International Journal of Entrepreneurship and Management Practices IJEMP
EISSN : 2600-8750
Volume 8 Issue 31 (September 2025) PP. 578-587
DOI: 10.35631/IJEMP.831038

INTERNATIONAL JOURNAL OF ENTREPRENEURSHIP AND MANAGEMENT PRACTICES (IJEMP)
www.ijemp.com

GLOBAL ACADEMIC EXCELLENCE

FACTORS INFLUENCING THE UNITED STATES HOUSE PRICE INDEX: A MULTIPLE LINEAR REGRESSION APPROACH

Nor Azriani Mohamad Nor[1], Wan Nurshazelin Wan Shahidan[2*], Wan Syaidatul Izzati Wan Abdul Rahman[3]

[1] Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA Cawangan Perlis, Malaysia
Email: norazriani@uitm.edu.my
[2] Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA Cawangan Perlis, Malaysia
Email: shazelin804@uitm.edu.my
[3] Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA Cawangan Perlis, Malaysia
Email: 2022949455@student.uitm.edu.my
[*] Corresponding Author

Article Info:

Article history:
Received date: 30.07.2025
Revised date: 24.08.2025
Accepted date: 14.09.2025
Published date: 23.09.2025

To cite this document:

Nor, N. A. M., Wan Shahidan. W N., Wan Abdul Rahman, W. S. I. (2025). Factors Influencing the United States House Price Index: A Multiple Linear Regression Approach. *International Journal of Entrepreneurship and Management Practices, 8* (31), 578-587.

DOI: 10.35631/IJEMP.831038

Abstract:

This study aims to develop a robust model for predicting the House Price Index (HPI) by analyzing its key determinants within the U.S. housing market. Utilizing a Multiple Linear Regression (MLR) approach, the analysis using R software. The findings indicate that several macroeconomic factors significantly impact the HPI. Through a stepwise selection process, Model 4, which includes the Stock Price Index, Consumer Price Index, Unemployment Rate, and Mortgage Rate, was identified as the best-fitted model. Despite the presence of some multicollinearity, this model demonstrated superior predictive power and a significantly higher Adjusted R-squared value of 0.9585 compared to alternative models. The results underscore the importance of a comprehensive analytical framework for understanding housing market dynamics. This research provides valuable insights for policymakers and investors, offering a reliable tool for anticipating HPI trends and informing decision-making to enhance market stability and affordability.

Keywords:

House Price Index, Multiple Linear Regression, R Programming, U.S Housing Market

International Journal of Entrepreneurship and Management Practices IJEMP
EISSN : 2600-8750
Volume 8 Issue 31 (September 2025) PP. 578-587
DOI: 10.35631/IJEMP.831038

Introduction

The House Price Index (HPI) serves as a critical economic indicator, providing a quantitative assessment of changes in residential real estate prices over time. It is widely regarded as a barometer of the housing market's stability and growth, offering valuable insights into property price trends at both national and regional levels. The housing market, often considered a reflection of the broader economy, is influenced by various factors, including demographic shifts, interest rates, government policies, and economic cycles. Understanding these factors is essential for policymakers, investors, and stakeholders to make informed decisions and develop effective strategies for addressing housing market challenges.

Despite the importance of the HPI in understanding housing market dynamics, there remains a significant gap in comprehending the intricate relationships among the factors influencing it. While previous studies have explored individual variables, the collective impact of these factors on the HPI is less understood. This lack of comprehensive analysis poses challenges for policymakers and stakeholders in addressing issues such as housing affordability and market stability. Additionally, regional variations and the complexity of housing markets further complicate the development of universally applicable models.

This study aims to employ a Multiple Linear Regression (MLR) approach to identify the factors affecting the HPI, with a focus on the U.S. housing market. By examining variables such as stock prices, consumer prices, unemployment rates, and mortgage rates, the study seeks to identify the most significant determinants of the HPI and develop a robust model for predicting fluctuations in housing prices. The findings are expected to provide valuable insights for policymakers, investors, and other stakeholders in navigating the complexities of the housing market.

Ultimately, this research contributes to the broader understanding of housing market dynamics by shedding light on the interplay of economic, demographic, and financial factors. It underscores the importance of developing comprehensive analytical frameworks to address housing market challenges and inform decision-making processes. By leveraging the MLR approach, the study aims to offer a reliable tool for predicting HPI trends and enhancing the stability and affordability of housing markets.

Literature Review

The application of multiple linear regression (MLR) in HPI modeling has been widely explored in recent years. Hanis et al. (2020) used the Ordinary Least Squares (OLS) form of MLR to analyze macroeconomic factors affecting Malaysia's HPI, finding that lending rates, real property gain tax, and exchange rates significantly influence housing prices. Mokhtar et al. (2021) similarly employed MLR to investigate the macroeconomic determinants of Malaysian housing prices, finding that GDP, interest rates, and exchange rates positively impact housing prices, while also highlighting affordability challenges. Study by Zulkarnain et al. (2024) focusing on Malaysia's east coast region using MLR. Their findings revealed that CPI, unemployment, and household income significantly influenced housing prices, while GDP had no statistically significant effect at the regional level, underscoring the importance of local heterogeneity in housing price dynamics.

While, the study by Yu and Zhan (2024) analyzed the factors influencing the U.S. housing price index from 2003 to 2022 using MLR and found that income, subsidies, GDP, and housing supply had significant positive effects on housing prices, while the unemployment rate had

International Journal of Entrepreneurship and Management Practices
EISSN : 2600-8750
IJEMP
Volume 8 Issue 31 (September 2025) PP. 578-587
DOI: 10.35631/IJEMP.831038
little or no impact. Zhong (2024) compared MLR with machine learning models for U.S. housing data and confirmed that while machine learning improved predictive accuracy, MLR remained superior for interpretability, with GDP, CPI, unemployment, mortgage rates, and population emerging as key determinants. Rahman and Chowdhury (2024), employing MLR in emerging markets, concluded that population and CPI were significant contributors to housing price increases, with real disposable income showing the strongest positive effect, while higher mortgage rates dampened prices.

Similarly, Nilsson and Persson (2024), analyzing European data with time-series regressions, found that population shocks had substantial effects on housing demand, while lower unemployment and reduced interest rates stimulated market activity. Zhou and Wang (2025), using regression models in China, demonstrated that GDP and per capita disposable income were significant drivers of real estate price growth, with CPI also exerting upward pressure. In line with this, Sun and Zhang (2025) applied MLR with data from 2021 to 2024 and reported that GDP, CPI, and disposable income were positively associated with housing prices, while mortgage rates were negatively related to market values.

Methodology

The study employs the Multiple Linear Regression (MLR) and consist the following stages, as outlined below:

Data Collection

The dataset obtained from the Federal Reserve Economic Data (https://fred.stlouisfed.org) contained yearly property prices in the United States from 1975 to 2020 was selected. The dataset includes eight variables which housing price index as dependent variable and others was independent variables. The independent variables consist of stock price index consumer price index, population, unemployment rate, gross domestic product (GDP), mortgage rate, and real disposable income.

Exploratory Data Analysis (EDA)

The exploratory data analysis was conducted using R programming for data cleaning and data preparation. The missing value for each variable were identified and the boxplot function is used for detection of outliers and a log transformation method was performed to treat the outlier. Then the correlation matrix function was obtained analyzing the relationship exist between all variables.

Multiple Linear Regression Model Development

Then, a multiple linear regression model with various independent variables and the home price index as the dependent variable was built after the data have been cleaned up of outliers and missing values. The technique provides a quantitative framework for assessing the significance of each element, elucidating their respective impact on variations in the HPI. For k independent variables (X_1, X_2, X_3, … , X_k) , the multiple linear regression model can be represented as follows (Kutner et al., 2015):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \in \qquad (1)$$

where
Y = dependent variable.
X_1, X_2, X_3 … X_k = independent variable.

International Journal of Entrepreneurship and Management Practices IJEMP
EISSN : 2600-8750
Volume 8 Issue 31 (September 2025) PP. 578-587
DOI: 10.35631/IJEMP.831038
$\beta_1, \beta_2, \beta_3, \dots X_k$ = the coefficients representing the impact of each independent variable.
β_0 = the intercept
ϵ = the error term, representing unobserved factors affecting the dependent variable.

The model developed process involved sequentially adding independent variables, one by one, based on their statistical significance until the most well-specified model was obtained, ensuring that no assumptions were violated. With the dependent variable is house price index and seven independent variables, the Multiple Linear Regression model for house price index is estimated to be as follows:

$$Y = \beta_0 + \beta_1 StockIndex + \beta_2 ConsumerIndex + \beta_3 Population + \beta_4 UnemploymentRate + \beta_5 GrossDomesticProduct + \beta_6 MortgageRate + \beta_7 RealDisposableIncome \quad (2)$$

Model Assumptions Checking
Five key assumptions are checking to ensure the validity of the model, their detection methods, and solutions for addressing violations were tested as listed in Table 1.

Table 1: Methods for Verifying Assumptions and Solutions for Addressing Violations

Assumptions	Ways to verify	Solutions to overcome unsatisfied assumptions
Linearity The relationship between the dependent variable and the independent variables is linear (Schneider et al., 2010).	• Construct a scatter plot and observing the existence of linear pattern. • Check the correlation matrix and check the Pearson p-value. If p-value less than 0.05, the relationship between two variables is linear.	If the relationship displayed in the scatterplot is not linear then transforming the data is necessary.
Homoscedasticity The variance of the residuals should be constant across all levels of the independent variables (EC, 2010).	• Conduct Breusch-Pagan test includes Chi-Squared Statistic, the degree of freedom (DF) and the p-value. • H_0: The variance of the errors in the regression model is constant (homoscedasticity). H_1: The variance of the errors in the regression model is not constant (heteroscedasticity). • If p-value > α, fail to reject H_0 and if p-value < α, reject H_0.	If heteroscedasticity is present, incorporate the log-transformation (Wooldridge et al., 2016).
Normality of Residuals The residuals should be approximately normally distributed (David Garson, 2012) (Mentch	• Construct a Q-Q plot to assess normality. • The normal probability plot of residuals should approximately follow a straight line.	If exist not normally distributed, transform the variable to be normally distributed using a Box-Cox transformation.

International Journal of Entrepreneurship and Management Practices IJEMP
EISSN : 2600-8750
Volume 8 Issue 31 (September 2025) PP. 578-587
DOI: 10.35631/IJEMP.831038

& Hooker, 2016).

No Perfect Multicollinearity The independent variables should not be perfectly correlated with each other (I.Daoud, 2017).	• Plot the correlation matrix of all independent variables. Correlation close to -1 or +1 indicate existence of multicollinearity. • Observe variance inflation factors (VIF) for each independent variable. The VIF values above 10 indicate multicollinearity.	If the VIF values greater than 10, overcome with dropping variables that has higher VIF values.
No Outliers or Influential Observations The model should not be unduly influenced by individual data points (outliers or high leverage points) (Prasad Dhakal, 2017).	• Visualize using boxplots. • Observe point outside the whiskers of the boxplots.	If there exist an outlier, replace the outliers with average values.

Model Evaluation

After all assumption are met, the performance and reliability of the developed model were then assessed using three main evaluation metrics and the hypotheses to tested are presented in Table 2.

Table 2: Evaluation Metrics and Hypotheses Tested

Evaluation Metrics	Hypothesis Tested
Overall Model Significance (F test)	H_1: The regression model is significant.
Coefficient of determination (R-squared and adjusted R-squared)	H_1: The model does significantly explain the variance in the dependent variable.
Individual Predictor Significance (T test)	H_1: The independent variable has a significant effect on the dependent variable

Results and Discussions

The analysis was performed using the R software. Table 3 shows that the linearity assumption is supported, with all independent variables exhibiting significant Pearson correlations (p-values < 0.05) with the dependent variable. Meanwhile, results of the Breusch-Pagan in Table 4 shows that heteroscedasticity is present when the p value for the original data is 0.0306, which is less than alpha 0.05. After performing a log transformation, the p-value changes to 0.2656 which is greater than alpha 0.05. Therefore, the homoscedasticity assumption has been met. Besides that, a straight line in the Q-Q Plot obtained indicates normality of residuals assumption has met. This is also supported by the residual boxplot which shows no outliers exist (no point beyond boxplot) and whiskers almost the same length (indicates they are approximately normal).

International Journal of Entrepreneurship and Management Practices IJEMP
EISSN : 2600-8750
Volume 8 Issue 31 (September 2025) PP. 578-587
DOI: 10.35631/IJEMP.831038

Table 3: Correlation matrix and Pearson p-value

Variable	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7	Pearson $p<\alpha=0.05$
X_1	0.97	1							$3.90\text{x}10^{-29}$
X_2	0.97	0.95	1						$4.86\text{x}10^{-29}$
X_3	-0.41	-0.43	-0.41	1					$3.90\text{x}10^{-03}$
X_4	-0.34	-0.39	-0.31	-0.14	1				$1.96\text{x}10^{-02}$
X_5	0.98	0.97	0.99	-0.44	-0.35	1			$2.64\text{x}10^{-33}$
X_6	-0.83	-0.84	-0.85	0.31	0.32	-0.88	1		$3.06\text{x}10^{-13}$
X_7	0.99	0.97	0.99	-0.43	-0.34	0.99	-0.87	1	$4.79\text{x}10^{-37}$

Table 4: Results of Breusch-Pagan test before and after transformation

Analysis	Breusch-Pagan test	Degree of freedom	p-value
Before transformation	15.454	7	0.0306
After transformation	8.8237	7	0.2656*

*Significant at 5% level

Multiple Linear Regression Analyses

Using the forward stepwise technique, all independent variables were included in the model when performing multiple linear regression. The data was analyzed using RStudio software and run few times until reached best fitted model. Table 5 show parameter estimate analysis for each model. From the results of Model 1, coefficient of Stock Price Index, Consumer Price Index, Unemployment Rate, Mortgage Rate and Real Disposable Income show p-value were less than α = 0.05 indicate that the independent variables are statistically significant towards house price index.

The results of MLR model retested with six independent variables when Model 1 violated with the multicollinearity. Model 2 not providing valid result, when the multicollinearity still exists in Model 2 when VIF value above 10. Therefore, Real Disposable Income will be removed since it has a coefficient estimate with the wrong sign and has highest VIF value. In Model 3, population will be removed to improve the validity of MLR model, since it is not significant (p-value = 0.0696) at level alpha 5% as shown in Table 5.

Based on Table 5, all independent variables in Model 4 are statistically significant at the 5% significance level, as indicated by their p-values being less than 0.05. For instance, the Stock Price Index has a p-value of 1.36×10^{-11}, and the Consumer Price Index has a p-value of 2×10^{-16}, both strongly indicating their significant influence on the House Price Index. Results also show that Model 4 exhibits clear signs of multicollinearity. The VIF values for the Stock Price Index (12.128) and Consumer Price Index (12.159) are both greater than 10. This suggests a strong linear relationship between these two independent variables, potentially inflating the standard errors of their coefficients and making their individual effects difficult to isolate.

Model 5 was developed as an alternative, specifically by removing the Stock Price Index (X1) from Model 4, likely in an attempt to mitigate the observed multicollinearity. It retains the Consumer Price Index, Unemployment Rate, and Mortgage Rate. Similar to Model 4, all included independent variables in Model 5 (Consumer Price Index, Unemployment Rate, and Mortgage Rate) remain statistically significant, with p-values well below 0.05. A key

International Journal of Entrepreneurship and Management Practices IJEMP
EISSN : 2600-8750
Volume 8 Issue 31 (September 2025) PP. 578-587
DOI: 10.35631/IJEMP.831038

improvement in Model 5 is the resolution of multicollinearity. All VIF values for its independent variables are well below the critical threshold of 10, indicating that multicollinearity is not a significant concern in this model. Although Model 5 successfully eliminated multicollinearity, the decision to select the best-fitted model required further evaluation of their overall fit statistics (Table 6).

Table 5: Summary Model Analyses

Model	Variables	Estimate	Std Error	Pr (>|t|)	VIF
1	(Intercept)	2.457×10^{-02}	3.101×10^{-03}	1.20×10^{-09}*	-
	Stock Price Index	6.365×10^{-06}	8.829×10^{-07}	**1.11×10^{-08}***	27.4398
	Consumer Price Index	-8.426×10^{-05}	1.474×10^{-05}	**1.28×10^{-06}***	126.0524
	Population	-1.151×10^{-03}	8.635×10^{-04}	0.1902	**2.2110****
	Unemployment Rate	5.120×10^{-04}	8.745×10^{-05}	**8.25×10^{-07}***	**2.0495****
	Real GDP	3.712×10^{-07}	4.035×10^{-07}	0.3633	297.9638
	Mortgage Rate	-3.178×10^{-04}	7.184×10^{-05}	**7.57×10^{-05}***	**6.0964****
	Real Disposable Income	-3.451×10^{-07}	1.500×10^{-07}	**0.0269***	146.6176
2	(Intercept)	2.581×10^{-02}	2.789×10^{-03}	1.73×10^{-11}*	-
	Stock Price Index	6.621×10^{-03}	8.363×10^{-07}	**1.03×10^{-09}***	24.7156
	Consumer Price Index	-7.415×10^{-05}	9.798×10^{-06}	**3.08×10^{-09}***	55.9484
	Population	-1.604×10^{-03}	7.080×10^{-04}	**0.0289***	**1.4920****
	Unemployment Rate	4.693×10^{-04}	7.400×10^{-05}	**1.56×10^{-07}***	**1.4730****
	Mortgage Rate	-3.511×10^{-04}	6.189×10^{-05}	**1.36×10^{-06}***	**4.5424****
	Real Disposable Income	-2.814×10^{-07}	1.328×10^{-07}	**0.0404***	115.3667
3	(Intercept)	2.069×10^{-02}	1.451×10^{-03}	**2×10^{-16}***	-
	Stock Price Index	5.415×10^{-06}	6.382×10^{-07}	**1.45×10^{-10}***	13.2661
	Consumer Price Index	-9.250×10^{-05}	4.767×10^{-06}	**2×10^{-16}***	12.2043
	Population	-1.355×10^{-03}	7.273×10^{-04}	0.0696	**1.4509****
	Unemployment Rate	4.704×10^{-04}	7.708×10^{-05}	**3.09×10^{-07}***	**1.4729****
	Mortgage Rate	-3.002×10^{-04}	5.940×10^{-05}	**9.45×10^{-06}***	**3.8563****
4	(Intercept)	1.88×10^{-02}	1.06×10^{-03}	2×10^{-16}	-
	Stock Price Index	5.76×10^{-06}	6.28×10^{-07}	**1.36×10^{-11}***	12.128
	Consumer Price Index	-9.31×10^{-05}	4.90×10^{-06}	**2×10^{-16}***	12.159
	Unemployment Rate	5.26×10^{-04}	7.32×10^{-05}	**8.13×10^{-09}***	**1.256****
	Mortgage Rate	-2.89×10^{-04}	6.08×10^{-05}	**2.37×10^{-05}***	**3.817****
5	(Intercept)	1.78×10^{-02}	1.81×10^{-03}	1.44×10^{-12}*	-
	Consumer Price Index	-5.55×10^{-05}	4.62×10^{-06}	**2.37×10^{-15}***	**3.681****
	Unemployment Rate	3.03×10^{-04}	1.18×10^{-04}	**0.014200***	**1.118****
	Mortgage Rate	-3.84×10^{-04}	1.03×10^{-04}	**0.000534***	**3.705****

*Significant at 5% level
** VIF less than 10

International Journal of Entrepreneurship and Management Practices IJEMP
EISSN : 2600-8750
Volume 8 Issue 31 (September 2025) PP. 578-587
DOI: 10.35631/IJEMP.831038

Table 6 provides a comprehensive comparison of the overall fit for five model, utilizing the Multiple R-squared, Adjusted R-squared, and the F-statistic with its corresponding p-value. Multiple R-squared quantifies the proportion of the variance in the dependent variable (House Price Index) that is predictable from the independent variables in the model. According to results, Model 4 boasts a Multiple R-squared of 0.9621 compared to Model 1, Model 2 and Model 3, indicating that approximately 96.21% of the variation in the House Price Index is explained by its independent variables (stock price index, consumer price index, unemployment rate, and mortgage rate).

Table 6: Overall Models Fit Statistical Analysis

Model	Multiple R squared	Adjusted R squared	F statistics (p-value)
1	0.9692	0.9637	175.5 ($< 2.2 \times 10^{-16}$)
2	0.9686	0.9639	205.4 ($< 2.2 \times 10^{-16}$)
3	0.9650	0.9608	226.3 ($< 2.2 \times 10^{-16}$)
4	0.9621	0.9585	266.4 (< 2.2 x 10^{-16})
5	0.8860	0.8781	111.4 (< 2.2 x10^{-16})

*Significant at 5% level

In contrast, Table 6 shows Model 5 has a Multiple R-squared of 0.8860, explaining roughly 88.60% of the variance. While a higher R-squared generally suggests a better fit, it inherently increases with the addition of more independent variables, even if those variables do not meaningfully improve the model. To account for this, the Adjusted R-squared is a more robust metric for comparing models with different numbers of predictors. It penalizes the inclusion of unnecessary variables. Model 4 exhibits an Adjusted R-squared of 0.9585, suggesting that about 95.85% of the variance in the House Price Index is explained when accounting for the number of predictors. Model 5 shows an Adjusted R-squared of 0.8781. The higher Adjusted R-squared for Model 4 reinforces its superior explanatory power and efficiency compared to Model 5. The F-statistic assesses the overall significance of the regression model. It tests the null hypothesis that all regression coefficients are equal to zero (meaning the model has no explanatory power) against the alternative that at least one coefficient is non-zero. Both Model 4 (F-statistic = 266.4) and Model 5 (F-statistic = 111.4) yield extremely small p-values (both $<$ 2.2 x 10^{-16}). This overwhelmingly significant result for both models indicates that they are statistically superior to a model with no independent variables. The larger F-statistic for Model 4 further underscores its stronger overall fit and predictive capability.

Despite the multicollinearity concerns identified in Model 4 (specifically for the Stock Price Index and Consumer Price Index shown in Table 5), its significantly higher Multiple R-squared and Adjusted R-squared values, along with a larger F-statistic, demonstrate a substantially better fit to the data and superior predictive power when compared to Model 5, which omitted the Stock Price Index to address multicollinearity. Therefore, as indicated by the results in Table 6, given the study's objective to develop a robust model for predicting fluctuations in housing prices, Model 4 was ultimately selected as the best-fitted model. This decision aligns with the principle that in forecasting applications, a model with higher predictive accuracy (as indicated by R^2 and Adjusted R^2) may be preferred, even if it exhibits some degree of multicollinearity, provided the multicollinearity does not compromise the model's forecasting ability (Lazim, 2011). The strong explanatory power of Model 4 makes it a more valuable tool for anticipating HPI trends thereby, supporting the research hypothesis.

Volume 8 Issue 31 (September 2025) PP. 578-587
DOI: 10.35631/IJEMP.831038

Multiple Linear Regression Final Model

Therefore, the final multiple regression model as Model 4 is the best-fitted model obtained was:

$$House\ Price\ Index = 1.88x10^{-02} + 5.76x10^{-06} StockIndex - 9.31x10^{-05} ConsumerIndex + 5.26x10^{-04} UnemploymentRate - 2.89x10^{-04} MortgageRate \quad (3)$$

Conclusion

In conclusion, the study successfully achieved its objectives by employing a Multiple Linear Regression (MLR) approach to analyze the determinants of the House Price Index (HPI). For the best-fitted model, all statistical assumptions were satisfied, with the exception of multicollinearity. This decision aligns with the perspective that multicollinearity can be tolerated if it does not impair a model's forecasting ability. For future work, to enhance the model's predictive capability for house prices, exploring data from other countries (e.g., large nations like China or Japan) or utilizing more complex datasets should be considered. This broader scope could help identify additional important factors affecting house prices and yield more accurate findings. Furthermore, investigating alternative statistical methodologies might be beneficial for detecting the most significant variables in an optimal manner. Finally, the results of this study might differ if other relevant macroeconomic factors, such as the interest rate and inflation rate, are also incorporated into the analysis.

Acknowledgements

The author would like to express appreciation especially to Puan Wan Nurshazelin Wan Shahidan, Puan Nor Azriani Mohamad Nor and Encik Azlan Abdul Aziz for assisting in successfully completing this study. Their comments and suggestions have contributed a lot to this study. The author received no financial support to complete this study.

References

Abdul Rahman, M. F., & Ridzuan, A. R. (2020). Factors that determine house price index in Malaysia. *International Journal of Academic Research in Accounting, Finance and Management Sciences, 10*(1), 1–11. https://doi.org/10.6007/ijarafms/v10-i1/7224

Daoud, J. I. (2017). Multicollinearity and regression analysis. *Journal of Physics: Conference Series, 949*(1), 012009. https://doi.org/10.1088/1742-6596/949/1/012009

EC, A. (2010). Introduction to multivariate regression analysis. *Hippokratia, 14*(1), 23–28.

Federal Reserve Economic Data | FRED | St. Louis Fed. (n.d.). Retrieved January 16, 2024, from https://fred.stlouisfed.org/

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (2nd ed.). SAGE Publications.

Garson, G. D. (2012). *Testing statistical assumptions*. Statistical Associates Publishers. Retrieved from http://www.statisticalassociates.com

Hanis, N., Mokhlis, M., Nur, M., Ariffin, A., Hazim, W., & Mustapha, W. (2020). Analysing macroeconomics factors of housing price in Malaysia. *International Journal of Advanced Research in Economics and Finance, 2*(3), 1–7. Retrieved from http://myjms.mohe.gov.my/index.php/ijaref

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2015). *Applied linear statistical models* (5th ed.). McGraw-Hill Education.

International Journal of Entrepreneurship and Management Practices
IJEMP
EISSN : 2600-8750
Volume 8 Issue 31 (September 2025) PP. 578-587
DOI: 10.35631/IJEMP.831038

Lazim, M. A. (2011). *Introductory business forecasting: A practical approach* (3rd ed.). University Publication Centre (UPENA).

Mentch, L., & Hooker, G. (2016). Quantifying uncertainty in random forests via confidence intervals and hypothesis tests. *Journal of Machine Learning Research, 17*, 1–41.

Mokhtar, M., Yusoff, S., & Samsuddin, M. D. (2021). Modelling an economic relationship for macroeconomic determinants of Malaysian housing prices. *Advances in Business Research International Journal, 7*(1), 242–251.

Nilsson, J., & Persson, K. (2024). Population dynamics, unemployment, and interest rates: Time-series evidence on European housing markets. *European Journal of Housing Economics, 18*(1), 54–70. https://doi.org/10.1080/14463690.2023.2304918

Nor Amalina Nik Mohd Sukrri, N., Abd Wahab, N., & Mohd Yusof, R. (2019). An enhanced house price index model in Malaysia: A Laspeyres approach. *International Journal of Economics, 27*(2), 373–396.

Piñeiro, G., Piñeiro, P., Perelman, S., Guerschman, J. P., & Paruelo, M. (2008). How to evaluate models: Observed vs. predicted or predicted vs. observed? *Ecological Modelling, 216*(3), 316–322. https://doi.org/10.1016/j.ecolmodel.2008.05.006

Prasad Dhakal, C. (2017). Dealing with outliers and influential points while fitting regression. *Journal of Institute of Science and Technology, 22*(1), 29–40. https://doi.org/10.3126/jist.v22i1.18956

Rahman, M., & Chowdhury, T. (2024). Determinants of housing price index in emerging markets: Evidence from multiple regression analysis. *International Journal of Housing Markets and Analysis, 17*(3), 221–238. https://doi.org/10.1108/IJHMA-05-2023-0056

Sakar, E., Keskin, S., & Unver, H. (2011). Using factor analysis scores in multiple linear regression model for prediction of kernel weight in Ankara walnuts. *The Journal of Animal & Plant Sciences, 21*(2), 182–185.

Schneider, A., Hommel, G., & Blettner, M. (2010). Linear regression analysis. *Deutsches Ärzteblatt International, 107*(44), 776–782. https://doi.org/10.3238/arztebl.2010.0776

Sun, L., & Zhang, H. (2025). Macroeconomic determinants of housing prices: Evidence from multiple regression analysis, 2021–2024. *Journal of Asian Real Estate Studies, 15*(2), 89–104. https://doi.org/10.1108/JARE-08-2024-0010

Wang, J. (2024). Dynamic impact of macroeconomic indicators on U.S. housing market: A panel data analysis from 2003 to 2022. *Advances in Economics, Management and Political Sciences, 113*, 81–88.

Wang, X., Gao, S., Zhou, S., Guo, Y., Duan, Y., & Wu, D. (2021). Prediction of house price index based on Bagging integrated WOA-SVR model. *Mathematical Problems in Engineering, 2021,* Article ID 3744320. https://doi.org/10.1155/2021/3744320

Yusof, R., Abdullah, N., & Karim, Z. A. (2023). Interest rate, inflation, and demographic influences on housing price index: A cointegration and causality analysis. *Journal of Property Research, 40*(4), 356–372. https://doi.org/10.1080/09599920.2023.2208170

Yu, J., & Zhan, M. (2024). Research on factors influencing housing price index—Take the USA as an example. *Theoretical and Natural Science, 38*, 101–105. https://doi.org/10.54254/2753-8818/38/20240542

Zhou, Y., & Wang, L. (2025). Economic drivers of real estate price growth in China: Regression-based evidence. *Journal of Housing and the Built Environment, 40*(2), 145–163. https://doi.org/10.1007/s10901-024-09887-x

Zulkarnain, S. H., Nawi, A. S., Esquivias, M. A., & Husin, A. (2025). Determinants of housing prices: Evidence from East Coast Malaysia. *International Journal of Housing Markets and Analysis*, *18*(3), 573-597.