**GEMARIMATIKA JOURNAL**
***(Media Innovation and Development of Mathematics Learning)***
Volume 01 Number 01, April 2026

# ANALYSIS OF STUDENTS' STATISTICAL THINKING ABILITY IN SOLVING CONTEXTUAL PROBLEMS ON CORRELATION AND LINEAR REGRESSION

**Helmi Rahmawati[1], Saiful Fadli[2], Baiq Mauli Aprianti[3]**

1,2,&3Mathematics Education Study Program, Universitas Qamarul Huda Badaruddin Bagu

Correspondence Author: helmirahmawati18@gmail.com

**Abstract**

This study aims to analyze the statistical thinking ability of eleventh-grade students in solving contextual problems on correlation and linear regression. The study employed a descriptive qualitative approach supported by quantitative data. The subjects of this research were 40 students of class XI MIPA 2 at MAN 1 Mataram who had learned correlation and linear regression material. Data were collected through a statistical thinking ability test and semi-structured interviews with students representing high, medium, and low ability categories. The results showed that 30.40% of students were in the high category, 43.20% in the medium category, and 26.40% in the low category. Students in the high category were able to calculate, interpret, and relate the results of analysis to real-life contexts. Students in the medium category were generally able to perform calculations but still experienced difficulties in interpretation and concept application. Meanwhile, students in the low category showed difficulties in understanding basic concepts, performing calculations, and interpreting results. These findings indicate that students' statistical thinking abilities are still not optimal, especially in terms of interpretation and contextual application. Therefore, learning should emphasize conceptual understanding, data analysis, and contextual approaches to improve students' statistical thinking abilities.

**Keywords**: Statistical Thinking, Correlation, Linear Regression, Contextual Learning

## 1. INTRODUCTION

In today's information- and data-driven era, the ability to understand and use data effectively is an essential competency that younger generations must possess. Statistics is not merely a collection of formulas and computational techniques; it also encompasses the ability to interpret, evaluate, and draw meaningful conclusions from real-world data (Andriatna et al., 2021). This shift reflects the demands of 21st-century skills known as statistical literacy, namely the ability to read, analyze, and use data critically in solving real-world problems. Recent literature reviews show a growing trend of attention to developing students' statistical literacy from primary through secondary education as part of modern education, because this competency is fundamental to addressing data-driven challenges in everyday life (Hafiyusholeh et al., 2018; Ritmayanti et al., 2025).

In the context of mathematics education at the senior high school level, particularly Grade XI, correlation and linear regression are among the important topics in statistics. This material involves not only procedural skills in calculating correlation coefficients or regression equations, but also requires students to understand relationships between two variables, recognize the direction and strength of those relationships, and apply linear regression models to predict values in real-life situations. Real-life contexts, such as the relationship between study time and grades or between production volume and profit, are examples of applications frequently found in textbooks and examinations. Studies focusing on statistical thinking emphasize that students' abilities involve not only performing calculations but also analyzing data in real-life contexts and drawing logical conclusions (Alem, n.d.; Mat Nor Azmay et al., 2023).

However, various empirical studies show that students' statistical thinking abilities are still not optimal (Aini & Warli, 2025; Takaria & Talakua, 2018). Several educational studies indicate that students often have difficulty interpreting data, understanding relationships between variables, and explaining statistical meaning in relevant contexts. This suggests that although students may be able to carry out computational procedures, they are often not yet able to apply statistical concepts appropriately when faced with problem contexts that require higher-order thinking skills (Mohamad Hasim et al., 2024).

In addition, recent literature reviews indicate that the focus of research in statistics education is increasingly developing toward statistical reasoning and statistical thinking, which emphasize students' ability to reason deeply about data and relationships between variables rather than merely applying mathematical techniques. These studies show that mastery of statistics is not only academically important but is also related to students' ability to make decisions based on data evidence in real-life situations (Mat Nor Azmay et al., 2023).

On the other hand, research that specifically measures students' statistical thinking abilities using contextual problems on correlation and linear regression remains relatively limited, especially at the senior high school level in Indonesia. This creates opportunities for further research not only to determine students' skill levels, but also to examine how they think when solving statistical problems presented in real-life contexts. This contextual approach is important because it provides authentic and relevant learning situations that can stimulate students to think critically and reflectively when solving data-based problems (Ritmayanti et al., 2025). Contextual learning emphasizes students' active engagement in understanding material by connecting it to real-world situations they encounter, thereby encouraging them to apply it in everyday life (Hidayana & Lianingsih, 2025). Furthermore, Alfirawati Lumahu et al. (2025) emphasize that contextual learning can develop deeper conceptual understanding and build students' learning motivation.

Furthermore, the educational context in Indonesia shows that data and statistical literacy often pose particular challenges (Prihastari et al., 2022). This is reinforced by research trends indicating that students' statistical literacy generally needs to be improved in reading data, interpreting graphical representations, and making decisions based on relationships between variables (Fitri et al., 2023; Pujiati et al., 2023). This condition shows that statistics learning oriented solely toward computational procedures is insufficient to develop students' statistical thinking abilities comprehensively (Ritmayanti et al., 2025).

Based on the description above, it is important to analyze the statistical thinking ability of Grade XI MIPA students at MAN 1 Mataram in solving contextual problems on correlation and linear regression. This study is expected to provide an empirical overview of the extent to which students can apply statistical concepts in meaningful contexts, as well as assist educators and policymakers in designing more effective learning strategies to improve this competency in accordance with the demands of 21st-century education.

## 2. RESEARCH METHODS

This study employed a descriptive qualitative approach supported by quantitative data. This approach was selected to analyze in depth the statistical thinking ability of Grade XI students in solving contextual problems on correlation and linear regression. The subjects of this study were 40 students of class XI MIPA 2 at MAN 1 Mataram who had studied correlation and linear regression. The subjects were selected using purposive sampling. Several students were selected as interview informants based on the statistical thinking ability categories (high, medium, and low) obtained from the test results.

The main research instrument was a statistical thinking ability test containing contextual problems on correlation and linear regression, including the interpretation of bivariate data, determination of the type and direction of correlation, analysis of data trends, use of linear regression models, and prediction of values based on the problem context. To obtain more in-depth data, semi-structured interviews were conducted with selected subjects to explore their thinking processes and the difficulties they experienced in solving the problems.

The data were analyzed using quantitative descriptive analysis to categorize students' statistical thinking abilities based on test scores, as well as qualitative analysis involving data reduction, data presentation, and conclusion drawing regarding patterns in students' statistical thinking abilities based on the test and interview results. Data validity was ensured through technique triangulation by comparing test and interview results, as well as source triangulation across subjects with different ability levels. The indicators of statistical thinking ability are presented in Table 1 below.

Table 1. Indicators of Statistical Thinking Ability

| Statistical Thinking | Indicators |
|---|---|
| Data Description | 1. Identifying the uniqueness of data<br>2. Recognizing different displays of the same data<br>3. Evaluating the effectiveness of data presentation |
| Data Organization and Reduction | 1. Grouping data<br>2. Presenting measures of central tendency |
| Data Representation | 1. Constructing a data presentation from a data set or a given data presentation<br>2. Constructing a data display based on the data representation shown |
| Data Analysis and Interpretation | 1. Comparing data<br>2. Comparing across data sets or data presentations<br>3. Drawing conclusions |

(Aini & Warli, 2025)

The classification of students' statistical thinking abilities is presented in Table 2 below.

Table 2. Categories of Statistical Thinking Ability

| Statistical Thinking Test Score | Category |
|---|---|
| 71-100 | High |
| 41-70 | Medium |
| 10-40 | Low |

(Masjudin et al., 2020).

## 3. RESULTS AND DISCUSSION

### 3.1 Research Results

#### 3.1.1 Results of the Analysis of Students' Statistical Thinking Ability Test

The results of the statistical thinking ability test of class XI MIPA 2 students at MAN 1 Mataram are presented in Figure 1 below.

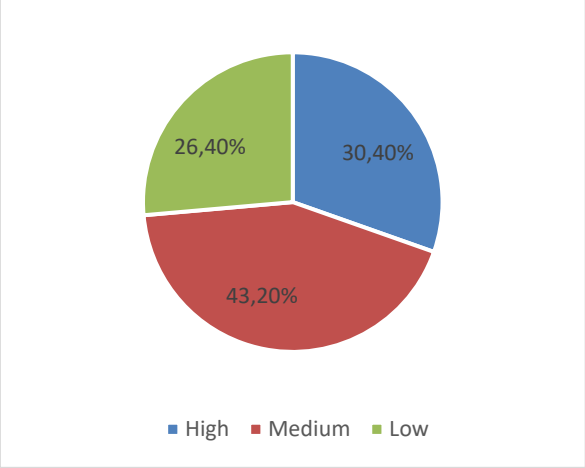

Figure 1. Students' Statistical Thinking Ability

Figure 1 illustrates the distribution of students' statistical thinking abilities across three categories: High, Medium, and Low. A smaller proportion of students (30.40%) demonstrated high statistical thinking ability, while the largest proportion (43.20%) was in the medium category, indicating that they had an adequate, although not very deep, understanding of statistics. Meanwhile, 26.40% of students demonstrated low statistical thinking ability, indicating difficulties in understanding statistical concepts or applying data analysis. Overall, the graph shows that most students were in the medium category, highlighting the need to improve statistical understanding more evenly, particularly among students in the low category, so that students' statistical thinking skills can develop more effectively.

### 3.1.2 Interview Results for Students with High Ability (SKT)

The following is the interview transcript of a student with high ability:

**Interview Transcript with SKT**

P: "How did you understand the task when you were asked to calculate the correlation coefficient between production volume and the percentage of defects?"

SKT: "I started by looking at the two variables. After calculating, I saw that the correlation was positive, although the value was small. This means there is a relationship, but it is not strong, between production volume and defects. I know that many other factors can affect the result, such as the quality of the materials or the machinery."

P: "Then, how did you explain the result of that weak correlation coefficient?"

SKT: "Because the correlation coefficient is only 0.26, I concluded that even if production increases, the number of defects does not necessarily increase substantially. There may be other factors that affect it, such as the production process or the quality of the materials used."

P: "Did you encounter any challenges when applying the linear regression model to predict defects based on production volume?"

SKT: "Actually, it was not too difficult. I used the regression model to estimate how many defects would occur if production increased. However, I also realized that this prediction is accurate only if other factors remain constant, such as the quality of the materials or the type of machinery used."

Students in this category demonstrated good mastery of correlation coefficient calculation, as SKT was able to calculate the correlation value accurately and understand that the positive correlation between production volume and the percentage of defects indicated a positive but weak relationship. In addition, SKT demonstrated good interpretation of the results by relating the correlation coefficient to a real-world context and explaining that increased production is not always accompanied by a significant increase in defects. Furthermore, SKT demonstrated competence in using linear regression, as indicated by the ability to use a linear regression model to predict the percentage of defects based on production data, although in some cases the student had difficulty explaining the rationale behind the prediction.

### 3.1.3 Interview Results for Students with Medium Ability (SKS)

The following is the interview transcript of a student with medium ability.

**Interview Transcript with SKS**

P: "How did you understand the problem about the correlation between production volume and the percentage of defects?"

SKS: "I can calculate the correlation well, but I am somewhat confused about what the number means. I know that the result is positive, but I do not really understand why the relationship is weak."

P: "What do you think about the weak relationship between production volume and defects?"

SKS: "I am not sure. Maybe if production increases, the quality could decrease, but I am not sure why the relationship is not very strong. I think there are other factors that affect it, but I do not know what those factors are."

P: "Did you have difficulty using linear regression to predict the number of defects?"

SKS: "Yes, I was somewhat confused when I was asked to predict defects based on production volume. I know there is an equation, but I do not really understand how to use the numbers I have obtained."

Students in this category were able to calculate the correlation coefficient. Most students were able to perform the correlation calculation well, although minor errors still occurred during the calculation process. Regarding the determination of the type and direction of the relationship, although SKS could identify a positive relationship, their understanding of the strength of the relationship tended to be limited. They had difficulty explaining why the relationship was weak even though it was mathematically valid. Furthermore,

SKS experienced difficulty interpreting data, as indicated by the tendency to state numerical results without providing an in-depth explanation of their meaning in a real-world context.

### 3.1.4 Interview Results for Students with Low Ability (SKR)

The following presents the interview results for a student with low ability (SKR).

**Interview Transcript with SKR**

P: "How did you feel when you were asked to calculate the correlation coefficient between production volume and defects?"

SKR: "I do not really understand correlation. I just followed the calculation steps in the book, but I do not understand what the result I obtained means. I only know that it is a number."

P: "What do you think about the relationship between production volume and defects based on the calculation you performed?"

SKR: "Honestly, I do not know much about it. I only know that it is a positive relationship, but I do not know whether that means that if production is higher, there will also be more defects. I do not understand what the number means in real life."

P: "Did you experience difficulty using the linear regression model for prediction?"

SKR: "Yes, I do not know how to make predictions using a regression model. I only followed the formula, but I do not know why I have to calculate it or what it is useful for."

Students in this category experienced substantial difficulties in almost all aspects of statistics, beginning with correlation calculations, where SKR failed to calculate the correlation value correctly or made errors in organizing the data. In interpreting correlation, when asked to explain the direction or strength of the relationship, they tended to be confused or mentioned relationships that were irrelevant to the context. In applying linear regression, SKR in this category could not use the linear regression model correctly and could not explain the meaning of the resulting prediction.

## 3.2 Discussion

The interviews conducted with three students with different levels of statistical thinking ability revealed several important findings that illustrate the difficulties students faced in understanding and applying contextual

concepts of correlation and linear regression. Students' understanding differed according to their ability levels. Students with High Ability (SKT) were able to calculate the correlation coefficient and relate it to real-world contexts. They were also able to explain that although there was a positive relationship between production and defects, the relationship was weak and did not indicate a strong association. This shows that students with high ability were able to use their statistical understanding to explain calculation results contextually, as described by Mat Nor Azmay et al. (2023), who state that statistical thinking requires students to understand data within a broader context. This is consistent with the findings of Hidayati et al. (2022), who reported that students with high statistical literacy scores were able to read data presented in tables or diagrams.

Furthermore, Students with Medium Ability (SKS) demonstrated a basic understanding of correlation calculations but had difficulty interpreting the results and relating them to other factors that might affect the outcome. These students showed that although they were able to follow computational procedures, they were less able to draw relevant conclusions from the data. Students with Low Ability (SKR), meanwhile, faced substantial difficulties in understanding the concepts of correlation and regression. They could not explain the meaning of the calculation results or connect them to real-world contexts, indicating that applying statistical techniques in real-world contexts remains a major challenge for students with low ability.

Another difficulty experienced by students involved interpretation and contextual application. As stated by Ritmayanti et al. (2025), although students can follow basic statistical procedures, they often have difficulty interpreting the results and applying them in real situations. This was clearly evident among Students with Medium Ability (SKS) and Students with Low Ability (SKR), who, although able to follow procedural steps, were unable to connect them to real-world contexts, such as the case of production volume and defects.

In addition, the contextual approach is important in statistics learning. Based on the interviews, it is clear that a contextual approach to statistics learning can help students connect theory with reality, as demonstrated by Students with High Ability (SKT). This approach is consistent with the findings of Hafiyusholeh et al. (2018), who reported that a contextual approach can improve students' statistical literacy by making learning more relevant and applicable. A contextual approach has also been shown to improve students' thinking abilities because it connects the material being taught with problems from students' everyday lives, thereby making mathematics learning more relevant and easier for students to understand (Rahmawati & Fadli, 2023). The findings of this study reinforce the need for changes in statistics teaching strategies. At present, learning that focuses on computational procedures needs to be complemented by contextual approaches that involve more analysis and

real-life application. This will also help students develop higher-order thinking skills (HOTS), which are essential for addressing data-driven challenges in the future.

## 4. CONCLUSION

Based on the research findings, it can be concluded that the statistical thinking ability of Grade XI students in solving contextual problems on correlation and linear regression is still not optimal. Although some students can perform calculations correctly, they often have difficulty interpreting the results and relating them to real-world contexts. Therefore, statistics learning should place greater emphasis on conceptual understanding, contextual analysis, and the application of statistical models in real life to improve students' statistical thinking abilities.

## 5. ACKNOWLEDGMENTS

The authors would like to thank the students of class XI MIPA 2 at MAN 1 Mataram for their time and participation in this study, which enabled the research to be carried out successfully.

## 6. RECOMMENDATIONS [font size 11pt]

Future research is recommended to develop and test more innovative contextual learning models to improve students' statistical thinking abilities more effectively.

## 7. REFERENCES

Aini, C. N., & Warli. (2025). *ANALISIS KEMAMPUAN BERPIKIR STATISTIK SISWA YANG BERGAYA KOGNITIF REFLEKTIF*. *4*.

Alem, D. D. (n.d.). *An Overview of Data Analysis and Interpretations in Research*.

Alfirawati Lumahu, Anetha L. F. Tilaar, & Santje M. Salajang. (2025). Pendekatan Kontekstual dalam Pembelajaran Matematika: Studi Eksperimentasi PMRI di Kelas VII SMP Negeri 3 Tondano: Penelitian. *Jurnal Pengabdian Masyarakat dan Riset Pendidikan*, *3*(4), 4546–4556. https://doi.org/10.31004/jerkin.v3i4.1245

Andriatna, R., Kurniawati, I., & Wulandari, A. N. (2021). PROFIL KEMAMPUAN LITERASI STATISTIK MAHASISWA CALON GURU MATEMATIKA. *FIBONACCI: Jurnal Pendidikan Matematika dan Matematika*, *7*(1), 19. https://doi.org/10.24853/fbc.7.1.19-28

Fitri, I., Setyaningrum, W., & Pulungan, D. A. (2023). FENOMENA LITERASI STATISTIK PADA PEMBELAJARAN MATEMATIKA SISWA SMA DI LHOKSEUMAWE ACEH. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *12*(2), 1927. https://doi.org/10.24127/ajpm.v12i2.7000

Hafiyusholeh, M., Budayasa, K., & Siswono, T. Y. E. (2018). Statistical Literacy: High School Students in Reading, Interpreting and Presenting Data. *Journal of Physics: Conference Series*, *947*, 012036. https://doi.org/10.1088/1742-6596/947/1/012036

Hidayana, R., & Lianingsih, N. (2025). [Contextual Learning as a Means to Improve Elementary School Students' Mathematical Literacy Skills]. *International Journal of Ethno-Sciences and Education Research*, *Vol. 5, No. 2*.

Hidayati, N. A., Prabowo, A., & Muharyati, T. (2022). *Literasi Statistik: Kemampuan Siswa SMP dalam Membaca Data* (Pt. 821-824).

Masjudin, M., Muzaki, A., Abidin, Z., & Ariyanti, I. A. P. (2020). Analysis of student's statistical thinking ability in understanding the statistical data. *Journal of Physics: Conference Series*, *1521*(3), 032063. https://doi.org/10.1088/1742-6596/1521/3/032063

Mat Nor Azmay, N. A., Rosli, R., Maat, S. M., & Mahmud, M. S. (2023). Educational Research Trends on Statistical Reasoning and Statistical Thinking: A Systematic Literature Review. *International Journal of Academic Research in Progressive Education and Development*, *12*(2), Pages 586-600. https://doi.org/10.6007/IJARPED/v12-i2/16921

Mohamad Hasim, S., Rosli, R., & Halim, L. (2024). A quantitative case study of secondary school students' level of statistical thinking. *Eurasia Journal of Mathematics, Science and Technology Education*, *20*(4), em2421. https://doi.org/10.29333/ejmste/14358

Prihastari, E. B., Sukestiyarno, S., & Kartono, K. (2022). Kajian Literasi Statistik pada Jenjang Pendidikan di Indonesia. *MENDIDIK: Jurnal Kajian Pendidikan dan Pengajaran*, *8*(2), 290–299. https://doi.org/10.30653/003.202282.250

Pujiati, A., Nusantari, D. O., & Nursa'adah, F. P. (2023). Literasi Statistik Mahasiswa Pendidikan Matematika ditinjau dari Resiliensi Matematik. *Prosiding Diskusi Panel Nasional Pendidikan Matematika*, 155–160.

Rahmawati, H., & Fadli, S. (2023). Prototype of Cartoon Concept Teaching Book Based on Contextual Approach to Train Students' Reflective Abstraction Skills. *Jurnal Kependidikan: Jurnal Hasil Penelitian Dan Kajian Kepustakaan Di Bidang Pendidikan, Pengajaran Dan Pembelajaran*, *9*(4), 1345. https://doi.org/10.33394/jk.v9i4.9660

Ritmayanti, I. R., Turmudi, T., & Dasari, D. (2025). Kemampuan Literasi Statistik Siswa dalam Pembelajaran Abad 21: Systematic Literature Review dan Bibliometric Analysis. *Juring (Journal for Research in Mathematics Learning)*, *8*(2), 127. https://doi.org/10.24014/juring.v8i2.37382

Takaria, J., & Talakua, M. (2018). THE ABILITY OF STATISTICAL LITERACY STUDENT TEACHER CANDIDATE IN TERMS OF PRIOR-ABILITY ON MATHEMATICS. *Jurnal Kependidikan: Penelitian Inovasi Pembelajaran*, *2*(2), 395–408. https://doi.org/10.21831/jk.v2i2.18768