e-ISBN : 978-629-98755-5-0
SIG : e-Learning@CS
https://appspenang.uitm.edu.my/sigcs/
Publication Date : 24 – Mar - 2025

A MULTIPLE LINEAR REGRESSION APPROACH TO FORECASTING MALAYSIA'S GDP USING MACROECONOMIC VARIABLES

Siti Nurhanani Shamsuddin[1], Siti Nur Aishah Razali[2], *Mahanim Omar[3] and Siti Nurleena Abu Mansor[4]
2021494756@student.uitm.edu.my [1], 2021886048@student.uitm.edu.my [2], mahanim@uitm.edu.my[3], sitin140@ uitm.edu.my [4]

[1,2]College of Computing, Informatics and Mathematics,
Universiti Teknologi MARA Cawangan Negeri Sembilan, Malaysia

[3,4] Jabatan Sains Komputer dan Matematik (JSKM),
Universiti Teknologi MARA Cawangan Pulau Pinang, Malaysia

**Corresponding author*

ABSTRACT

This study examines the impact of macroeconomic variables on Malaysia's Gross Domestic Product (GDP) using the Multiple Linear Regression (MLR) method and develops a predictive model through multiple linear regression. Due to data availability constraints, this study utilizes secondary data from the World Bank and World Development Indicators, covering the period from 2010 to 2019. Based on a review of previous literature, six macroeconomic variables were selected: inflation, exports, imports, Foreign Direct Investment (FDI), population growth, and unemployment rate. Using EViews 12 Student Lite software, the findings reveal that only exports have a statistically significant positive impact on GDP growth, while inflation, imports, FDI, population growth, and unemployment rates are found to be insignificant. These results provide valuable insights for policymakers in formulating strategies to enhance economic growth, identifying key economic trends, and mitigating potential risks. Additionally, the findings contribute to improved decision-making for firms, investors, and government agencies by supporting risk management and strategic economic planning.

Keywords: GDP, Multiple Linear Regression, Macroeconomic Variables, World Development Indicators

Introduction

Gross Domestic Product (GDP) measures the monetary value of goods and services produced within a country over a given period, such as a quarter or a year (OECD, 2009). It includes both market-driven and non-market output, such as government services. GDP is widely used to assess economic size, growth, and performance, but it has limitations, such as not accounting for environmental sustainability or wealth distribution (OECD, 2024). Alternative indices are thus needed for a more comprehensive evaluation of economic progress.

In Asia, GDP growth remains a key driver of global economic expansion, but its pace is slowing. The IMF estimates that in 2023, Asia contributed nearly two-thirds of global growth. However, while GDP in the Asia-Pacific region grew from 4.2% in 2022 to 4.6% in 2023, concerns remain over declining investment in Southeast Asia and Japan, as well as China's real estate slump. Malaysia's GDP outperformed expectations in Q1 2023 with a 5.6% growth rate but slowed to 2.9% in Q2 before

improving slightly to 3.3% in Q3, reflecting global economic challenges and declining external demand (Goh, 2023).

Multiple linear regression is a statistical method used to assess the impact of multiple independent variables on a single dependent variable. This approach estimates coefficients that minimize prediction errors, commonly using the Ordinary Least Squares (OLS) method, which minimizes the total squared differences between actual and predicted values (Gujarati & Porter, 2009; Pohlmann & Leitner, 2003).

In this study, macroeconomic indicators such as population growth, FDI, unemployment rate, inflation rate, imports, and exports are used as independent variables (Agalega & Antwi, 2013). Population growth is measured as the exponential rate of midyear population increase. FDI refers to net inflows aimed at acquiring a long-term managerial stake in a foreign enterprise. The unemployment rate represents the percentage of the labour force actively seeking jobs. Inflation is calculated using the consumer price index (CPI) as an annual percentage change in consumer costs. Exports and imports denote the total value of goods and services traded internationally (World Bank, 2024).

As a rapidly growing economy, Malaysia faces ongoing challenges in accurately predicting GDP to guide policy decisions, strategic planning, and sustainable economic growth. The interplay of various macroeconomic factors complicates economic forecasting, highlighting the need for a robust predictive model. While previous studies have attempted to forecast GDP, a more comprehensive approach incorporating multiple macroeconomic variables is necessary for improving accuracy (Omar & Nor, 2020). For instance, studies on Pakistan's GDP by Irshad et al. (2022) and Memon et al. (2021) yielded differing results regarding the impact of exports. This discrepancy arose due to variations in methodology—both studies conducted the Augmented Dickey-Fuller (ADF) test, but only Irshad et al. (2022) performed additional multicollinearity, heteroscedasticity, and autocorrelation tests to refine their model. Such methodological differences can lead to conflicting conclusions, making it difficult for firms and policymakers to evaluate risks and predict economic trends accurately.

Given these challenges, this study proposes a model that identifies key factors influencing Malaysia's GDP using a more appropriate econometric approach. By improving GDP prediction accuracy, this model aims to assist policymakers in making informed decisions that support long-term economic stability and growth.

Literature Review

Stani´c and Ra˘ci´c (2019) used the multiple linear regression model to investigate the macroeconomic variables influencing Bosnia and Herzegovina's GDP from 2005 until 2018. From the study, growth rate, unemployment rate, and inflation rate are shown to have a negative association with GDP, whereas imports, exports, and FDI are shown to have a positive relationship. In other study investigating whether

US' GDP can be predicted annually using multiple linear regression using macroeconomic variables, three models have been proposed (Samiyu, 2021). The first one is called the basic model. In this model, population, maturity rate, disposable income, and housing price index are positively associated to GDP. Next, in the second model, corporate profit has been included. This is called model 1. The result shows that it has a positive and significant correlation with GDP. In the third model, 10-year breakeven inflation rate was added as an extended to the model 1. However, the result revealed that inflation rate has a negative correlation with GDP.

Omar and Nor (2020) study the linkage between macroeconomic variables (population, unemployment and export) with the economic growth in Malaysia based on time series quarterly data from 2006 until 2016 using multiple linear regressions (MLR) analysis. The study revealed that population is significant and negatively related to economic growth, while export is significant but positively related to economic growth. However, this study stated that the unemployment is insignificant towards economic growth. In a similar study to examine the impact of macroeconomic variables on the expansion of Pakistan's economy from 1991 to 2020, the results of the Ordinary Least Square (OLS) method show that, while the inflation rate has a negative association with GDP, household consumption, government spending, the investment rate, and net exports have positive relationships (Memon et al., 2021).

Using four machine learning techniques — principal component regression (PCR), ridge regression (RR), lasso regression (LR), and OLS — Agu et al. (2022) conducted a study to predict Nigeria's GDP based on macroeconomic indicators and identify the most likely major macroeconomic variables that could affect GDP growth. The study's data set spans from 1981 to 2019. The first approach, PCR, is a PCA-based regression analysis methodology. However, PCA cannot be used to forecast GDP since it does not offer an economic interpretation of the data. Other than that, the RR, LR, and OLS approaches result in the same outcome: population, federal government expenditure, import rate, and exports have a positive connection with GDP, whereas the foreign exchange rate, FDI, and oil revenue have a negative link with GDP. When comparing MSE, PCR outperformed OLS, LR, and RR in terms of predictive accuracy. However, since PCA does not offer economic interpretation, hence PCR unable to impose any relationship between dependant and independent variables. Due to that issue, RR model is utilized for predicting accuracy since it has second highest predictive accuracy.

Methodology

Data of GDP and other macroeconomic variables from 2010 to 2019 were retrieved from World Development Indicator (WDI), The World Bank (2024). Regarding the macroeconomic variables, the dependent variable GDP would be predicted using the following independent variables; population growth (POP), imports (IMP), exports (EXP), unemployment rate (UNEMP), inflation rate (INF), and

foreign direct investment (FDI). Next, all the data are checked for stationarity test. It is a method used to identify whether a time series statistical properties hold true over time. In this study, Augmented Dickey-Fuller (ADF) test is used to determine whether the data is stationary or non-stationary. All the analysis were executed using EViews 12 Student Lite software.

Multicollinearity in a multiple regression model is the presence of strong intercorrelations between two or more independent variables. If there is multicollinearity in the data, the common solution would be to identify and remove those variables from the regression (Tan et al., 2021). This problem can be tested using Variance Inflation Factor (VIF). The degree to which multicollinearity increases the variances in the regression estimates is shown by the linear regression's VIFs. A VIF value of 5 to 10 will mean that variables are highly correlated, a VIF between 1 and 5 indicates that variables are moderately correlated, and a VIF of 1 show that the variables are not correlated (Statistics, 2013).

Multiple linear regression (MLR) is used to explain the impact of GDP on the dependent variable. MLR is a linear regression model in which only one dependent variable (Y) and two or more independent variables (X) are regressed. The purpose of using this model is to create a linear equation that best fits the data by estimating the coefficients of the independent variables that minimize the error in predicting the dependent variable (Gujarati & Porter, 2009). The general form of multiple linear regression equation is:

$$GDP_t = B_0 + B_1POP_t + B_2IMP_t + B_3EXP_t + B_4UNEMP_t + B_5INF_t + B_6FDI_t + \xi_t$$

Where;

GDP_t = the dependent variable, GDP, at time t .

B_0 = the intercept, representing the value of GDP when all the independent variables are zero.

$B_1, B_2, ..., B_n$ = the coefficients for the independent variables representing how much the dependent variable changes for a one-unit change in each respective independent variable, while holding other variables constant.

ξ = the error term, representing the difference between the predicted and actual values of the dependent variable.

Result and Discussion

The study examined the relationship between GDP and macroeconomic variables; inflation, exports, imports, FDI, population growth, and unemployment using multiple linear regression. After conducting Multicollinearity Tests, no multicollinearity was detected, and the data was homoscedastic.

Table 1: Variance Inflation Factor (VIF) result

Variables	VIF values
Inflation	2.096297
Export	2.230797
Import	1.902861
FDI	1.230503
Population growth	1.933110
Unemployment	1.995800

The result for F-statistics for heteroscedasticity test is 1.508703 with p-value of 0.3955 indicating that the data is homoscedasticity. Initially, all six independent variables were included, but only variable exports (EXP) were found to be statistically significant in influencing GDP.

Table 2: Multiple Linear Regression result

Variables	Coefficient	p-values
Intercept	4.982692	0.0032
Inflation	-0.074520	0.8030
Export	0.228831	0.0472*
Import	-0.000685	0.9860
FDI	-0.167898	0.4640
Population growth	3.969650	0.5438
Unemployment	1.065465	0.3628

From the result, only export data is significant variables and F-statistic shown to have value of 3.2080 with p-value of 0.1833. With p-value (F-statistic) more than 0.05, the whole regression proves to be insignificant. Hence, to achieve model that is significant, insignificant variables need to be excluded starting from highest p-value. After few steps execution in finding the good model, finally, only export variable is left with statistically significant p-value (F-statistics=15.23762, p-value=0.004522). Other variables were excluded due to high p-values. Therefore, the final regression model obtained was:

$$GDP = 4.824162 + 0.179571EXP$$

This indicates that a 1% increase in exports is associated with a 0.18% increase in GDP. The model was statistically significant, aligning with previous studies that support the positive relationship between exports and GDP. However, the study's small sample size (10 years) may have contributed to the insignificance of other variables.

Overall, the findings highlight the strong impact of exports on GDP, while other factors like inflation, FDI, and population growth were not significant in this dataset. These results align with previous studies, such as Omar and Nor (2020), which also found exports to be a key driver of Malaysia's GDP. The insignificance of unemployment and FDI is consistent with findings by Assaf (2014) and Memon

et al. (2021), suggesting that these factors may have a weaker direct influence on economic growth within the selected time frame. Moreover, in the study on Nigeria's GDP of data set spans from 1981 to 2019 has identified exports as a major factor positively affecting GDP, while FDI had a negative relationship (Agu et al., 2022).

Conclusion

This study aimed to examine the impact of macroeconomic variables (inflation, exports, imports, FDI, population growth, and unemployment) on GDP using a multiple linear regression model. The analysis covered data from [years] to determine which factors significantly influence economic growth. The regression results revealed that among the six independent variables, only exports (EXP) had a significant positive relationship with GDP. A 1% increase in exports was associated with a 0.18% increase in GDP, indicating a strong linkage between international trade and economic growth. In contrast, inflation, imports, FDI, population growth, and unemployment were found to be statistically insignificant in this study.

The findings underscore the crucial role of exports in driving economic growth. Policymakers should focus on enhancing trade policies, strengthening export-oriented industries, and expanding international market access to sustain GDP growth. Meanwhile, further investigation into the role of other macroeconomic factors may be needed to understand their long-term impact on economic performance. This study is limited by its relatively small sample size of 10 years, which may have influenced the statistical significance of certain variables. Future research could extend the dataset to include a longer period or explore additional macroeconomic indicators to provide a more comprehensive analysis. Furthermore, alternative econometric techniques, such as time-series modelling, could be applied to enhance the robustness of the findings.

Overall, this study highlights exports as a vital contributor to GDP growth. As globalization continues to shape economic trends, further research into trade dynamics and macroeconomic policies will be essential in ensuring sustainable economic development.

References:

Agalega, E., & Antwi, S. (2013). *The impact of macroeconomic variables on gross domestic product: Empirical evidence from Ghana*. International Business Research, 6(5), 108–116.

Agu, S. C., Onu, F. U., Ezemagu, U. K., & Oden, D. (2022). *Predicting Gross Domestic Product to macroeconomic indicators*. Intelligent Systems with Applications, 14, 200082.

Goh, N. (2023). *Malaysia's GDP growth inches up to 3.3% but remains lacklustre* [Accessed: 2024-04-20]. https://asia.nikkei.com/Economy/Malaysia-s-GDP-growth-inchesup-to%203.3-but-remains-lackluster

Gujarati, D. N., & Porter, D. C. (2009). Basic econometrics. McGraw-hill.

Irshad, M., Hussain, M., & Baig, M. A. (2022). *Macroeconomic variables the indicators for the economic growth of Pakistan*. Pakistan Social Sciences Review, 6(2), 58–72.

Memon, A. W., Phul, M. H., & Kanhar, N. A. (2021). *The impact of macroeconomic variables on economic growth of Pakistan*. Turkish Online Journal of Qualitative Inquiry, 12(8), 7802–7830.

OECD. (2009). National accounts of OECD countries 2009, volume 1, main aggregates. https://doi.org/10.1787/na vol 1-2009-en-fr

OECD. (2024). Gross Domestic Product (GDP) (indicator) [Accessed: 2024-05-26]. https://doi.org/10.1787/dc2f7aec-en

Omar, M. S., & Nor, A. M. (2020). *Macroeconomic variables and economic growth: The Malaysian perspectives*. International Journal of Academic Research in Business and Social Sciences, 10(3), 272–280.

Pohlmann, J. T., & Leitner, D. W. (2003). *A comparison of Ordinary Least Squares and Logistic Regression (1)*. The Ohio journal of science, 103(5), 118–126.

Samiyu, M. (2021). *Multiple regression model for predicting GDP using macroeconomic variables (part 1)*. Available at SSRN 3895177.

Stani´c, S., & Ra˘ci´c, ˘ Z. V. (2019). *Analysis of macroeconomic factors effect to Gross Domestic Product of Bosnia and Herzegovina using the multiple linear regression model.* Economics, 7(2), 91–97.

Statistics, L. (2013). Multiple regression analysis using SPSS statistics. Laerd Research Ltd.

Tan, V. N., Md Yusof, Z., Misiran, M., & Supadi, S. S. (2021). *Assessing youth unemployment rate in malaysia using multiple linear regression*. Journal of Mathematics and Computing Science (JMCS), 7(1), 23–34.

The World Bank. (2024). World development indicators [Accessed: 2024-01-08]. https://databank.worldbank.org/source/world-development-indicators