a view of Shanghai's landmarks, and a further 4% if they are south-facing. The study also found that people preferred nicer, high-floor apartments, especially those with elevators. The orientation and location of apartments emphasizes the importance of the location of an apartment on its price [11]. Providing an in-depth look at the key factors affecting rental prices in Shanghai, this analysis provides a comprehensive view of the current state and trends of the Shanghai housing market, analyzing both geographic location and the size of the home.

2. Methods

2.1. Data Sources

Shanghai Lane House Rentals 2021 as the data for this literature is collected from the Kaggle website, “Lane House” apartment listings in Shanghai.

2.2. Variable Selection

The data set is represented by table 1 List of Variables as follows:

Table 1. List of Variables

Variable	Logogram	Meaning
Sq meters	X1	Housing area
District	X2	Huangpu, Xuhui and other
Latitude	X3	The distance north or south of the equator that is angular
Longitude	X4	The celestial standard meridian is angular to the west
Bedrooms	X5	Total number of bedrooms
Living-Dining	X6	Number of Living-Dining
Bathrooms	X7	Number of bathrooms
Loft	X8	The spatial design of the house (Yes = 1, No = 0)
Outdoor-Space	X9	All outdoor sections, such as courtyards
Floor-Heat	X10	Electric or hydronic systems to heat the floor
Rent	Y	Housing rental price (from 2.5k-220k)

Since the original data has 22 variables, this table chooses to leave 10 variables (Sq meters, district, latitude, longitude, bedrooms, living-dining, bathrooms, loft, outdoor-space and floor-heat) and rent as dependent variables.

2.3. Method Introduction

This paper will utilize the method to investigate the quantitative dependence between the dependent variable and multiple independent variables, use multiple linear regression. Regression coefficients in multiple linear regression analysis are also estimated by the method of least square (MLS). By using the appropriate coefficients to reduce the total amount squares of the residuals of the dependent variable. After modeling the dependent and independent variables, the sum of squared deviations between the theoretical and observed values of the model is minimized.

3. Results and Discussion

3.1. Mutiple Linear Regression

The correlation coefficients of 10 variables as the table 2 Pearson Correlation shows:

Table 2. Pearson Correlation

Variable	rent
district	0.063
Latitude	-0.001
Longitude	0.01
bedrooms	0.75
living-dining	0.579
bathrooms	0.689
loft	-0.037
sq meters	0.811
outdoor space	0.294
floor heat	0.149

From table 2, the relationship between rent and district, bedrooms, bathrooms, living-dining has significant positive correlation, sq meters, outdoor space, and floor heat has significant positive correlation. The relationship between rent and latitude has no significant correlation and p-value is 0.944 (>0.05). The relationship between rent and longitude has no significant correlation and p-value is 0.617(>0.05). The relationship between rent and lot has no significant correlation and p-value is 0.057(>0.05). After examining the Pearson correlation matrix for the different variables, a multiple

regression analysis was performed. The standard formula for multiple linear regression is expressed as:

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

Where β_0 is a constant and e is a residual term. According to table 3 basic indicators, descriptive analysis uses the mean or median to describe the overall data situation. As can be seen in the table above, of the five variables, rent, square meters, latitude, longitude, and bathrooms, the maximum value exceeds the mean by more than three standard deviations. The data is highly variable, and this suggests that the median may be a better indicator of the overall level than the mean.

Table 3. Basic Indicators

Name	Sample	Minimum	Maximum	Mean	SD	Median
rent	2608	2500.000	220000.000	13904.957	13507.051	10000.000
sq meters	2608	1.000	600.000	80.706	69.770	60.000
district	2602	1.000	13.000	7.539	3.370	6.500
Latitude	2608	31.031	121.562	31.248	1.769	31.213
Longitude	2607	121.036	121.691	121.549	0.042	121.562
bedrooms	2608	1.000	10.000	1.669	1.116	1.000
living-dining	2608	0.000	9.000	1.304	0.730	1.000
bathrooms	2608	0.000	15.000	1.332	0.810	1.000
loft	2608	0.000	1.000	0.093	0.290	0.000
outdoor space	2608	0.000	1.000	0.523	0.500	1.000
floor heat	2608	0.000	1.000	0.144	0.351	0.000

3.2. Mutiple Linear Regression with Interaction Terms

As the table 4 linear regression analysis presented, a linear regression analysis was conducted with sq meters, district, Latitude, Longitude, bedrooms, living-dining, bathrooms, loft, outdoor space, and floor heat as independent variables, and rent as the dependent variable.

Table 4. Linear regression analysis

	Unstandardized Coefficients		Standardized Coefficients	t	p	collinearity diagnostics	
	B	Std. Error	Beta			VIF	Tolerance
constant	-3516482.878	782233.191	-	-4.495	0.000	-	-
sq meters	101.226	5.190	0.517	19.506	0.000	5.854	0.171
district	168.097	44.364	0.042	3.789	0.000	1.020	0.980
latitude	79339.113	21946.085	0.040	3.615	0.000	1.019	0.981
longitude	8524.679	3575.494	0.026	2.384	0.017	1.009	0.991
bedrooms	2352.285	273.783	0.193	8.592	0.000	4.187	0.239
living-dining	1372.184	272.522	0.074	5.035	0.000	1.793	0.558
bathrooms	1434.797	311.567	0.085	4.605	0.000	2.842	0.352
loft	-128.358	515.128	-0.003	-0.249	0.803	1.022	0.978
outdoor space	1245.723	315.454	0.046	3.949	0.000	1.134	0.882
floor heat	1105.341	438.552	0.029	2.520	0.012	1.071	0.933
R2			0.689				
adjusted R2			0.688				
F			F (10,2590)=574.075,p=0.000				

The resulting model equation is as follows:

$$E(Y) = -3516482.878 + 101.226\,x_1 + 168.097\,x_2 + 79339.113\,x_3 + 8524.679\,x_4 + 2352.285x_5 + 1372.184x_6 + 1434.797\,x_7 - 128.358\,x_8 + 1245.723x_9 + 1105.341\,x_{10} \quad (2)$$

According to the R-squared value of the model, these variables explain 68.9% of the variation in rent. The F-test (F = 574.075, p < 0.05) confirmed that the rent is greatly impacted by at least one of the variables. Additionally, a check for multicollinearity in the model revealed VIF values greater than 5 but less than 10, indicating potential multicollinearity issues that could be addressed using ridge regression or stepwise regression.

Table 5 Stepwise Regression

	Unstandardized Coefficients		Standardized Coefficients	t	p	collinearity diagnostics	
	B	Std. Error	Beta			VIF	Tolerance
constant	-2253.783	649.299	-	-3.471	0.001**	-	-
district	172.831	63.046	0.043	2.741	0.006**	1.003	0.997
living-dining	10029.397	303.603	0.540	33.035	0.000**	1.083	0.924
loft	-2834.980	731.321	-0.061	-3.877	0.000**	1.003	0.997
outdoor space	3919.951	442.150	0.145	8.866	0.000**	1.084	0.923
R2	0.360						
Adjusted R2	0.359						
F	F (4,2596)=364.512,p=0.000						
D-W value	0.537						

Dependent variable: rent, p<0.05 p<0.01

From table 5, a stepwise regression analysis was performed with district, Latitude, Longitude, living-dining, loft, outdoor space, and floor heat as independent variables, and rent as the dependent variable. Following the model's automatic selection process, district, living-dining, loft, and outdoor space emerged as the significant variables in the model. These variables account for 36% of the variability in rent based on the R-squared value of 0.360. The model's validity was confirmed by passing the F-test (F = 364.512, p < 0.05). The regression formula is as follows:

$$E(Y) = -2253.783 + 172.831x_2 + 10029.397x_6 - 2834.980x_8 + 3919.951x_9 \quad (3)$$

In terms of specific impacts for table 5: The district's regression coefficient (t = 2.741, p = 0.01) suggests a significant positive impact on rent. Living-dining has a coefficient of 10029.397 (t = 33.035, p < 0.01), which indicates a significant positive influence on rental.. Loft has a coefficient of -2834.980 (t = -3.877, p < 0.01), Rent is negatively impacted by outdoor space with a coefficient of 3919.951, but has a notable positive effect on rent. In summary, the variables district, living-dining, and outdoor space have significant positive effects on rent, while loft has a significant negative impact.

4. Conclusion

The study provides a comprehensive look at the impact on the rental market in Shanghai, focused on analyzing Shanghai alleyway housing rentals in 2021, with data sourced from the Kaggle website. The selected variables such as square meters, area, latitude, longitude, bedroom, living-dining room, bathroom, loft, outdoor space, and geothermal heat to understand their impact on rents. The study identifies the key variables that indicate through the use of multiple linear regression analysis and stepwise regression methods. Correlation analysis and multiple regression modeling indicated that there was a strong relationship between rent and the variables of square meters, bedrooms, living-dining rooms and bathrooms. Regression models with R-squared values of 0.691 and 0.360, respectively, indicated that these variables collectively explained a significant portion of the variation in rents. Notably, the study emphasizes the positive impact of area, living-dining, and outdoor space

on rents, while lofts show a significant negative impact. In summary, this study provides a comprehensive and nuanced understanding of the factors that influence rental prices in Shanghai. Results of the study emphasize location and property characteristics. In addition to enriching the existing knowledge base on urban housing markets, this study provides a valuable resource for policy makers, real estate professionals and urban planners.

References

[1] Jie Chen, et al. The User Cost of Housing and the Price-Rent Ratio in Shanghai. Regional Science and Urban Economics, North-Holland, 2021.

[2] Zhou Zhengyi, Chen Hong, et al. The Effect of a subway of House Prices: evidence from Shanghai. Real Estate Ecomomic, 2021, 49(S1). 199-234.

[3] Li Han, et al. Analyzing the Private Rental Housing Market in Shanghai with Open Data. Land Use Policy, Pergamon, 2019, 103738.

[4] Zou, Gao Lu, and Kwong Wing Chau. Determinants and Sustainability of House Prices: The Case of Shanghai, China. MDPI, Multidisciplinary Digital Publishing Institute, 2015, 7(4), 4524-4548.

[5] Zhou, Jian, et al. Affordable Levels of House Prices Using Fuzzy Linear Regression Analysis: The Case of Shanghai - Soft Computing. SpringerLink, Springer Berlin Heidelberg, 2018, 5407-5418.

[6] Hu Jin, et al. The Ripple Effect and Spatiotemporal Dynamics of Intra-Urban Housing Prices at the Submarket Level in Shanghai, China. MDPI, Multidisciplinary Digital Publishing Institute, 2020, 12(12), 5073.

[7] Li Han, et al. Analyzing Housing Prices in Shanghai with Open Data: Amenity, Accessibility and Urban Structure. Cities, Pergamon, 2018.

[8] Zhai Dong, Shang Yishu, Wen Haizhen, et al. Housing Price, Housing Rent, and Rent-Price Ratio: Evidence from 30 Cities in China. Journal of Urban Planning and Development, 2017.

[9] Shen Hang, Li Lin, et al. Exploring a Pricing Model for Urban Rental Houses from a Geographical Perspective. Land, 2022, 11(1), 4.

[10] Wang Yiming, et al. Transit Premium and Rent Segmentation: A Spatial Quantile Hedonic Analysis of Shanghai Metro. Transport Policy, Pergamon, 2016.

[11] Lu Jiajun, et al. The Value of a South-Facing Orientation: A Hedonic Pricing Analysis of the Shanghai Housing Market. Habitat International, Pergamon, 2018.