UJM 12 (1) 2023
UNNES Journal of Mathematics
http://journal.unnes.ac.id/sju/index.php/ujm

Prediction of final grade in linear algebra course with multiple linear regression approach

Andika Ellena Saufika Hakim Maharani*, Siti Soraya, Gilang Primajati , Habib Ratu Perwira Negara, Ahmad

Universitas Bumigora, Indonesia
Ismail Marzuki Street No.22, Cilinaya, Cakranegara, Mataram, 83127, Indonesia

Article Info

Article History:
Recived November 2022
Accepted March 2023
Published May 2023

Keywords:
Final Grade Analysis,
Multiple Linear Regression,
Attendance,
Daily Grade.

Abstract

This article examines the analysis of the final grade of Linear Algebra course with multiple linear regression approach. The research was conducted by collecting attendance data, daily grades, and final grades from students of the Bumigora University Computer Science Study Program who took Linear Algebra courses in the odd semester 2022/2023. The collected data were analyzed using multiple linear regression techniques. The purpose of this study is to determine the relationship between variables that have a significant effect on students' final grades and how to predict these variables using multiple linear regression models. The results of the analysis show that both independent variables, namely attendance and daily grades, have a significant effect on the dependent variable, namely the final grade of students, with a significance value smaller than 0.05. The resulting multiple linear regression model can also be used to predict students' final grades with an accuracy rate of 70.4%. In addition, the results of this analysis also show that daily grades have a greater influence than attendance in predicting final grades. The results of this study can provide useful information for lecturers in improving teaching and for students to improve their performance in the course..

How to cite:

Maharani, A.E.S.H., et al. 2023. Prediction of final grade in linear algebra course with multiple linear regression approach. *UNNES Journal of Mathematics*. 12(1): 1-11.

*Correspondence Address
E-mail: a.ellena.saufika@universitasbumigora.ac.id

p-ISSN 2252-6943
e-ISSN 2460-5859

1. Introduction

Linear Algebra course is one of the courses that must be taken by students of Computer Science major. This course has a varying level of difficulty depending on the curriculum and the basic ability of students in understanding linear algebra. For students, the final grade of each course taken is very important in addition to being proof of learning outcomes, it can also be a guideline in selecting other courses in the following semester (Zharbaini et al., 2023). The final grade of a course is also one of the bases for calculating the Grade Point Average (GPA) (Kurnia, 2014) (Prasetyo & Hasyim, 2022). Students with high final course grades assume that students have a good understanding of the course (Aswan & Fadhillah, 2022). To obtain a good understanding of this course, students are required to attend class and complete daily assignments. In the teaching process, lecturers need to know the factors that influence a student to get a high final grade in this course. These factors can be in the form of student activeness in class as shown by the number of attendance and the acquisition of daily grades. By knowing these factors, lecturers can improve teaching and provide appropriate guidance to students.

Several studies have shown that attendance variables and daily grades can be used to predict final course grades. One of the studies conducted by (Pamungkas & Mustafidah, 2016) (Muktiadi & Mustafidah, 2012) (Wibowo et al., 2013) found that there is an influence of learning interest and student attendance level on learning achievement in students. In their research, each of them showed that students with more than 9, 10, and 11 attendance tended to have higher grades than students who attended less. Research conducted by (Fahmi & Hidayat, 2014) found that with an increase or decrease in the value of daily test scores, assignment scores, mid-semester test scores, and final semester test scores will be followed by an increase and decrease in final grades. Similarly, research conducted by (Adna, 2016) shows that student learning outcomes are an accumulation of several aspects of assessment, including: task observations, assignments, midterm exams, and semester final exams.

The use of statistics when processing research materials affects the level of analysis of research results. Regression analysis is one of the most popular and widely used analyses. In statistics, linear regression is an approach to modeling the causal relationship between (one or more) dependent variables and (one or more) independent variables (Mardiatmoko, 2020). In other words, this analysis is used to understand which independent variables are related to the dependent variable and find the shape of the relationship. When there is only one independent variable and one dependent variable, the regression is called simple linear regression. If there is more than one independent variable, the regression test is called multiple linear regression.

Multiple linear regression is an equation model that explains the relationship of one independent/response variable/response (Y) with two or more independent/predictor variables $(X_1, X_2, \ldots, X_n)$. The purpose of the multiple linear regression test is to predict the value of the independent variable/response (Y) with two or more independent variables/predictors $(X_1, X_2, \ldots, X_n)$ known (Wasilaine et al., 2014). The multiple linear regression equation is mathematically expressed by:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n \qquad (1)$$

where:
Y = independent variable (the value of the variable to be predicted),
α = constant,
$\beta_1, \beta_2, \ldots, \beta_n$ = regression coefficient value,
$X_1, X_2, \ldots, X_n$ = independent variable.
Conditions when the regression coefficients, namely $\beta_1, \beta_2, \ldots, \beta_n$ have a value:

- Value = 0. In this case, the variable Y is not influenced by $X_1, X_2, \ldots, X_n$,
- Value < 0. here is a relationship in the inverse direction between the independent variable Y and the variables $X_1, X_2, \ldots, X_n$,
- Value > 0. Here there is a relationship in the same direction between the independent variable Y and the variables $X_1, X_2, \ldots, X_n$.

Several studies related to multiple regression have been conducted, among others, by (Magfirah et al., 2015) to see the effect of self-concept and learning habits on math learning outcomes. Related research was also conducted by (Syam et al., 2016). Furthermore, (Kurnia, 2014) related to the effect of organizational activeness on the cumulative grade point average of students. In addition, there is also research conducted by (Wanti Wulan Sari et al., 2019) to determine the factors that influence GPA. However, in collecting data on students' final grades, there are several obstacles that may occur, such as variability in the grading system, student absences in class, and lack of accurate and complete data. Therefore, an accurate and precise analysis is needed to determine the factors that influence students' final grades in Linear Algebra courses. A multiple linear regression approach can be used to analyze the variables that affect students' final grades in Linear Algebra courses. This approach takes into account the influence of several interrelated variables, making it more accurate in analyzing the factors that affect the final grade.

Therefore, this article aims to analyze the final grade of the Linear Algebra course using a multiple linear regression approach to determine the relationship between the variables that affect students' final grades and how these variables can be predicted using multiple linear regression models. Thus, this article is expected to provide useful information for lecturers and students in improving the quality of teaching and learning.

2. Method

1. Research Method

Analysis of final grades in linear algebra courses with a regression approach can provide useful information in understanding the factors that affect student academic achievement in the course. The regression approach can be used to analyze the relationship between the dependent variable (final grade) and one or more independent variables (factors that affect the final grade).

The steps that can be taken in analyzing the final grade of linear algebra courses with a regression approach include (Mardiatmoko, 2020):

1. Identify independent variables: Factors that can affect the final grade in linear algebra courses can vary, such as the number of student attendance and daily grades based on the accumulation of class activity scores, assignment scores, and quiz scores. These independent variables are selected based on their relevance to the linear algebra course and available data.
2. Data collection: Data on final grades and independent variables were collected from students of the computer science study program at Bumigora University who had taken linear algebra courses in one odd semester of the 2022/2023 academic year.
3. Descriptive analysis: Descriptive analysis was conducted to determine the basic statistics of the dependent and independent variables, such as data range, mean, standard deviation, and variance.
4. Multiple Linear Regression Analysis: Multiple linear regression analysis is conducted to determine the effect of the selected variables on the final grade. In this analysis, the significant variables will be the independent variables and the final score will be the dependent variable. Regression analysis must be carried out by paying attention to classical assumptions, namely the assumptions of normality, multicollinearity, and heteroscedasticity. In addition, regression hypothesis testing is also carried out, namely the F-test (Simultaneous ANOVA) and t-test (Partial). The analysis conducted using SPSS version 22 software.
5. Model Evaluation: The regression model should be evaluated to determine how well it explains the variation in the final values. Evaluation can be done by calculating the standardized coefficients beta and R2 values. The standardized beta coefficients for each independent variable are used to compare which independent variable has the greatest contribution to the dependent variable of final value. Meanwhile, the R2 value is used to show the percentage of variation in final value that can be explained by the independent variables. A significance test of the regression coefficients also needs to be conducted to ensure that the model is statistically significant.
6. Interpretation of Results: The results of the regression analysis can be used to interpret the factors affecting final grades in linear algebra courses. Interpretation of multiple linear regression results is done by looking at the regression coefficient value of each independent variable. A positive regression coefficient indicates that the greater the value of the independent variable, the greater the final grade. Conversely, a negative regression coefficient indicates that the greater the value of the independent variable, the smaller the final grade.

By conducting this analysis, it is hoped that significant variables can be found and affect the final grade of the Linear Algebra course, so that it can help decision making in improving the quality of learning and improving student performance in the course.

2. Data Description

The data used in this study are based on the grades of linear algebra courses of computer science study program students in the odd semester of the 2022/2023 academic year at Bumigora University. The data consists of Student Identification Number/*Nomor Induk Mahasiswa* (NIM), student name, number of attendance, daily grades, and final grade. There are 223 students in the computer science study program who take linear algebra courses in the odd semester of the 2022/2023 academic year. In this data, the final grade is influenced by 2 factors, namely the number of attendances and daily grades obtained based on the accumulation of activeness scores, assignment scores, and quizzes in class. Taking into account the existence of 2 predictor variables that are thought to affect the response variable in this research data, multiple linear regression analysis is used to model the relationship between the final grade and the factors that are thought to cause variation in the achievement of the final grade.

3. Variables in the Research

This study uses data on the final grade of linear algebra courses and the factors that influence it. The data is then analyzed using multiple linear regression methods. The final grade of students as the dependent variable (Y) with a value range of 0 to 100, while the independent variables include:

- Total attendance as variable X_1 with a range of numbers between 0 and 16
- Daily grades as variable X_2 with a range of values from 0 to 96.

Both dependent and independent variables are ratio scaled. The following Table 1 provides a statistical description for the research data using SPSS software version 22.

Table 1. Data Statistics Description Table

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic
Y	223	100	0	100	66.61	1.327	19.815	392.618
X1	223	16	0	16	12.66	.260	3.882	15.071
X2	223	96	0	96	64.13	1.395	20.839	434.254
Valid N (listwise)	223							

3. Results and Discussions

3.1. Classical Assumption Test

The classical assumption test is a process to check whether the multiple linear regression model meets the basic assumptions of regression analysis, such as the assumptions of normality, multicollinearity, and heteroscedasticity.

3.1.1. Normality Test

The normality test aims to test whether the residual variables in the regression model are normally distributed (Fahmeyzan et al., 2018). As is known, the t-test and F-test assume that the residuals follow a normal distribution. If the assumption is violated, the small sample statistical test is invalid. In this study, the normality test used a normal probability plot (P-P plot) and the Kolmogorov-Smirnov test.

3.1.1.1. Normal P-P Plot Test

The normality test is carried out by looking at the normal P-P plot which compares the cumulative distribution with the normal distribution (Permana & Ikasari, 2023). The normal distribution forms a straight diagonal line, and the residual data plot is compared to the diagonal line. If the distribution of the residual data is normal, the line that describes the actual data follows the diagonal line (Oktaviani & Sarkawi, 2017). The normality test curve shows that the sampling error used is normally distributed, allowing multiple linear regression models to be used in the analysis.

The following is a graph of the normality test results using the P-P Plot for research data using SPSS software version 22 in Figure 1.

The format of the results of the study and discussion is not separated, considering the number of pages available to the author is limited. The results of the study can be presented with the support of tables, graphs or images as needed, to clarify the presentation of the results verbally.

Table titles and graphs or image captions are concisely arranged in phrases (not sentences). The image/graphic caption is placed below the image/graphic, while the table title is placed on top of it. The title begins with a capital letter. An example as follows. Notice the elastic pendulum on the Figure 1.

From the normal P-P *plot of regression standardized residuals* in Figure 1, it shows that the points are spread around the diagonal line and the spread is not too long or too wide. It can be concluded that the data is distributed around the line and the data pattern is usually linear between **Y** and **X**, so the data distribution can be said to be normal. In other words, the regression model is normal and can be used.

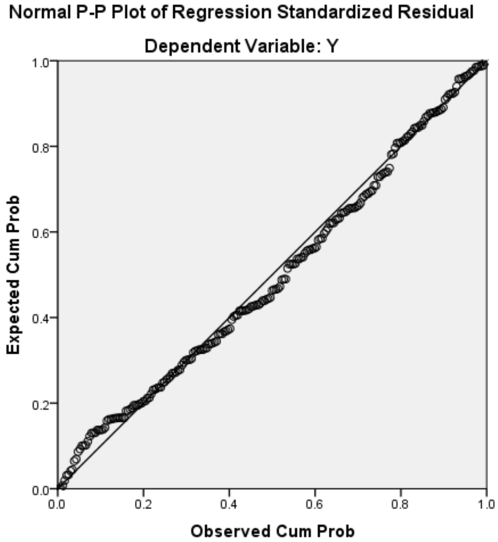

Figure 1. Image of Normality Test Results Using P-P Plot

3.1.1.2. **Kolmogorov-Smirnov Test**

In addition to the normality test using the P-P Plot, another statistical test that can be used to test normality is the Kolmogorov-Smirnov non-parametric statistical test. If the significance value of the Kolmogorov-Smirnov test is greater than 0.05, the data is normally distributed (Sintia et al., 2022). The following Table 2 provides a table of Kolmogorov-Smirnov test results for research data using SPSS version 22 software.

Table 2. Kolmogorov-Smirnov Test Result Table

One-Sample Kolmogorov-Smirnov Test

		Unstandardized Residual
N		223
Normal Parameters[a,b]	Mean	.0000000
	Std. Deviation	10.77338828
Most Extreme Differences	Absolute	.053
	Positive	.051
	Negative	-.053
Test Statistic		.053
Asymp. Sig. (2-tailed)		.200[c,d]

a. Test distribution is Normal.

b. Calculated from data.

c. Lilliefors Significance Correction.

d. This is a lower bound of the true significance.

Based on the results in Table 2, a significant value of 0.200 was obtained. This significant value is greater than 0.05 which indicates that the data is normally distributed.

3.1.2. Multicollinearity Test

Multicollinearity is the presence of a linear relationship between independent variables in a multiple linear regression model (Wasilaine et al., 2014). The linear relationship between independent variables can take the form of perfect (perfect) and imperfect (incomplete) linear relationships. A regression model is said to experience multicollinearity if some or all of the independent variables in the linear function have a perfect linear function. To detect multicollinearity in multiple linear regression models, the variance inflation factor (VIF) and tolerance values can be used. If the VIF value < 10 dan *tolerance* > 0,1 it is said that there is no multicollinearity in the regression model.

The following Table 3 provides a table of multicollinearity test results for research data using SPSS version 22 software.

Table 3. Multicollinearity Test Result Table

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	T	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	10.505	2.598		4.044	.000		
	X1	1.638	.267	.321	6.144	.000	.493	2.030
	X2	.551	.050	.580	11.102	.000	.493	2.030

a. Dependent Variable: Y

Based on the results in Table 3, the VIF value is 2,030 < 10 and tolerance 0,493 > 0,1. This means that there is no multicollinearity in the regression model.

3.1.3. Heteroscedasticity Test

Heteroscedasticity is a condition where the variance of the residuals of all observations in the regression model is not the same (Hanifah et al., 2015).

3.1.3.1. Scatterplot Test

One way to test for the presence of heteroscedasticity problems in a regression model is to plot the data where the X-axis is the prediction of Y. And the Y-axis is the residual (Y prediction-Y actual) being examined. If the results of the data plot show a patterned distribution or there is a certain pattern, for example a pattern of dots, or a certain regular pattern (bubbles, flares, then narrowing), then heteroscedasticity occurs. Conversely, if the data plot does not show a clear pattern and the dots are scattered above and below the number 0 on the Y axis, it can be said that there is no heteroscedasticity problem. The following is the condition that there is no heteroscedasticity problem based on the scatterplot:

- Data points spread above and below or around the number 0,
- Data points do not just gather on the upper or lower side,
- The distribution of data points does not form a wave pattern, widening, narrowing, and widening again,
- The distribution of data points is not patterned.

The following is given the results of the data plot using Scatterplot for research data using SPSS software version 22 in Figure 2.

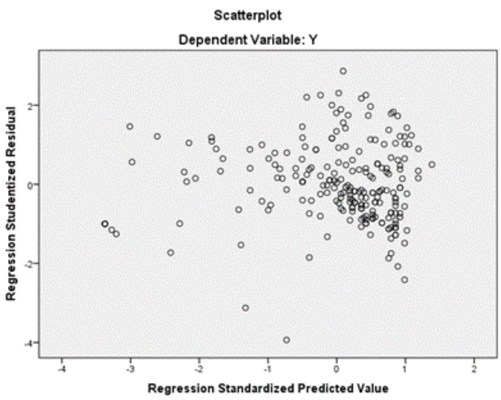

Figure 2. Scatterplot Test Result Image

From Figure 2, it can be seen that the data is randomly distributed, meaning that there is no heteroscedasticity problem, so the multiple linear regression model can be used in this analysis. \

3.1.3.2. Glejser Test

In addition to testing for heteroscedasticity problems using Scatterplot, another statistical test that can be used to test for heteroscedasticity problems is the Glejser test. The test is performed by regressing the independent variable on the absolute residual value. Residuals are the difference between the value of variable Y and the predicted value of variable Y, and absolute values are absolute values (all values are positive). When the significance value between the independent variable and the absolute residual > 0.05, there is no heteroscedasticity.

The following in Table 4 provides a table of Glejser test results for research data using SPSS version 22 software.

Table 4. Table of Glejser Test Results

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	T	Sig.	Tolerance	VIF
1	(Constant)	8.900	1.604		5.548	.000		
	X1	.223	.165	.129	1.357	.176	.493	2.030
	X2	-.052	.031	-.161	-1.683	.094	.493	2.030

a. Dependent Variable: Abs RES

Based on the results in Table 4, the significance values between the independent variables X_1 and X_2 are each greater than 0.5. This means that there is no heteroscedasticity problem in the regression model. In other words, it can be concluded that the variation of the dependent variable (final grade of Linear Algebra course) does not vary at each level of the independent variables used in the regression model. This shows that the multiple linear regression model used has homoscedasticity, namely that the variance of the final grade of Linear Algebra course is relatively constant at each level of the independent variable, and this fulfills one of the basic assumptions of multiple linear regression analysis.

3.2. Hypothesis Test

Hypothesis testing aims to determine whether there is a significant relationship between the independent variable and the dependent variable. The null hypothesis in this hypothesis test states that there is no significant relationship between the independent variable and the dependent variable, while the alternative hypothesis states that there is a significant relationship between the independent variable and the dependent variable.

3.2.1. F-Test (Simultaneous ANOVA)

F-test or simultaneous ANOVA in multiple linear regression analysis is a statistical technique used to test whether all independent variables together have a significant effect on the dependent variable in multiple linear regression models. This technique is done by comparing the regression variance (variability that can be explained by the regression model) with the residual variance (variability that cannot be explained by the regression model). The results of the F-test or simultaneous ANOVA are important to determine whether the multiple linear regression model built can accurately describe the relationship between the dependent variable and the independent variable.

The following Table 5 provides a table of F-test results (simultaneous ANOVA) for research data using SPSS software version 22.

Table 5. F-Test Results Table (Simultaneous ANOVA)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	61394.645	2	30697.322	262.099	.000[b]
	Residual	25766.629	220	117.121		

Total	87161.274	222			

a. Dependent Variable: Y

b. Predictors: (Constant), X2, X1

Based on the results in Table 5, the F-count value is 262.099 with a significant level of 0.000. This significant value is below 0.05, which indicates that the independent variables X_1 dan X_2 simultaneously have a significant effect on the independent variable Y at a significant level of 5%.

3.2.2. T-test (Partial)

In multiple linear regression analysis, the main objective is to build a mathematical model that can be used to predict the final value based on the given independent variables. The t-test (partial) in multiple linear regression analysis is a statistical technique used to test whether an independent variable individually has a significant effect on the dependent variable in a multiple linear regression model. In analyzing the final grade of Linear Algebra course with multiple linear regression approach, partial t-test can be used to evaluate the effect of each independent variable on the final grade. In this case, a large t-value indicates that the independent variable has a significant influence on the final grade, while a small t-value indicates otherwise.

The following Table 6 provides a table of t-test results (partial) for research data using SPSS software version 22.

Table 6. Table of T-test Results (Partial)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	T	Sig.
1	(Constant)	10.505	2.598		4.044	.000
	X1	1.638	.267	.321	6.144	.000
	X2	.551	.050	.580	11.102	.000

a. Dependent Variable: Y

Based on the results in Table 6, the significant value for X_1 and X_2 are 0,000. This significant value is less than 0,05 which means that the independent variables X_1 dan X_2 partially affect the independent variable Y at the 5% significant level.

Furthermore, based on Table 6, the multiple linear regression equation can be arranged, namely:

$$\boldsymbol{Y = 10,505 + 1,638X_1 + 0,551X_2} \qquad (\mathbf{2})$$

with

Y=final grade,

X_1=number of attendance

X_2=daily grades.

Based on equation (2), the constant value is $\boldsymbol{\alpha = 10,505}$. This states that if there is no influence from the independent variables, namely the number of attendance and daily grades, the final grade is 10,505. Meanwhile, the regression coefficient value for the independent variable number of attendance (X_1) is positive, namely $\boldsymbol{\beta_1 = 1,638}$, indicating that for every 1 increase in the number of attendance, the final grade increases by 1,638. Furthermore, the regression coefficient value for the independent variable daily grade (X_2) is positive, namely $\beta_2 = 0,551$, indicating that for every 1 increase in daily grade, the final grade increases by 0,551.

The comparison between the results of the final value prediction model in equation (2) and the actual data can also be presented in a linear graph in Figure 3.

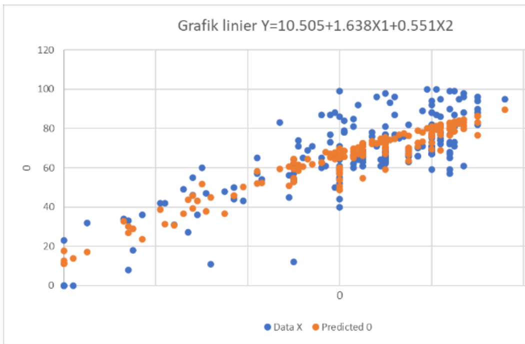

Figure 3. Linear Graph of Prediction Model Equation (2)

3.3. Dominance Test

Dominance test is a method to determine which independent variable is the most dominant or contributes the most to the variation in the dependent variable. The dominance test is carried out by calculating the standardized coefficients for each independent variable and comparing them to determine which independent variable has the greatest contribution.

Based on Table 6, the standardized coefficients beta value for variable X_2 is 0,580. This value is greater than the standardized coefficients beta value for variable X_1 which is 0,321. In this case, it indicates that the independent variable X_2 namely daily grades, is more influential than the independent variable X_1 namely the number of attendances.

3.4. R^2-Test

The R^2 test in multiple linear regression analysis is a statistical technique used to measure how much variation or variability in the dependent variable can be explained by the multiple linear regression model. This test produces a value between 0 and 1, where the higher the R^2 value, the greater the proportion of variation in the dependent variable that can be explained by the multiple linear regression model. The R^2 test results are important for evaluating the quality of multiple linear regression models.

The following Table 7 provides a table of R^2 test results for research data using SPSS software version 22.

Table 7. R^2-Test Result Table

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.839[a]	.704	.702	10.822

a. Predictors: (Constant), X2, X1

Table 7 shows that R = 0.839 and R^2= 0,704. This means that the ability of the independent variables to explain the variance of the independent variables is 70.4%. There is still 29.6% of the variance of the independent variables that cannot be explained by the independent variables. The final grade (Y) that can be explained by the number of attendance (X_1) and daily grades (X_2) in equation (2) is 70,4%. The remaining 29.6% is explained by other factors outside the variables in equation (2).

4. Conclusion

Based on the results of the analysis using the Multiple Linear Regression approach, it can be concluded that the number of attendance and daily grades of students have a significant effect on the final grade in the Linear Algebra course. Both independent variables have significance values smaller than 0.05 and the coefficient of determination (R^2) value shows that 70.4% of the variation in the final grade can be explained by the independent variables. In this case, daily grades have a greater influence than the number of attendances in predicting the final grade. Therefore, it is recommended for students to focus more on doing daily assignments in order to improve their academic performance in Linear Algebra courses. In addition, attendance in class is also very important to gain a good understanding of the material being taught. This research can be a reference for lecturers or teachers in developing more effective learning strategies for Linear Algebra courses. In addition, similar research can be conducted for other courses with different independent variables to find out what factors affect students' final grades in the course.

References

Adna, S. F. 2016. Descriptive Analysis of Student Learning Outcomes in Mathematical Statistics I Course. Delta: *Scientific Journal of Mathematics Education*, *4*(2), 39–50.

Aswan, N., & Fadhillah, Y. 2022. The Effect of PACE Model Learning Implementation on Statistics Learning Outcomes. *Journal of Cultural Education Science*, *2*(1), 1–7.

Fahmeyzan, D., Soraya, S., & Etmy, D. 2018. Normality Test of Monthly Turnover Data for Microeconomic Actors in Senggigi Village Using Skewness and Kurtosi. *VARIAN*, *2*(1), 31–36. https://doi.org/10.30812/varian.v2i1.331

Fahmi, D. A., & Hidayat, M. A. 2014. Correlation of Daily Test Scores, Mid-Semester Test Scores, and Final Semester Test Scores, with Fourth Grade Mathematics Report Card Scores of Fourth Grade Students of SD Negeri 1 Bae Kudus. *JPPK (Jurnal Pendidikan Dan Pembelajaran Khatulistiwa), 4*(1), 19–28.

Hanifah, N., Herrhyanto, N., & Agustina, F. 2015. Application of the Weigthed Least Square Method to Overcome Heteroscedasticity in Linear Regression Analysis. *EurekaMatika Journal*, *3*(1), 105–114.

Kurnia, H. 2014. The Effect of Organizational Activity on Cumulative Grade Point Average of Students of Universitas Cokroaminoto Yogyakarta. *Academy of Education Journal*, *5*(2), 91–103. https://doi.org/10.47200/aoej.v5i2.120

Magfirah, I., Rahman, U., & Sulasteri, S. 2015. The Effect of Self-Concept and Study Habits on Mathematics Learning Outcomes of Class Viii Students of SMP Negeri 6 Bontomatene Selayar Islands. *MaPan: Journal of Mathematics and Learning*, *3*(1), 103–116.

Mardiatmoko, G. 2020. The Importance of the Classical Assumption Test in Multiple Linear Regression Analysis. *BAREKENG: Journal of Mathematical and Applied Sciences*, *14*(3),333–342. https://doi.org/10.30598/barekengvol14iss3pp333-342

Muktiadi, R., & Mustafidah, H. 2012. Fuzzy Quantification System to Analyze the Effect of Study Interest and Student Attendance Rate on Student Learning Achievement of Muhammadiyah Purwokerto University. *JUITA*, *II*(1), 65–72.

Pamungkas, N. A., & Mustafidah, H. 2016. Analysis of Student Learning Discipline and Student Attendance on Course Grades Using Fuzzy Quantification Theory. *SAINTEKS*, *XIII*(1), 71–82.

Permana, R. A., & Ikasari, D. 2023. Data Normality Test Using Empirical Distribution Function Method Utilizing Matlab and Minitab 19. *Semnas Ristek (National Seminar on Research and Technology Innovation)*, *7*(1), 7–12. https://doi.org/10.30998/semnasristek.v7i1.6238

Prasetyo, M. H., & Hasyim. 2022. Analysis of the Effect of Study Hours on Grade Point Average (GPA) of Civil Engineering Students of Batam International University. *Nusantara Hasana Journal*, *1*(11), 22–32.

Sintia, I., Pasarella, M. D., & Nohe, D. A. 2022. Comparison of the Consistency Level of Normality Distribution Test in the Case of Unemployment Rate in Java. *National Seminar on Mathematics, Statistics, and Its Applications,*2*(*2), 322–333.

Syam, S., Rahman, U., & Nursalam, N. 2016. The Effect of Metacognition Knowledge and Visual Learning Style on Mathematics Problem Solving Ability of Class IX Students of SMP Negeri 2 Barombong, Gowa Regency. *MaPan: Journal of Mathematics and Learning*, *4*(2), 231–244. https://doi.org/10.24252/mapan.2016v4n2a7

Wanti Wulan Sari, N., Suyitno, & Mirwansyah, D. 2019. Factors Affecting the GPA of STMIK Sentra Pendidikan Bisnis Graduates. *National Seminar on Mathematics, Statistics, and Their Applications 2019*, *1*(5), 199–205.

Wasilaine, T. L., Talakua, M. W., & Lesnussa, Y. A. 2014. Ridge Regression Model to Overcome Multiple Linear Regression Models Containing Multicollinearity. *BAREKENG: Journal of Mathematical and Applied Sciences*, *8*(1), 31–37. https://doi.org/10.30598/barekengvol8iss1pp31-37

Wibowo, S. A., Mustafidah, H., Wicaksono, A. P., & Aryanto, D. 2013. Analysis of Learning Motivation and Attendance on Student Course Grades Using Fuzzy Quantification Theory. *JUITA*, *II*(3), 175–181.

Zharbaini, F., Saputra, H. D., & Putra, D. S. 2023. The Relationship between Active Organizing and Cumulative Grade Point Average, Case Study: Automotive Engineering Students of Padang State University. *JTPVI (Jurnal Teknologi Dan Pendidikan Vokasi Indonesia)*, *1*(1), 119–124