ENHANCED KNOWLEDGE IN SCIENCES AND TECHNOLOGY
e-ISSN: 2773-6385

EKST

Vol. 6 No. 1 (2026) 264-272
https://publisher.uthm.edu.my/periodicals/index.php/ekst

Factors Affecting Housing Prices in Malaysia Using Regression and Correlation Analysis

Kyau Chin Wen[1], Cik Sri Mazzura Muhammad Basri[1]*

[1] *Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, UTHM Kampus Cawangan Pagoh, Hab Pendidikan Tinggi Pagoh, KM 1, Jalan Panchor, 84600 Pagoh, Muar, Johor, MALAYSIA.*

*Corresponding Author: mazzura@uthm.edu.my
DOI: https://doi.org/10.30880/ekst.2026.06.01.032

Article Info

Received: 2 January 2026
Accepted: 19 January 2026
Available online: 9 July 2026

Keywords

Factors, Housing Prices, Malaysia, Regression, Correlation, Analysis

Abstract

Housing is a fundamental requirement of every citizen in every country. Thus, it is important to understand what factor affects housing prices. Regression and correlation analysis are the techniques used in this research to investigate the factors influencing housing prices in Malaysia. There are four identified influencing factors are examined: Gross Domestic Product (GDP), inflation rate, population and labour force. Simple linear regression is applied to study the impact of each factor individually on housing prices, whereas multiple linear regression is applied to investigate the combined effects of all factors collectively. All analyses involved secondary data ranging from year 2020 to 2024, collected from various reliable sources. Correlation coefficients, R^2 and significance testing are used to measure the reliability of the regression models and correlation analysis. The results show that the effects of each factor on housing prices are different, with GDP showing the strongest positive relationship. These findings can guide policymakers, investors and stakeholders in developing policies and strategies to address the Malaysian housing market effectively.

1. Introduction

The Government of Malaysia recognises that housing is an important element of the urban economy as well as a basic need of every citizen [1]. A house serves as both a place to live and an investment because house ownership is often considered a well valued asset in most regions of the world. Homebuyers make a major decision when purchasing a house since expenditure on homeownership will greatly affect the standards of living of the homeowner. In Malaysia, housing demand is expected to increase due to the positive correlation between housing performance and occupants' productivity, safety, well-being and satisfaction [2].

Housing is unaffordable in many countries around the world including Malaysia [3]. Despite strong demand, Malaysia faces housing shortage problems, with a population of 31 million and a total housing stock of 7.2 million, resulting in fewer than 250 housing units per 1000 people [4], [5]. This is because accelerated economic development has increased the demand for residential housing in Malaysia, leading to significantly higher housing prices. Between 2001 and 2010, Malaysia's house price index increased from 101.9 to 146.9, while average house prices surged from RM141,494 to RM203,495. Data from the Valuation and Property Services Department (JPPH) also reported an increment of 45% in the average price of all houses between 2001 and 2010 [1]. Research also shows that when there are more buyers than sellers due to speculation, housing prices tend to increase and may form a price bubble [6].

The Malaysian government has implemented various schemes, programmes, and policies to increase housing supply and homeownership. For example, Perumahan Rakyat 1Malaysia (PR1MA), Rumah Mesra Rakyat (RMR1M), Program Rumah Mampu Milik (RMM) and others. Additional initiatives including subsidies and tax reliefs, relaxed Employees Provident Fund (EPF) regulations, removal of Goods and Services Tax (GST) and loosening labour policies, have further boosted the housing industry. Developers also offered discounts, staged payments, and bridging loans, while the bureaucratic processes for land allocation have become easier in recent years [2].

Gross Domestic Product (GDP) has been found to have a significant impact on housing prices. Past research mentioned that GDP will always remain a vital macroeconomic indicator in understanding and forecasting housing price trends in Malaysia [7]. Besides that, inflation is also a driver of housing prices. Over half of the total variation in housing prices over a five-year period across different nations can be attributed to inflation on average, suggesting a strong link between inflation and housing prices [8]. Inflationary pressures also push prices upward, particularly in metropolitan regions [9]. The statement "People will move to regions where houses are built, but houses are not necessarily built in regions where people would want to live." also suggests that population is one of the contributing factors to influence housing prices [10]. Since the housing industry is labour-intensive, some aspects such as government restrictions on foreign workers and labour costs would also contribute to increasing housing prices as well [2].

Fluctuations in Malaysia's housing prices have raised concerns among potential homebuyers, developers and policymakers as the relatively slow growth in the population's annual income is completely out of sync with the high annual increases in property prices. Despite these concerns, there remains a limited amount of research that comprehensively examines the factors influencing housing prices in Malaysia, leaving the underlying factors insufficiently explored. Several economic and non-economic factors are cited as key contributing factors, yet their precise impact remains unclear [1].

In this research, we investigated the factors affecting housing prices in Malaysia using regression and correlation analysis. To achieve this purpose, secondary data was collected from various reputable sources such as the Department of Statistics Malaysia (DOSM) and the National Property Information Centre (NAPIC), covering the most recent five years since the latest data will capture the current trends and behaviour of the housing market. Accordingly, the research objectives are: (a) to identify the factors affecting housing prices in Malaysia, (b) to determine the strength and direction of the relationship between housing prices and the identified factors using correlation analysis, and (c) to develop a regression model for predicting housing prices based on significant influencing factors.

2. Research Methodology

2.1 Research Design

The research design employed in this study was a quantitative research design because it was appropriate for analysing numerical data and determining statistical relationships between variables. It relied on the analysis of secondary data in which the existing numerical data were used to investigate the housing prices in Malaysia and their influencing factors. With the use of quantitative methods, regression and correlation analysis were used to evaluate the strength and direction of relationships among variables and determine how various factors impacted housing prices.

2.2 Data Collection Methods and Source

Data for this study were obtained from multiple reputable, reliable and publicly accessible secondary sources including the Department of Statistics Malaysia (DOSM) and the National Property Information Centre (NAPIC). NAPIC provided the data for housing prices in Malaysia [11] while DOSM provided the data for Gross Domestic Product (GDP) [12], inflation rate [13], population [14] and labour force [15]. This study focused on data covering the period of the most recent five years, which is between 2020 and 2024, to observe the latest trends and changes in Malaysia's housing market.

2.3 Data Cleaning

Data cleaning is the process of removing or making amends to inaccurate, inconsistent, improperly formatted, incomplete or duplicate data in a dataset. When data are combined from several datasets of different data sources, a greater likelihood of duplicated records, missing values or incorrectly labelled data is possible. Hence, the data collected was cleaned before being subjected to the analysis process to ensure the accuracy, consistency and overall quality of data. This process involved the identification and handling of missing values and the detection of duplicates. In this study, data were collected from different sources and combined in Microsoft Excel. To check for missing values, each column was filtered using Excel's "Filter" function to identify empty cells. Duplicated data were identified using Excel's "Remove Duplicates" function and deleted to ensure data accuracy and consistency. No missing or duplicated data were found. Effective data cleaning is crucial in preventing misleading results and improving the overall quality and reliability of the analysis.

Enhanced Knowledge in Sciences and Technology Vol. 6 No. 1 (YEAR) p. 264-272

2.4 Data Analytical Tools

Primarily, Microsoft Excel was used for data recording. For regression and correlation analysis, Minitab was chosen since it is a specialised statistical software for analysis such as regression, chi-square, analysis of variance (ANOVA) and so on. In this study, Minitab had the ability to employ larger and complex datasets and provide detailed outputs including p-values, R-squared, ANOVA tables, as well as Variance Inflation Factor (VIF), which were essential for interpreting the results of regression and correlation analysis. Although Microsoft Excel can perform basic statistical analysis, it has limitations compared to Minitab, such as the inability to directly generate the VIF output for checking multicollinearity. However, the Microsoft Excel outputs were used for cross-checking purposes.

2.5 Data Analysis Techniques

The main techniques employed in this study were regression and correlation analysis, which assisted in identifying relationships between the dependent variable, housing price, and the independent variables, various influencing factors. In this study, a significance level of 5% was used to determine statistical significance. The 0.05 threshold is widely accepted in research largely due to the historical precedent set by Fisher and it balances the risk between Type I and Type II errors. Although it is acknowledged as an arbitrary choice, this conventional level of significance has been adopted in many fields as a standard [16].

2.5.1 Correlation Analysis

Correlation analysis examines the relationship between two or more quantitative variables, measuring the strength and direction of the relationships between dependent and independent variables. A change in one variable is normally accompanied by another variable experiencing an increase, decrease, or no changes at all.

The output of correlation analysis is called the correlation coefficient, denoted by r, and it ranges from -1 to +1. An r value of -1 reflects a perfect negative correlation, meaning that an increase in one variable results in a decrease of the other. Conversely, an r value of +1 reflects a perfect positive correlation, meaning that when one variable increases, the other increases as well. An r value of 0 suggests no linear relationship in which changes in one variable do not correlate with changes in another variable.

The r sign is used to show the direction of the relationship, where a positive sign (+) is used to show a direct relationship, that is, both variables move in the same direction, while a negative sign (−) is used to show an inverse relationship, that is, the variables move in opposite directions. The absolute value of r represents the strength of the relationship. An absolute value of the correlation coefficient closer to 1 implies a stronger relationship while a value closer to 0 implies a weaker relationship.

2.5.2 Regression Analysis

Regression analysis generally models the relationship between one or more response variables and one or more predictor variables. Response variables are also known as dependent variables while predictor variables are known as independent variables. In this study, the response variable was housing price, while the predictor variables included factors that affect housing prices.

Simple linear regression analysis was applied to examine the effect of each factor that affect housing prices individually. Each factor was analysed separately in different models to determine its strength, direction and statistical significance in explaining the variation in Malaysia's housing price. For each regression model, the output included regression statistics such as correlation coefficient, coefficient of determination, Significance F, p-value and so on. It also provided an Analysis of Variance (ANOVA) table and estimated regression coefficients. These models showed the independent relationship between each of the identified factors and Malaysian housing prices. The general form of the simple linear regression equation is as follows, where Y represents the dependent variable, X represents the independent variable, α represents the intercept, β represents the regression coefficient and ε represents the error term.

$$Y = \alpha + \beta X + \varepsilon \tag{1}$$

Besides, multiple linear regression analysis was also conducted to explore the overall combined effect of all identified factors collectively on housing prices in Malaysia. Unlike the simple linear regression analysis, the multiple linear regression analysis enables the modelling of the housing price against two or more factors which influence the housing prices simultaneously. This gives a better chance of understanding the interaction between the factors and their contribution to affecting Malaysia's housing prices. The output will consist of the coefficient of determination, Significance F, p-value and others. The general form of the multiple linear regression equation is as:

$$Y=\beta_0+\beta_1X_1+\beta_2X_2+\cdots+\beta_nX_n+\varepsilon \tag{2}$$

2.5.3 Coefficient of Determination (R^2)

The effectiveness of various factors contributed to the variation in housing prices in Malaysia was evaluated using the coefficient of determination, R^2. An R^2 can have a value of 0 to 1, with a value closer to 1 indicating a stronger explanatory power. When an R^2 is below 0.5, it implies a weak relationship between housing prices and factors affecting housing prices, whereas an R^2 between 0.5 and 0.8 indicate that the model is not sufficiently explaining the variability of the data. For example, an R^2 of 0.80 implies that the model can explain 80% of the variation in housing prices, while the remaining 20% being a result of other factors not included in the model. However, a larger R^2 value does not directly indicate that the factors have a statistically significant effect on housing prices. The significance of each factor is determined using p-values from regression analysis, where a significant effect is indicated by a p-value less than 0.05. Additionally, an R^2 can also be used to evaluate the goodness of fit of the regression model as well as determine the model's predictive power in this study [17]. The formula for the coefficient of determination is as follows, where SSE represents Sum of Squares of Error and SST represents Sum of Squares of Total:

$$R^2=1-\frac{SSE}{SST} \tag{3}$$

2.5.4 Analysis of Variance (ANOVA)

Analysis of Variance (ANOVA) is an alternative approach that can be used to test the significance of a regression model. The ANOVA table is a crucial part of the regression results since it provides a summary of the distribution of variation in the dependent variable into regression and error terms. Here, the $\hat{y}$ represents the predicted value of the dependent variable based on the estimated regression equation, while the $\bar{y}$ represents the mean of the dependent variable. The degrees of freedom, the sum of squares, the mean square and the F-statistic were provided in the table and Table 1 summarised the formula for the ANOVA table.

Table 1 ANOVA Table for Testing Significance of Regression Model

Factors	Degrees of Freedom	Sum of Squares	Mean Square	*F*-statistic
Regression	$df_R=k-1$	$SSR=\sum_{i=1}^{n}\left(\hat{y}-\bar{y}\right)^2$	$MSR=\frac{SSR}{k-1}$	$F=\frac{MSR}{MSE}$
Error	$df_E=N-k$	$SSE = SST-SSR$	$MSE=\frac{SSE}{N-k}$	
Total	$df_T=N-1$	$SST=\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2$		

Finally, the p-value as shown in the ANOVA table as Significance F, indicates whether the regression model is statistically significant or not. A p-value of less than 0.05 provides sufficient evidence to reject the null hypothesis, indicating that the overall regression model is statistically significant.

3. Results and Discussion

3.1 Results of Simple Linear Regression and Correlation Analysis

This section presents the summarised results of simple linear regression and correlation analysis, where each factor affecting housing prices was analysed separately to determine its individual relationship with housing prices in Malaysia. The results showed the strength, direction, and statistical significance of each influencing factor on housing prices, supported by fitted line plots that visually illustrated the relationship.

To determine the statistical significance of the regression model, hypothesis testing was performed at the chosen significance level of 0.05 in this study. The null hypothesis, H_0 stated that there is no linear relationship between housing prices and each factor affecting housing prices. H_0 will be rejected if Significance $F < 0.05$.

Table 2 Summary of Regression Statistics of Simple Linear Regression and Correlation Analysis

Factors	Multiple *R*	R^2	Standard Error	Significance *F*
GDP	0.75269	0.56654	97,927.14316	6.86030×10^{-15}
Inflation Rate	0.28416	0.08075	176,774.29080	0.01064
Population	0.12330	0.01520	182,967.72683	0.27587

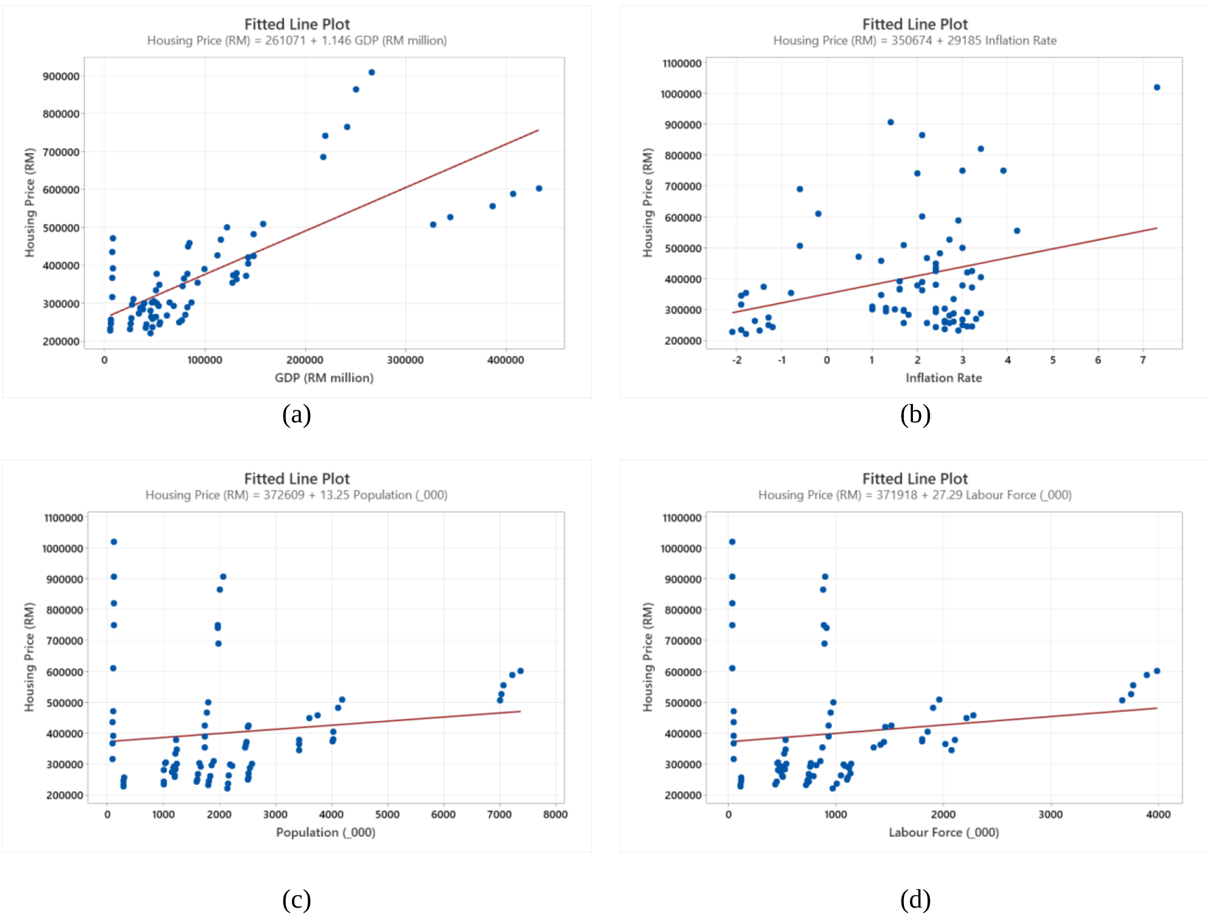
Labour Force	0.13848	0.01918	182,598.21313	0.22057

The regression statistics for the relationship between housing prices and each factor that affects housing prices are summarised in Table 2. Gross Domestic Product (GDP) exhibited the strongest positive relationship with housing price, with a sample correlation coefficient, Multiple *R* of 0.75269 and a coefficient of determination, R^2 of 0.56654, indicating that approximately 56.65% of the variation in housing prices was explained by GDP within this model. The standard error of 97,927.14316 indicated that, on average, the predicted housing prices deviated from the actual prices by approximately RM 97,927.14. The model was statistically significant with Significance *F* of 6.86030×10^{-15}. In comparison, the inflation rate showed a weaker positive relationship, followed by population and labour force. The model of inflation rate was statistically significant but both population and labour force were not statistically significant with Significance $F > 0.05$.

Table 3 Summary of Estimated Regression Coefficients

Factors	Intercept (RM)	Coefficient
GDP (RM million)	261,070.79653	1.14580
Inflation Rate (%)	350,673.58350	29,184.91725
Population ('000)	372,608.62790	13.24695
Labour Force ('000)	371,918.27783	27.29366

As shown in Table 3, the estimated regression coefficients for each factor affecting housing prices, were used to construct the regression equation. For example, the regression equation for GDP is $Y = 261{,}070.79653 + 1.14580x$. This implied that, for every additional RM 1 million increase in GDP, the housing price is expected to increase by approximately RM 1.14580 million, assuming other factors remain constant.



(a) (b)

(c) (d)

Fig. 1 Fitted Line Plot of Housing Price and (a) GDP; (b) Inflation Rate; (c) Population; (d) Labour Force

Fig. 1 displays the fitted line plots that visually examine the relationships between housing prices and each factor that affects housing prices, with the estimated regression line. Fig. 1(a) revealed a clear upward trend, indicating that housing prices generally rise as GDP increases. Fig. 1(b), Fig. 1(c) and Fig. 1(d) showed a slight upward trend, suggesting that housing prices rise slightly when the inflation rate, population and labour force increase.

3.2 Results of Multiple Linear Regression and Correlation Analysis

This section presents the results of multiple linear regression and correlation analysis, which investigated the combined effects of four factors on housing prices in Malaysia. Two regression models were examined.

3.2.1 Model 1: GDP, Inflation Rate, Population and Labour Force

Table 4 Model Summary of First Multiple Linear Regression Model

Regression Statistics	
R Square	0.72346
F-statistic	45.78092
Significance *F*	7.62228×10^{-19}
Standard Error	79,877.70771

The model summary of the first multiple linear regression model is displayed in Table 4. The R^2 value of 0.72346 indicated that approximately 72.35% of the variation in housing price was explained collectively by the four factors in this model. This implied that the model had a relatively strong explanatory power, indicated a good overall model fit. The *F*-statistic of 45.78092 with corresponding Significance *F* of 7.62228×10^{-19} indicated that the model was statistically significant overall at the 5% significance level. The standard error of 79,877.70771 suggested that the predicted housing prices deviated from the actual prices by an average of approximately RM 79,877.71.

Table 5 Estimated Regression Coefficients of First Multiple Linear Regression Model

Factors	Coefficient	Standard Error	*p*-value	VIF
Intercept	314,162.99549	18,543.16662	1.80765×10^{-26}	-
GDP (RM million)	2.06743	0.17723	4.48951×10^{-18}	3.43
Inflation Rate (%)	1,487.74808	5,619.06234	0.79197	1.03
Population ('000)	-89.17041	38.08156	0.02205	47.57
Labour Force ('000)	47.17916	66.85625	0.48273	44.10

The estimated regression coefficients of the first multiple linear regression model is displayed in Table 5. Gross Domestic Product (GDP) showed a positive and statistically significant coefficient of 2.06743 with a *p*-value of 4.48951×10^{-18}, indicated that for every additional RM 1 million increase in GDP, the housing price was expected to increase by approximately RM 2.07 million, assuming other factors remained constant. This positive relationship aligned with economic theory, which considers high GDP as high economic performance, purchasing power, and eventually, high housing prices.

The coefficient of 1,487.74808 of the inflation rates with a *p*-value of 0.79197 was positive but statistically insignificant, suggested that the variations in housing prices in this model could not be significantly explained by the inflation rate when other variables were controlled. The reason for this lack of significance could be the annual change in the inflation rate might not have cause a direct effect on the housing prices in the short-term.

Population had a negative coefficient of -89.17041 and was statistically significant at 5% level with a *p*-value of 0.02205. This implied that an increased population was related to a lower housing price when other factors were held constant. This unexpected finding may be caused by multicollinearity, where population had a strong correlation with other factors, especially the labour force. This was observed in its very high Variance Inflation Factor (VIF) value of 47.57, which was a sign of severe multicollinearity. The model did not clearly distinguish the individual effects of these variables when they increased or decreased at the same time. Consequently, the estimated coefficients might have become misleading.

Similarly, the labour force had a positive coefficient of 47.17916 but was statistically insignificant with a *p*-value of 0.48273. This suggested that the labour force was not significant in explaining the model when other variables were taken into consideration such as GDP, inflation, and population. This insignificance was likely due to its high correlation with population, as reflected by its extremely high VIF value of 44.10.

Overall, although Model 1 demonstrated good explanatory power, the extremely high VIF values for population and labour force indicated serious multicollinearity, which reduced the model's overall reliability and produces unstable estimated coefficients. To cope with this situation, another multiple linear regression model was estimated where the variable with the highest VIF value, population, was removed to improve model stability and reduce multicollinearity.

3.2.2 Model 2: GDP, Inflation Rate and Labour Force

Table 6 Model Summary of Second Multiple Linear Regression Model

Regression Statistics	
R Square	0.83773
F-statistic	55.69690
Significance *F*	1.26777×10^{-18}
Standard Error	82,360.84023

The model summary of the second multiple linear regression model is displayed in Table 6. The R^2 value of 0.83773 indicated that approximately 83.77% of the variation in housing prices was explained collectively by Gross Domestic Product (GDP), the inflation rate and the labour force. The *F*-statistic of 55.69690 with corresponding Significance *F* of 1.26777×10^{-18} indicated that the model was statistically significant overall at the 5% significance level. The standard error of 82,360.84023 suggested that the predicted housing prices deviated from the actual prices by an average of approximately RM 82,360.84.

Table 7 Estimated Regression Coefficients of Second Multiple Linear Regression Model

Factors	Coefficient	Standard Error	*p*-value	VIF
Intercept	293,232.60643	16,751.58258	1.82464×10^{-27}	-
GDP (RM million)	1.95250	0.17559	3.32287×10^{-17}	3.17
Inflation Rate (%)	2,137.45724	5,786.67239	0.71295	1.03
Labour Force ('000)	-103.71769	18.35475	3.09673×10^{-7}	3.13

The estimated regression coefficients along with the standard error, *p*-values and Variance Inflation Factor (VIF) for the second multiple linear regression model is displayed in Table 7. The coefficient of GDP was 1.95250, indicated that for every additional RM 1 million increase in GDP, the housing price was expected to increase by approximately RM 1.95 million, assuming other factors remained constant. This coefficient was highly statistically significant, with a *p*-value of 3.32287×10^{-17}, confirming that GDP remained the strongest predictor of housing prices in Malaysia.

The inflation rate had a coefficient of 2,137.45724 with a *p*-value of 0.71295. This suggested that a 1% increase in the inflation rate was associated with an approximately RM 2,137.46 increase in housing price, holding other variables constant. Although the coefficient was positive, it was statistically insignificant. This lack of significance was consistent with the first model.

In this model, the labour force exhibited a negative coefficient of -103.71769 with a statistically significant *p*-value of 3.09673×10^{-7}. This suggested that, holding GDP and inflation rate constant, an increase in labour force was associated with a decrease in housing price. This outcome was opposite from the first model and this might have been due to the population, which is a highly correlated variable was eliminated.

After removing the variable with the highest VIF value, all VIF values of independent variables in the second model were lower than in the first model. VIF of the labour force dropped from 44.10 to 3.13. This demonstrated that the population had been eliminated out of the model so that multicollinearity was successfully reduced, resulting in a more accurate estimation of the coefficients.

In summary, the second multiple linear regression model showed that GDP remained the strongest and significant predictor of housing prices, while the inflation rate was insignificant. The labour force shifted from an insignificant to a significant predictor, showed a negative relationship with housing prices. There was probably a reflection of its independent effect after multicollinearity was taken into consideration. Overall, the removal of population successfully lowered VIF values and made the regression coefficients more interpretable. Consequently, this improved the second model's stability, making it a better representation of the factors affecting housing prices in Malaysia.

4. Conclusion

The results of simple linear regression and correlation analysis showed that Gross Domestic Product (GDP), inflation rate, population and labour force have positive relationships with housing prices. This indicated that when these variables increase, housing prices in Malaysia also tend to increase. Among these variables, GDP showed the strongest positive relationship with housing price, with a correlation coefficient of 0.75269. This was followed by the inflation rate with a correlation coefficient of 0.28416, the labour force with a correlation coefficient of 0.13848 and the population with a correlation coefficient of 0.12330, which showed a weak positive relationship with housing prices when compared to GDP. On the other hand, the results highlighted that only GDP and the inflation rate were statistically significant in explaining changes in housing prices, with Significance *F* values of 6.86030×10^{-15} and 0.01064, respectively. However, the population and labour force were not statistically significant with Significance *F* values of 0.27587 and 0.22057, respectively.

Next, the results of the multiple linear regression and correlation analysis showed that in the first model, all variables showed a positive relationship with housing price except for the population. However, only GDP and the population were statistically significant, while the inflation rate and the labour force were not statistically significant. The first model has a R^2 value of 0.72346, suggesting that the variables in this model explain approximately 72.35% of the variation in housing prices, and the overall model was statistically significant. In the second model, the population variable was removed due to its extremely high Variance Inflation Factors (VIF) of 47.57 to reduce multicollinearity and improve model stability. The remaining variables, GDP and inflation rate showed a positive relationship with housing prices while the labour force showed a negative relationship. In this model, GDP and labour force were statistically significant, while the inflation rate was not. The model's explanatory power improved with an R^2 of 0.83773 and the overall model was statistically significant.

Overall, the findings indicated that housing prices in Malaysia were largely influenced by the economic factor, particularly the GDP. This showed that Malaysia's housing market is highly tied to the country's overall economic performance, especially changes in GDP. Hence, we may come to the conclusion that a more balanced housing market may be supported with the help of a stable GDP. Through these relationships, government, policymakers, investors and potential homeowners are able to make informed decisions regarding the housing market in Malaysia.

Acknowledgement

The authors would like to thank the Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia for its support.

Conflict of Interest

The authors declare that there is no conflict of interests regarding the publication of the paper.

Author Contribution

The authors confirm contribution to the paper as follows: ***study conception and design:*** *Kyau Chin Wen, Cik Sri Mazzura Muhammad Basri;* ***data collection:*** *Kyau Chin Wen;* ***analysis and interpretation of results:*** *Kyau Chin Wen, Cik Sri Mazzura Muhammad Basri;* ***draft manuscript preparation:*** *Kyau Chin Wen, Cik Sri Mazzura Muhammad Basri. All authors reviewed the results and approved the final version of the manuscript.*

References

[1] T. S. Ong, "Factors affecting the price of housing in Malaysia," *Journal of Emerging Issues in Economics, Finance and Banking*, vol. 1, no. 5, pp. 414–429, 2013.

[2] A. L. Olanrewaju, X. Y. Lim, S. Y. Tan, J. E. Lee, and H. Adnan, "Factors affecting housing prices in Malaysia: Analysis of the supply side," *Planning Malaysia Journal*, vol. 16, no. 2, pp. 225–235, 2018, https://doi.org/10.21837/pm.v16i6.477

[3] W. Cox and H. Pavletich, *13th Annual Demographia International Housing Affordability Survey: 2017*. [Online]. Retrieved from https://www.demographia.com/dhi.pdf

[4] Department of Statistics Malaysia, *Report on Characteristics of Household 2010: Population and Housing Census of Malaysia*. Putrajaya, Malaysia, 2014. [Online]. Available: https://www.dosm.gov.my

[5] Department of Statistics Malaysia, *Current Population Estimates, Malaysia, 2016–2017*. Putrajaya, Malaysia, 2017. [Online]. Available: https://www.dosm.gov.my

[6] E. J. Levin and R. E. Wright, "The impact of speculation on house prices in the United Kingdom," *Economic Modelling*, vol. 14, no. 4, pp. 567–585, 1997, https://doi.org/10.1016/S0264 9993(97)00008-4

[7] J. H. W. Huu, I. AbdLatif, and A. M. Nasir, "Lead-lag relationship between housing and gross domestic product in Sarawak," presented at the International Real Estate Society Conf., Kuala Lumpur, Malaysia, Jul. 26–30, 1999. [Online]. Available:

https://www.prres.org/uploads/460/1018/Hii_Lead_Lag_Relationship_Between_Housing_And_GDP_In_Sarawak.pdf

[8] K. Tsatsaronis and H. Zhu, "What drives housing price dynamics: Cross-country evidence," *BIS Quarterly Review*, Mar. 2004. [Online]. Available: https://www.bis.org/publ/qtrpdf/r_qt0403f.pdf

[9] T. H. Tan, "Housing satisfaction in medium- and high-cost housing: The case of Greater Kuala Lumpur, Malaysia," *Habitat International*, vol. 36, no. 1, pp. 108–116, 2012, https://doi.org/10.1016/j.habitatint.2011.06.003

[10] W. Vermeulen and J.V. Ommeren, "Housing supply and the interaction of regional population and employment," *CPB Discussion Paper No. 65*, CPB Netherlands Bureau for Economic Policy Analysis, 2006. [Online]. Available: https://ideas.repec.org/p/cpb/discus/65.html

[11] National Property Information Centre (NAPIC), *"Residential prices update by state in Malaysia (2020–2024),"* Nov. 18, 2025. [Online]. Available: https://napic.jpph.gov.my/ms/visualisasi-data?category=15&id=235

[12] Department of Statistics Malaysia (DOSM), *"Gross domestic product (GDP) by state in Malaysia (2020–2024),"* 2025. [Online]. Available: https://www.dosm.gov.my/portal-main/release-content/gross-domestic-product-gdp-by-state-2024

[13] Department of Statistics Malaysia (DOSM), *"Inflation: Consumer price index (CPI) – 2024,"* 2025. [Online]. Available: https://storage.dosm.gov.my/cpi/cpi_annual_2024.pdf

[14] Department of Statistics Malaysia (DOSM), *"Population by state,"* 2025. [Online]. Available: https://open.dosm.gov.my/data-catalogue/population_state

[15] Department of Statistics Malaysia (DOSM), *"Labour force by state,"* 2025. [Online]. Available: https://open.dosm.gov.my/data-catalogue/lfs_qtr_state

[16] L. V. Manderscheid, "Significance levels—0.05, 0.01, or?" *Journal of the American Statistical Association*, vol. 47, no. 5, pp.1381–1385, 1964, https://doi.org/10.2307/1236396

[17] A. Di Bucchianico, "Coefficient of determination (R^2)," *Wiley StatsRef: Statistics Reference Online*, John Wiley & Sons, 2014, https://doi.org/10.1002/9780470061572.eqr173