International Proceeding
Universitas Tulungagung
2025

MULTIVARIATE REGRESSION ANALYSIS OF FACTORS INFLUENCING STUDENTS' MATHEMATICS ACHIEVEMENT

Amazia Ananta Prabaswara[1], Adi Setiawan[2]
[1]Master of Data Science, Faculty of Science and Mathematics,
Universitas Kristen Satya Wacana, Indonesia
amazia.ananta@uksw.edu

Abstract

Mathematics achievement serves as a crucial indicator for assessing students' logical reasoning, analytical thinking, and problem-solving abilities. However, students' performance in mathematics is influenced by various internal and external factors. This study aims to analyze the impact of multiple factors on students' mathematics achievement using a multivariate regression approach. The research involved 129 students from Satya Wacana Christian Junior High School, utilizing questionnaire responses and mid-semester test scores as the dependent variable. A total of 24 independent variables were examined, including demographic factors, family background, learning motivation, and learning environment. The regression analysis revealed a coefficient of determination of $R^2 = 0.3983$ and an adjusted $R^2 = 0.1384$, indicating that approximately 13.8% of the variance in mathematics performance could be explained by the variables in the model. Three factors were found to have a significant influence: mother's education ($\beta = +68.4431$; $p = 0.0138$), father's education ($\beta = +44.981$; $p = 0.0230$), and the availability of Wi-Fi at home ($\beta = -17.8457$; $p = 0.0129$). These findings highlight that parental educational background plays a pivotal role in supporting students' academic success, whereas uncontrolled internet access may negatively affect learning outcomes. The study underscores the importance of family involvement in the educational process and the need for guidance in utilizing home internet access to foster effective learning.

Keywords: *Machievement; Multivariate regression; parental education; digital technology; learning factors*

INTRODUCTION

Mathematics holds a fundamental role within the Indonesian education system (Amalia et al., 2024), serving as a foundation for the development of logical reasoning, analytical thinking, and problem-solving skills (Siregar et al., 2024). Mathematics achievement is often regarded as a benchmark for evaluating students' overall academic competence (Waritsman, 2020). Nevertheless, students' learning outcomes in mathematics vary considerably, influenced by a wide range of internal and external factors.

Internal factors arise from within the students themselves, including motivation, intelligence, interest, and attitudes toward mathematics (Sinlae et al., 2023). External factors, on the other hand, encompass family support, teacher roles, learning environments, and the availability of learning facilities and technology (Arrosyad et al., 2023; Sembiring, 2025). Understanding which factors exert the greatest influence on students' mathematics

performance is essential for schools and parents in designing effective strategies to enhance academic achievement (Amalia et al., 2024; Syafi'i et al., 2018).

This study employs a multivariate regression analysis to examine the simultaneous influence of multiple factors on students' mathematics achievement. The novelty of this research lies in its application of a comprehensive predictive regression model to identify the most dominant determinants and to provide data-driven recommendations for educational practice. Furthermore, this study offers a distinctive perspective by emphasizing contextual and non-academic variables—such as type of residence and family size—that are often overlooked but may substantially affect mathematics performance.

The contributions of this research are twofold: (1) **Theoretically**, it enriches empirical literature on the determinants of mathematics achievement through a quantitative analytical approach; and (2) **Practically**, it provides valuable insights for teachers, parents, and educational policymakers to develop strategies aimed at improving students' mathematics learning outcomes.

LITERATURE REVIEW

Mathematics achievement is defined as the level of students' mastery of mathematical concepts, procedures, and problem-solving abilities, which is generally assessed through standardized tests or academic evaluations (Achdiyat, 2016).

Previous studies have consistently demonstrated that learning motivation has a positive correlation with academic achievement (Lutfiwati, 2020; Ramadhani & Suriani, 2024). In addition, parental support, quality of instruction, and a conducive learning environment have been found to play a vital role in shaping students' academic performance (Ritonga et al., 2025). The integration of technology has also become increasingly relevant in the digital era, influencing students' learning styles and engagement (Nimah et al., 2025; Setiyawan et al., 2025).

Multivariate regression analysis is a statistical technique used to evaluate the influence of multiple independent variables on a single dependent variable (Dukalang et al., 2025; Kawano et al., 2025). Unlike simple regression, which focuses on a single predictor, multivariate regression allows researchers to assess the combined and relative effects of several predictors simultaneously. This approach provides a more comprehensive understanding of the relationships among variables and produces more accurate estimations of the factors influencing the dependent variable.

In the context of education, multivariate regression has been widely applied to identify the interplay between demographic, psychological, and environmental factors affecting students' academic performance. Therefore, employing this method enables researchers to capture the multifaceted nature of educational outcomes, offering both explanatory and predictive insights for improving teaching and learning processes.

RESEARCH METHOD

This study employed a quantitative approach with an inferential analytical design.

- **Research Sample:**

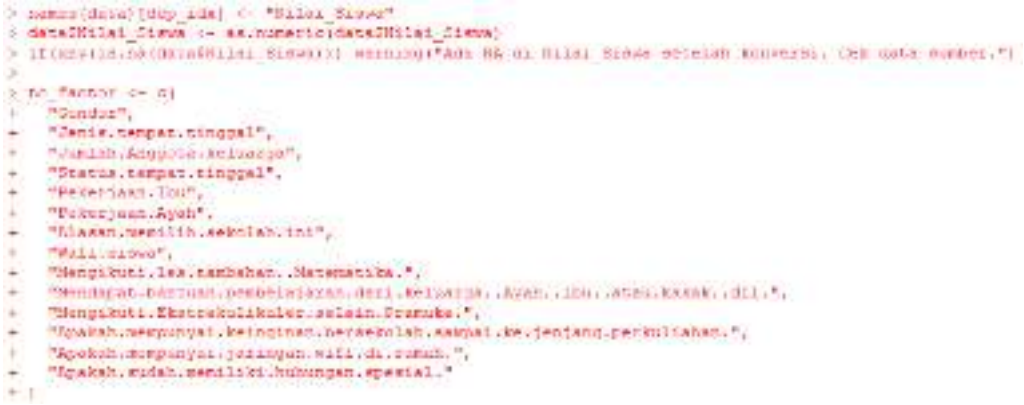
  The sample was determined using a random sampling technique involving students of Satya Wacana Christian Junior High School from Grades VII to IX, with a total of 129 participants.

- **Research Variables:**
  - The **dependent variable** in this study was *students' mathematics achievement*, represented by their Mid-Semester Examination scores for the 2025/2026 academic year.
  - The study also included **24 independent variables**, encompassing both demographic and contextual factors: gender, age, type of residence (urban/rural), number of family members, living arrangement (with parents or separately), mother's education, father's education, mother's occupation, father's occupation, reason for choosing the current school, guardian status, commuting time to school, weekly study hours, participation in private tutoring, family learning support, involvement in extracurricular activities (other than scouting), aspiration for higher education, Wi-Fi availability at home, relationship status, quality of family relationships, after-school leisure time, time spent with friends, health condition, and number of school absences.
- **Research Instruments:**

  Two instruments were employed to collect data.
  - A **Google Form questionnaire** was used to gather students' responses on independent variables.
  - **School records and report cards** (legers) were utilized to obtain the mathematics scores representing the dependent variable.
- **Data Preparation and Preprocessing:**
  1. Data Preprocessing

     Prior to statistical analysis, data preparation was conducted to ensure accuracy and consistency. One crucial step was **data conversion**, in which non-numeric (categorical) variables were transformed into numeric or factor formats. This process is essential since statistical analyses such as multivariate regression require numerical inputs. Examples of categorical variables converted include gender, type of residence, and living arrangement. In the R software, this conversion was performed using functions such as as.factor() or encoding techniques targeting specific variables.

Figure 1. Coding conversi variable



2. Instrument Validity and Reliability:

   Initial analyses were carried out to verify that all instruments met the criteria for validity and reliability. Valid instruments ensure that the data collected accurately represent the constructs being measured, while reliable instruments produce consistent results across repeated measurements.

3. Descriptive Analysis.

   Descriptive statistics were used to provide a general overview of the data, including respondents' characteristics and the distribution of variable scores. Statistical indicators such as measures of central tendency (mean, median, mode) and measures of dispersion (standard deviation, variance, range) were employed. These descriptive analyses helped researchers identify data patterns prior to further testing.

4. Classical Assumption Testing

   Before conducting the multivariate regression analysis, classical assumption tests were performed to ensure that the model satisfied the requirements of linear regression. These included:

   - **Normality test** of residuals, to verify that residuals were normally distributed.
   - **Multicollinearity test**, to detect any strong correlations among independent variables.
   - **Autocorrelation test** (when necessary), to identify potential correlations among residuals across observations.

   Meeting these assumptions ensures that regression estimates are valid and unbiased.

5. Multivariate Regression Analysis.

   After all assumptions were fulfilled, a multivariate regression analysis was performed using the R and Python statistical software. The objective was to identify the influence of multiple independent variables on students' mathematics achievement simultaneously and determine which factors exert the most significant effects.

RESULT

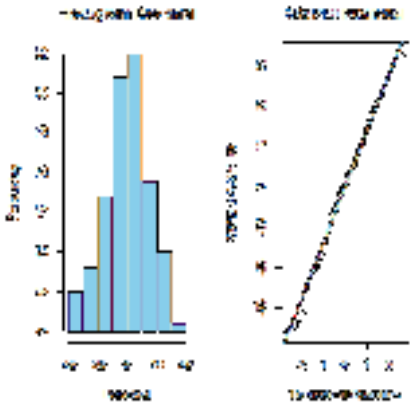
The multivariate regression analysis began with several classical assumption tests, including normality, multicollinearity, and autocorrelation tests.

The Shapiro–Wilk test was employed to assess the normality of the residuals. The test yielded a W value of 0.98862 with a p-value of 0.3778. Since the p-value exceeded the significance threshold ($p > 0.05$), the residuals were considered normally distributed. In addition to the statistical test, the distribution plots were visually examined. The histogram and Q–Q plot showed that the residuals were approximately aligned with the diagonal line, confirming that the normality assumption was satisfied.

Figure 2. Normality Test Results

```
        Shapiro-Wilk normality test

data:  residuals(model)
W = 0.98862, p-value = 0.3778
```



The multicollinearity test was performed to ensure that no independent variables were highly correlated with each other. The **Generalized Variance Inflation Factor (GVIF)** was used for this purpose. The results indicated that all independent variables had GVIF values below 5, with most values even lower than 2. These findings suggest that multicollinearity was not a concern in the regression model, and therefore, all independent variables were retained for further analysis.

Figure 3. Multicollinearity Test Results

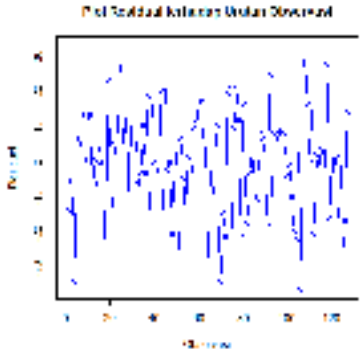
	GVIF^(1/(2*DF))
Gender	1.102940
Usia.siswa	1.218061
Jenis.tempat.tinggal	1.203800
Jumlah.Anggota.keluarga	1.095630
Status.tempat.tinggal	1.423314
Pendidikan.Ibu	1.213599
Pekerjaan.Ibu	1.153680
Pekerjaan.Ayah	1.200703
Alasan.memilih.sekolah.ini	1.167897
Wali.siswa	1.297104
Waktu.tempuh.rumah.ke.sekolah	1.422308
Waktu.belajar.selama.1.minggu	1.397469
Mengikuti.Les.tambahan..Matematika.	1.208694
Mendapat.bantuan.pembelajaran.dari.keluarga..Ayah..ibu..atau.kakak..dll.	1.301573
Mengikuti.Ekstrakulikuler.selain.Pramuka..	1.212131
Apakah.mempunyai.keinginan.bersekolah.sampai.ke.jenjang.perkuliahan..	1.242979
Apakah.mempunyai.jaringan.wifi.di.rumah..	1.218307
Apakah.sudah.memiliki.hubungan.spesial..	1.304087
Kualitas.hubungan.keluarga	1.472937
Waktu.luang.setelah.sekolah	1.236296
Jalan...jalan..main.bersama.dengan.teman..bertemu.langsung.	1.219621
Status.kesehatan	1.230903
Jumlah.ketidak.hadiran.dari.awal.masuk.sampai.sekarang	1.229335

Autocorrelation was tested using the **Durbin–Watson (DW) test**, which yielded a DW value of 2.0353 with a p-value of 0.764. Since the p-value was greater than 0.05, there was no evidence of autocorrelation among residuals. The random pattern observed in the residual plot also supported this conclusion, indicating that the independence assumption was met.

FIgure 4. Autocorrelation Test Results

```
        Durbin-Watson test

data:  model
DW = 2.0353, p-value = 0.764
alternative hypothesis: true autocorrelation is greater than 0
```



The results of the multivariate regression analysis revealed that the model had a relatively weak explanatory power for students' mathematics performance, with an R^2 value of 0.3983 and an adjusted R^2 of 0.1384. The overall F-statistic was 1.533 ($p = 0.052$), which was marginally above the 5% significance level, suggesting that the model was close to achieving overall statistical significance.

Figure 5. Multivariate Regression Test Results

```
Residual standard error: 17.38 on 88 degrees of freedom
Multiple R-squared:  0.3983,    Adjusted R-squared:  0.1384
F-statistic: 1.533 on 38 and 88 DF,  p-value: 0.05207
```

Most independent variables did not have significant effects ($p > 0.05$). However, three variables showed statistically significant influences on students' mathematics achievement, as detailed below:

- **Availability of Wi-Fi at home**: $\beta = -17.8457$ ($p = 0.0129$). Students with home Wi-Fi access tended to score approximately 17.8 points lower than those without such access.

```
Apakah.mempunyai.jaringan.wifi.di.rumah..Ya  -17.8457
```

- **Mother's education level**: $\beta = +68.4431$ ($p = 0.0138$). Students whose mothers had higher education levels achieved, on average, 68.4 points higher mathematics scores.

```
Pendidikan.Ibu.C      68.4431
```

- **Father's education level**: $\beta = +44.981$ ($p = 0.0230$). Students with fathers holding higher education degrees scored around 45 points higher than those whose fathers had lower educational attainment.

```
Pendidikan.Ayah.Q     44.9815
```

Other variables—such as gender, age, family size, parental occupation, reasons for school selection, commuting time, and extracurricular involvement—did not show significant relationships with mathematics performance ($p > 0.05$).

These results collectively indicate that parental educational background and access to digital resources are key determinants of students' academic performance, while other demographic factors appear to exert minimal influence.

DISCUSSION

The findings of this multivariate regression analysis indicate that among the 24 independent variables examined, only three factors—mother's education, father's education, and the availability of Wi-Fi at home—had significant effects on students' mathematics achievement. These results provide insights into the external dynamics influencing students' academic performance.

The significant positive relationship between **mother's education** and students' mathematics achievement ($\beta = +68.4431$; $p = 0.0138$) supports the view that parental education plays a critical role in shaping children's academic outcomes. Mothers with higher educational attainment are generally more capable of providing academic guidance, establishing structured learning environments, and fostering positive attitudes toward education. This finding aligns with Wahyuni (2022) and Ramadhan & Ichsan (2021), who emphasized that maternal education positively correlates with children's academic achievement. Similarly, **father's education** also demonstrated a positive association ($\beta = +44.981$; $p = 0.0230$), although to a lesser extent. Fathers often serve as role models and provide emotional and motivational support, which can strengthen students' engagement and perseverance in learning. This result resonates with Rahayu (2020), who found that educated parents possess broader perspectives and knowledge that enable them to guide their children more effectively and cultivate disciplined learning habits.

Interestingly, the **availability of Wi-Fi at home** exhibited a significant negative effect ($\beta = -17.8457$; $p = 0.0129$). Students with home internet access tended to have lower mathematics scores. This phenomenon may be attributed to uncontrolled internet use, where students allocate excessive time to non-academic online activities such as social media or online gaming. While digital access is crucial for modern learning, excessive or unregulated use can reduce students' focus and learning quality. Similar observations were reported by Ramadhan & Diah (2023) and Syifa et al. (2023), who noted that digital distractions often lead to decreased academic performance despite the educational potential of online resources.

Other variables—such as gender, age, family size, parental occupation, extracurricular participation, and study duration—were found to have no statistically significant impact on mathematics achievement. These findings are consistent with Watson (2009), who argued that demographic factors such as gender and family size do not directly affect academic outcomes but may interact with psychological or contextual mediators, such as motivation, self-efficacy, and learning strategies.

Overall, these results highlight the complexity of educational attainment, which cannot be fully explained by demographic characteristics alone. Instead, **family educational background and digital engagement patterns** emerge as critical determinants of students' mathematics achievement in the contemporary learning environment.

CONCLUSION

This study concludes that among the 24 independent variables analyzed, three factors—mother's education, father's education, and the availability of Wi-Fi at home—significantly influence students' mathematics achievement. The findings emphasize that **parental education** plays a crucial role in supporting children's cognitive development and creating a conducive

home learning environment. Conversely, **uncontrolled Wi-Fi access** may act as a distraction that diminishes students' focus and leads to lower mathematics performance.

These results suggest that strategies to enhance mathematics achievement should not only concentrate on classroom learning but also involve **active family engagement** in fostering a supportive educational ecosystem. Schools and parents must collaborate to provide structured guidance that encourages responsible internet use, ensuring that technology becomes an instrument for learning rather than an obstacle to academic success.

For future research, it is recommended to expand the study sample across multiple schools from diverse geographical areas—rural, urban, and interprovincial contexts—to strengthen the generalizability of the findings. Such efforts may contribute to a more comprehensive understanding of the social, familial, and technological factors shaping mathematics achievement among Indonesian students.

REFERENCES

Achdiyat, M. (2016). *Mathematics learning achievement in terms of students' self-confidence and classroom participation. Jurnal Pendidikan Matematika,* 6(1), 50–61.

Amalia, N., Hanifah, R. R., & Yani, A. (2024). *The role of mathematics education in strengthening national character. Proceedings of the National Panel Discussion on Mathematics Education,* 80, 701–705.

Arrosyad, M. I., Wahyuni, E., Kirana, D., & Sartika, M. (2023). *An analysis of factors affecting low mathematics achievement in elementary school students. Educativo: Jurnal Pendidikan,* 2(1), 222–228.

Dukalang, H. H., Otok, B. W., & Purhadi. (2025). Modified partial least squares structural equation model with multivariate adaptive regression spline: Parameter estimation technique and applications. *MethodsX, 14*(February), 103381. https://doi.org/10.1016/j.mex.2025.103381

Kawano, S., Fukushima, T., Nakagawa, J., & Oshiki, M. (2025). Multivariate regression modeling in integrative analysis via sparse regularization. *Japanese Journal of Statistics and Data Science.* https://doi.org/10.1007/s42081-025-00312-2

Lutfiwati, S. (2020). *Learning motivation and academic achievement. Jurnal Kependidikan Islam,* 10. http://ejournal.radenintan.ac.id/index.php/idaroh

Nimah, N. A., Nurullah, M., & Susilawati, S. (2025). The impact of digital technology on learning methods in primary schools. *Journal Neosantara Hybrid Learning,* 2(2), 455–461. https://doi.org/10.70177/jnhl.v2i2.1783

Rahayu, R. G. (2020). *The influence of parents' educational level on students' learning discipline at SDN 05 Kabawaten, Kepahiang Regency. Jurnal Pendidikan Dasar,* 2, 1–9.

Ramadhan, J., & Ichsan. (2021). *The effect of parental education on children's learning achievement in school. WANIAMBEY: Journal of Islamic Education,* 2(2), 69–78.

Ramadhan, M., & Diah, H. (2023). *The impact of online gaming on students' academic achievement: A case study at Universitas Islam Kebangsaan Indonesia. At-Tarbiyyah: Journal of Islamic Education,* 21(1), 39–46. https://doi.org/10.33830/jp.v21i1.7

Ramadhani, D., & Suriani, A. (2024). *The relationship between learning motivation and academic achievement. Jurnal Pendidikan Matematika dan IPA,* 3, 45–53.

Ritonga, R., Cahaya Putri, S., Huwaida, A., Amelia Nainggolan, N., Chatena Munte, C., William Iskandar, J., & Medan Estate, P. V. (2025). *The role of the learning environment in improving teaching quality among fifth-grade students at Universitas Negeri Medan Laboratory School. Aliansi: Journal of Law, Education, and Social Humanities,* 2, 127–132.

Sembiring, R. K. (2025). *The influence of grit on academic achievement in relation to internal and external factors. Jurnal Pendidikan,* 14(1), 148–158.

Setiyawan, H., Aulia, S. R., Firdaus, N. A., & Putri, E. A. (2025). *The effect of technology-based learning media on fifth-grade students' learning outcomes. Jurnal Pendidikan,* 25(2), 124–132. https://doi.org/10.52850/jpn.v25i2.15592

Sinlae, L. M., Kaluge, A. H., Manafe, H. A., Talok, D., Niha, S. S., & Yasinto, Y. (2023). *The influence of internal and external factors on learning achievement of tourism vocational students. Jurnal Ilmu Manajemen Terapan,* 4(3), 388–402.

Siregar, E. B., Hidayah, N., Karo, B., Samosir, D., Rajagukguk, W., & Medan, U. N. (2024). *Quality of mathematics education in Indonesia. Jurnal Ilmiah Widya Pustaka Pendidikan,* 12(2), 34–50.

Syafi'i, A., Marfiyanto, T., & Rodiyah, S. K. (2018). *A study of students' learning achievement and its influencing factors. Jurnal Komunikasi Pendidikan,* 2(2), 115. https://doi.org/10.32585/jkp.v2i2.114

Syifa, S. F., Nur Istirohmah, A., Lestari, P., & Nur Azizah, M. (2023). *The impact of social media use on students' academic performance. BELAINDIKA: Journal of Learning and Educational Innovation,* 5(1), 21–27. https://doi.org/10.52005/belaindika.v5i1.100

Wahyuni, C. (2022). Parental involvement and the achievement of students with special education needs in Indonesia. *Exceptionality Education International,* 32(1), 14–34. https://doi.org/10.5206/EEI.V32I1.14871

Waritsman, A. (2020). *The relationship between learning motivation and mathematics achievement among students. Tolis Ilmiah: Journal of Research,* 1(2), 124–129.

Watson, D. M. (2009). Determinants of parasitic plant distribution: The role of host quality. *Botany,* 87(1), 16–21. https://doi.org/10.1139/B08-105