*BRIDGE: The Multidisciplinary Research Portal*
*Vol. 4, No. 1 (2025), pp. 349-358 | p-ISSN: 3025-6585, e-ISSN: 3025-3640*
*https://jurnal.sttnlampung.ac.id/index.php/bridge*

# Statistical Literacy and Grouped Data Analysis Using SPSS in Undergraduate Education

***Radha Krisnamurti Sigamura[1], Dhian Nurul Istiqomah[2]**
[1,2]Fakultas Teknik, Universitas Musi Rawas, Sumatera Selatan, Indonesia
E-mail: *[1]radhakrisna964@gmail.com* *Corresponding Author

*Article History: Received: November, 15th 2025; Accepted: January, 18th 2026; Published: January, 31th 2026*

**ABSTRACT**

This study investigates undergraduate students' understanding of grouped data analysis using SPSS within an interdisciplinary framework integrating statistics education, computer science, and educational psychology. Employing a descriptive qualitative design, the study involved 18 computer science students who had completed coursework in grouped data computation. Data were collected through semi-structured interviews, classroom observations, and document analysis, and analyzed using iterative coding and thematic categorization. The findings reveal a gradient of understanding: 55.56% of students demonstrated high-level conceptual–procedural integration, 33.33% exhibited moderate understanding characterized by procedural competence with partial conceptual articulation, and 11.11% showed limited conceptual and interpretative ability. The results indicate that operational proficiency in statistical software does not automatically ensure statistical literacy. Students with higher understanding were able to justify class interval construction, interpret frequency distributions, and contextualize SPSS outputs within research applications. Conversely, lower-level understanding was associated with procedural dependency and limited interpretative reasoning. These findings reinforce contemporary perspectives emphasizing that technology-enhanced learning must integrate conceptual explanation, computational execution, and reflective interpretation. The study contributes to the discourse on interdisciplinary statistical education by highlighting the importance of aligning computational fluency with conceptual reasoning in higher education contexts. Strengthening pedagogical scaffolding in software-supported statistics instruction may enhance students' readiness for quantitative research and improve the quality of undergraduate academic work.

**Keywords:** *statistical literacy, grouped data analysis, SPSS, technology-enhanced learning, interdisciplinary education*

## INTRODUCTION

The digital transformation of higher education has shifted statistics instruction from manual procedural approaches toward computational and data-analytic paradigms. Within this global context, statistical literacy and data literacy are increasingly recognized as essential 21st-century competencies that underpin critical thinking, evidence-based decision-making, and professional readiness across disciplines (Gal, 2019; Gould, 2017; Engel, 2017). This shift necessitates the integration of conceptual statistical understanding with the meaningful use of statistical software.

In computer science programs, mastery of grouped data computation extends beyond mathematical procedures to include computational proficiency in statistical software such as SPSS. Conceptually, grouped data represent frequency distributions organized into class intervals, typically used to summarize large datasets prior to further analysis. A sound understanding of distribution structure, class intervals, cumulative frequency, and measures of central tendency and dispersion forms the foundation for both descriptive and inferential statistical analysis.

This study is situated at the intersection of four interrelated domains: statistics as applied mathematics, computer science as computational data processing, education as a pedagogical framework, and educational psychology as a lens for understanding cognitive and motivational dimensions of learning. International scholarship emphasizes that effective statistics education requires the integration of statistical reasoning, data visualization, and purposeful technological engagement (Ben-Zvi & Garfield, 2018; Zieffler et al., 2019). However, empirical evidence indicates that the mere use of statistical software does not automatically enhance conceptual understanding when students focus solely on procedural execution without interpreting the underlying statistical meaning of the outputs (Chance et al., 2016; Batanero et al., 2016).

From a technology-enhanced learning perspective, analytical software should function as a cognitive tool for constructing conceptual understanding rather than as a mechanism for automated calculation (Schindler et al., 2017). Prior studies have identified a persistent gap between operational software competence and the ability to interpret statistical results, particularly among students with limited statistical literacy (Weiland, 2017; Ridgway et al., 2018). These findings highlight the need for explicit pedagogical integration between statistical theory and computational practice.

Beyond cognitive dimensions, motivational factors and self-regulated learning processes significantly influence students' success in mastering technology-based quantitative content (Panadero, 2017; Broadbent & Poon, 2015). Consequently, examining students' understanding of grouped data computation using SPSS should be approached as an interdisciplinary phenomenon encompassing conceptual, technological, and psycho-pedagogical dimensions.

Preliminary observations of students in the Computer Science Program, Faculty of Engineering, Universitas Musi Rawas, revealed substantial variation in their understanding of grouped data computation using SPSS. While some students were able to execute procedural analyses within the software environment, they demonstrated limited comprehension of the statistical principles underlying the generated outputs. This condition aligns with international findings suggesting that technology-supported statistical literacy requires systematic evaluation of students' conceptual understanding (Gould, 2017; Ben-Zvi & Garfield, 2018).

Grounded in this interdisciplinary framework, the present study aims to analyze students' levels of understanding of grouped data computation using SPSS through a descriptive qualitative approach. Conceptually, this study contributes to the advancement of technology-integrated statistical literacy, the alignment of statistics instruction within computer science education, and the development of pedagogical strategies that foster deeper, reflective engagement with quantitative data analysis.

## METHODS

### Research Design

This study employed a descriptive qualitative design to examine students' understanding of grouped data analysis using SPSS within a higher education context. A qualitative approach is appropriate for exploring how learners construct meaning, articulate reasoning, and interpret statistical outputs in authentic learning environments (Creswell & Poth, 2018; Merriam & Tisdell, 2016). The design enables in-depth examination of conceptual and procedural dimensions of statistical understanding rather than measurement of causal relationships.

The study is grounded in an interdisciplinary framework integrating statistics education, computational practice in computer science, and educational psychology. Prior research in statistics education emphasizes the importance of investigating students' statistical reasoning processes and technology-mediated learning experiences through qualitative inquiry (Ben-Zvi & Garfield, 2018; Zieffler et al., 2019). This perspective is particularly relevant when evaluating how statistical software supports or fails to support conceptual understanding (Chance et al., 2016).

### Participants and Context

The participants were 18 undergraduate students enrolled in the Computer Science Program, Faculty of Engineering, Universitas Musi Rawas. The cohort consisted of eight male and ten female students who had completed coursework involving grouped data computation and SPSS-based statistical analysis. Purposive sampling was applied to ensure that participants possessed prior exposure to statistical concepts and hands-on experience with SPSS, which is consistent with qualitative research practices aiming to obtain information-rich cases (Merriam & Tisdell, 2016).

The institutional context represents a technology-oriented learning environment in which statistical software is integrated into research-oriented coursework. Such environments align with contemporary models of technology-enhanced learning in higher education, where digital tools function as cognitive instruments rather than merely computational utilities (Schindler et al., 2017).

### Data Collection Procedures

Data were collected through semi-structured interviews, classroom observations, and document analysis to ensure triangulation and enhance credibility (Nowell et al., 2017).

Semi-structured interviews were conducted to explore students' conceptual understanding of grouped data, including their ability to define class intervals, construct

*Vol. 4, No. 1 (2025), pp. 349-358 | p-ISSN: 3025-6585, e-ISSN: 3025-3640*
*https://jurnal.sttnlampung.ac.id/index.php/bridge*

frequency distributions, interpret cumulative frequencies, and explain SPSS-generated outputs. This format allows flexibility while maintaining alignment with core analytical themes and is widely used in studies examining statistical reasoning (Ben-Zvi & Garfield, 2018).

Classroom observations focused on students' engagement with SPSS during analytical tasks. Observational attention was directed toward procedural accuracy, error patterns, interpretative strategies, and alignment between conceptual explanation and software operation. Observational methods are particularly valuable in identifying discrepancies between stated understanding and enacted practice in technology-mediated learning (Weiland, 2017).

Document analysis included review of students' SPSS output files, frequency distribution tables, and coursework assignments. Documentary evidence provides objective artifacts of analytical performance and supports methodological triangulation (Bowen, 2009; Nowell et al., 2017).

**Data Analysis**

Data were analyzed using an interactive qualitative analysis framework involving data condensation, data display, and conclusion drawing (Miles et al., 2014). Interview transcripts, observation notes, and documentary artifacts were coded inductively to identify themes related to conceptual understanding, procedural competence, and interpretative ability.

The analytical process emphasized alignment between statistical reasoning and software-based execution, consistent with research frameworks in statistics education (Zieffler et al., 2019). Through iterative coding and cross-case comparison, students' understanding was categorized into three interpretative levels reflecting conceptual depth and analytical integration.

Low-level understanding was characterized by limited conceptual explanation and frequent misinterpretation of SPSS outputs. Moderate-level understanding indicated procedural competence accompanied by partial conceptual articulation. High-level understanding reflected coherent integration of conceptual explanation, accurate procedural execution, and appropriate interpretation of statistical results within research contexts. Such categorization aligns with models of progressive development in statistical literacy (Gal, 2019; Gould, 2017).

Cross-source comparison was conducted to ensure consistency across interview, observation, and documentary data. Analytical refinement occurred continuously to enhance coherence and explanatory validity.

**Trustworthiness**

Credibility was strengthened through methodological triangulation across interviews, observations, and document analysis (Nowell et al., 2017). Member checking was conducted to confirm interpretative accuracy, and peer debriefing was applied to reduce researcher bias. Thick description was employed to enhance transferability and allow readers to evaluate contextual applicability, consistent with established qualitative rigor criteria (Creswell & Poth, 2018).

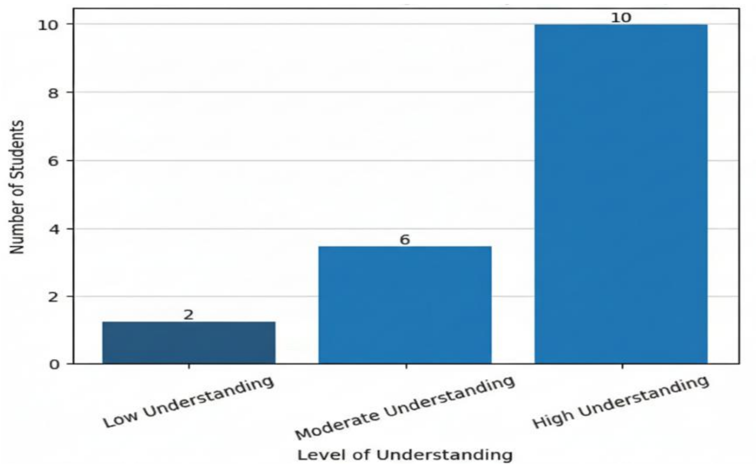
### Ethical Considerations

Participants were informed about the purpose of the study and provided voluntary consent prior to data collection. Confidentiality and anonymity were maintained throughout the research process. The study adhered to institutional ethical standards for educational research.

## RESULTS AND DISCUSSION

This study examined students' levels of understanding of grouped data analysis using SPSS within an interdisciplinary higher education context. The findings reveal variation not only in numerical distribution but also in conceptual depth, procedural competence, and interpretative reasoning. These differences reflect the extent to which students were able to integrate statistical concepts with computational execution.

Figure 1. Distribution of Students' Levels of Understanding of Grouped Data Analysis Using SPSS




As illustrated in Figure 1, ten students (55.56%) demonstrated high-level understanding, six students (33.33%) exhibited moderate understanding, and two students (11.11%) were categorized as having low-level understanding. Although the majority achieved strong conceptual–procedural integration, approximately 44.44% of students displayed partial or limited mastery, indicating that integration between statistical reasoning and software application remains uneven.

To clarify the structural differences across categories, Table 1 presents the conceptual indicators used in classifying students' levels of understanding.

Table 1. *Conceptual Indicators of Students' Understanding of Grouped Data Analysis Using SPSS*

| **Level of Understanding** | **Conceptual Explanation of Grouped Data** | **Procedural Use of SPSS** | **Interpretation of Statistical Output** | **Number of Students** | **Percentage (%)** |
| --- | --- | --- | --- | --- | --- |
| Low | Unable to clearly define | Frequent errors in | Misinterprets cumulative | 2 | 11.11% |

*Vol. 4, No. 1 (2025), pp. 349-358 | p-ISSN: 3025-6585, e-ISSN: 3025-3640*
*https://jurnal.sttnlampung.ac.id/index.php/bridge*



| | | | | | |
|---|---|---|---|---|---|
| | class intervals or frequency distribution concepts | generating tables and outputs | frequency and distribution patterns | | |
| Moderate | Able to define basic concepts but with partial explanation | Able to generate outputs correctly with minor guidance | Partially interprets outputs but lacks analytical depth | 6 | 33.33% |
| High | Accurately defines class intervals, frequency distribution, and statistical meaning | Accurately generates and navigates SPSS outputs independen tly | Correctly interprets outputs and connects to research context | 10 | 55.56% |

Table 1 demonstrates that differences across categories are not merely quantitative but conceptual and interpretative in nature. Students in the high-level category exhibited coherent integration between statistical explanation and computational implementation. They were able to justify class interval construction, interpret cumulative frequency distributions, and relate SPSS outputs to broader research contexts. This pattern reflects advanced statistical literacy characterized by conceptual–procedural alignment (Gal, 2019; Gould, 2017).

Students in the moderate category demonstrated operational fluency in SPSS but incomplete conceptual articulation. While they were capable of generating frequency tables and descriptive statistics accurately, they struggled to explain underlying statistical principles or critically interpret distributional characteristics. This finding supports previous research indicating a persistent gap between software proficiency and statistical reasoning (Chance et al., 2016; Weiland, 2017). When instructional emphasis prioritizes procedural execution, students may achieve technical correctness without deep conceptual understanding.

The low-level category highlights more fundamental difficulties. These students encountered challenges in identifying appropriate class intervals and misinterpreted cumulative frequencies and distribution patterns. Their performance reinforces concerns that statistical software, when used primarily as an automated calculation tool, may conceal conceptual weaknesses rather than resolve them (Ridgway et al., 2018). Effective technology-enhanced learning environments require that digital tools function as cognitive supports for reasoning, visualization, and interpretative dialogue (Schindler et al., 2017; Ben-Zvi & Garfield, 2018).

From an interdisciplinary perspective, the findings underscore the distinction between computational fluency and data literacy. Within computer science education,

**_BRIDGE: The Multidisciplinary Research Portal_**
_Vol. 4, No. 1 (2025), pp. 349-358 | p-ISSN: 3025-6585, e-ISSN: 3025-3640_
_https://jurnal.sttnlampung.ac.id/index.php/bridge_


students are typically comfortable navigating digital systems, yet such technical proficiency does not inherently guarantee statistical comprehension. International scholarship emphasizes that meaningful data analysis requires integration of algorithmic execution with conceptual reasoning about variability, distribution, and interpretation (Engel, 2017; Zieffler et al., 2019).

Motivational and self-regulated learning dimensions may further explain the observed variation. Students in the high-level group demonstrated reflective engagement during interviews and classroom observations, suggesting active monitoring of analytical decisions. This aligns with evidence that self-regulated learning strategies significantly influence performance in technology-mediated higher education settings (Panadero, 2017; Broadbent & Poon, 2015). Conversely, students in lower categories tended to focus on procedural completion rather than conceptual reflection.

Taken together, Figure 1 and Table 1 provide complementary evidence: the figure presents distributional clarity, while the table reveals structural conceptual differentiation. The results suggest that grouped data analysis represents a foundational threshold concept in undergraduate statistical education. Strengthening explicit pedagogical integration between statistical theory and SPSS-based practice may reduce disparities in understanding and foster more consistent development of statistical literacy in interdisciplinary academic programs.

## CONCLUSION

This study examined undergraduate students' understanding of grouped data analysis using SPSS within an interdisciplinary framework integrating statistics education, computer science, and educational psychology. The findings reveal a gradient of understanding characterized by varying degrees of conceptual–procedural integration. While more than half of the participants demonstrated high-level statistical literacy marked by coherent alignment between conceptual explanation, procedural accuracy, and interpretative reasoning, a substantial proportion exhibited partial or limited integration.

The results confirm that operational proficiency in statistical software does not automatically ensure conceptual understanding. Students who achieved higher levels of understanding were able to articulate the statistical meaning underlying class intervals, frequency distributions, and cumulative frequencies, rather than merely executing procedural commands. Conversely, students with moderate and low levels of understanding tended to rely on procedural familiarity without fully internalizing statistical reasoning.

These findings reinforce contemporary perspectives in statistics education that emphasize reasoning, interpretation, and contextualization as core components of statistical literacy (Gal, 2019; Gould, 2017). The study contributes empirical evidence to the ongoing discourse on technology-enhanced learning by demonstrating that digital tools such as SPSS must be pedagogically integrated to function as cognitive instruments rather than automated calculators (Ben-Zvi & Garfield, 2018; Schindler et al., 2017).

### Theoretical Implications

**_BRIDGE: The Multidisciplinary Research Portal_**
_Vol. 4, No. 1 (2025), pp. 349-358_ | _p-ISSN: 3025-6585, e-ISSN: 3025-3640_
_https://jurnal.sttnlampung.ac.id/index.php/bridge_

The study extends existing models of statistical literacy by highlighting the importance of conceptual–computational alignment in interdisciplinary higher education contexts. Within computer science programs, computational fluency may create an illusion of analytical competence; however, statistical reasoning requires explicit conceptual grounding. The findings therefore support theoretical distinctions between software proficiency and statistical literacy (Engel, 2017; Zieffler et al., 2019).

Furthermore, the results suggest that grouped data analysis functions as a threshold concept in undergraduate statistics learning. Mastery at this stage appears foundational for subsequent inferential reasoning and independent research competence. This conceptual positioning may inform future theoretical models linking descriptive statistics, computational tools, and advanced analytical development.

**Pedagogical Implications**

The findings indicate that statistics instruction in technology-oriented programs should incorporate explicit scaffolding that connects conceptual explanation, procedural execution, and interpretative discussion. Instructional strategies may include guided interpretation of SPSS outputs, reflective questioning, visualization-based reasoning, and structured dialogue on distribution patterns.

Educators should move beyond task completion models and adopt reasoning-centered pedagogies that encourage students to justify analytical decisions. Embedding metacognitive prompts and self-regulated learning strategies may further strengthen conceptual integration, given evidence linking reflective engagement to improved performance in digital learning environments (Panadero, 2017; Broadbent & Poon, 2015).

**Practical Implications**

At the institutional level, curriculum design should ensure balanced emphasis between computational training and statistical reasoning. Software workshops should be integrated with conceptual discussions rather than delivered as isolated technical sessions. Assessment strategies may also incorporate interpretative tasks that evaluate reasoning depth rather than procedural accuracy alone.

Strengthening this integration is particularly critical for undergraduate thesis preparation, where misinterpretation of grouped data may compromise research validity. By reinforcing conceptual clarity at the descriptive statistics stage, institutions may improve the overall quality of student research output.

**Limitations and Future Research**

This study was conducted within a single institutional context with a limited sample size, which may constrain generalizability. Future research may expand to multiple institutions or employ mixed-method approaches to examine longitudinal development of statistical literacy. Experimental designs investigating the impact of structured conceptual scaffolding on software-based statistical learning may further enrich the field.

## BIBLIOGRAPHY

Batanero, C., Chernoff, E. J., Engel, J., Lee, H. S., & Sánchez, E. (2016). Research on teaching and learning statistics. In D. Ben-Zvi & K. Makar (Eds.), _The_

*international handbook of research in statistics education* (pp. 1–24). Springer. https://doi.org/10.1007/978-3-319-66195-7

Ben-Zvi, D., & Garfield, J. (2018). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 3–15). Springer.

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40.

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies and academic achievement in online higher education learning environments: A systematic review. *Internet and Higher Education, 27*, 1–13. https://doi.org/10.1016/j.iheduc.2015.04.007

Chance, B., Ben-Zvi, D., Garfield, J., & Medina, E. (2016). The role of technology in improving student learning of statistics. *Technology Innovations in Statistics Education, 9*(1), 1–26.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal, 16*(1), 44–49.

Gal, I. (2019). Understanding statistical literacy: About knowledge of contexts and models. *Statistics Education Research Journal, 18*(1), 5–17.

Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal, 16*(1), 22–25.

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1), 1–13. https://doi.org/10.1177/1609406917733847

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8*, Article 422. https://doi.org/10.3389/fpsyg.2017.00422

Ridgway, J., Nicholson, J., & McCusker, S. (2018). Using technology to support statistical reasoning. *Technology Innovations in Statistics Education, 10*(1), 1–20.

Schindler, L. A., Burkholder, G. J., Morad, O. A., & Marsh, C. (2017). Computer-based technology and student engagement: A critical review of the literature. *International Journal of Educational Technology in Higher Education, 14*(1), 1–28. https://doi.org/10.1186/s41239-017-0063-0

Weiland, T. (2017). Problematizing statistical literacy: An intersection of critical and statistical literacies. *Educational Studies in Mathematics, 96*(1), 33–47. https://doi.org/10.1007/s10649-017-9764-5


Zieffler, A., Garfield, J., DelMas, R., & Reading, C. (2019). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal, 18*(1), 40–58.