**Journal on Mathematics Education**
Volume 15, No. 4, 2024, pp. 1335-1356

# Statistical literacy in primary education: An analysis of Indonesian fifth-graders' data interpretation and analysis skills

**Uswatun Hasanah[1,*], Susilahudin Putrawangsa[2], Farida Agus Setiawati[1], Edi Purwanta[1]**

[1]Faculty of Education and Psychology, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia
[2]Facutly of Education, University of Canberra, ACT, Australia
*Correspondence: uswatunhasanah.2022@student.uny.ac.id

Received: 7 July 2024 | Revised: 31 October 2024 | Accepted: 6 November 2024 | Published Online: 22 November 2024


## Abstract

Statistical literacy is an essential competency in today's data-driven society. However, research in this field often overlooks primary-level students, focusing predominantly on those in higher education. To address this gap, the present study evaluated the statistical literacy of 50 fifth-grade students in Indonesia, with a focus on their abilities in data analysis and interpretation. Utilizing a descriptive cross-sectional design, the study employed a written test grounded in real-world data scenarios, adapted from established theoretical frameworks. The assessment included various data representations—pictograms, bar charts, tables, and line charts—designed to evaluate students' skills in comprehension, comparison, and interpretation. The results indicate that the majority of students operated at an Idiosyncratic level, demonstrating limited capacity for data interpretation, while a smaller proportion reached the Transitional level; notably, no students attained Quantitative or Analytical proficiency. These findings underscore the urgent need for curriculum reforms to enhance both computational and analytical skills at an early stage, thereby equipping students with the foundational competencies required for advanced data reasoning. Future research should investigate pedagogical strategies that can bridge early-stage statistical literacy with the demands of higher-level statistical reasoning, ensuring students are better prepared for a data-intensive world.

**Keywords**: Data Analysis and Interpretation, Data Representation, Developmental Levels, Primary Education, Statistical Literacy

**How to Cite**: Hasanah, U., Putrawangsa, S., Setiawati, F. A., & Purwanta, E. (2024). Statistical literacy in primary education: An analysis of Indonesian fifth-graders' data interpretation and analysis skills. *Journal on Mathematics Education, 15*(4), 1335–1356. https://doi.org/10.22342/jme.v15i4.pp1335-1356

Statistical literacy is universally acknowledged as an essential competency in the contemporary, data-driven society, enabling individuals to critically interpret, assess, and make informed decisions based on statistical data (OECD, 2019). Within the realm of education, statistical literacy empowers students to comprehend and analyze data encountered in both academic and everyday contexts, thereby fostering the development of critical thinking and decision-making skills crucial for navigating complex and uncertain situations (Callingham & Watson, 2024; Gal, 2002; Mooney, 2002). As the global landscape continues to place increasing reliance on data for decision-making across diverse sectors, statistical literacy has emerged as a fundamental educational objective, particularly in the context of the rapidly advancing 5.0 era (Setiawan & Sukoco, 2021).

Consistent with global trends, Indonesia has incorporated statistical literacy into successive curricula, including the 2004, 2013, and 2016 KTSP, as well as the recent Merdeka Curriculum

(Mendikbudristek, 2022; Setiawan, 2019). In these frameworks, statistical literacy constitutes a core element of mathematics education, particularly within the domains of Data and Probability, equipping students with the skills necessary to effectively manage and interpret data. At the primary and intermediate levels, students are introduced to data collection, analysis, interpretation, and basic probability, with more advanced applications introduced at higher educational stages. This progression reflects Indonesia's dedication to cultivating a generation adept at engaging with data in both personal and professional settings, a competency increasingly vital in the global economy (OECD, 2019).

Despite these efforts, research on statistical literacy among primary school students in Indonesia remains limited (Setiawan, 2019; Setiawan, 2021). The majority of existing studies have concentrated on secondary and higher education, addressing the needs of older students while overlooking the development of foundational skills at the primary level. For example, studies by Kurnia et al. (2023) focus on statistical literacy assessments for junior and senior high school students, while others, such as those by Maryati and Priatna (2018) and Wildani et al. (2019), report low levels of statistical literacy among junior high school students. A review by Prihastari et al. (2022) identifies a significant gap in statistical literacy research for primary school students, underscoring the importance of studies aimed at this early stage of education to establish the foundational skills necessary for future academic growth.

Research on statistical literacy among primary school students is essential, as limited exposure to this skill during the early stages can hinder the development of more advanced statistical competencies in later years. Challenges in statistical literacy observed among secondary and tertiary students are often traced back to gaps in foundational education (Memisevic et al., 2017; Ow-Yeong et al., 2023; Siegler et al., 2012). Investigating statistical literacy in data analysis and interpretation at the primary level is particularly crucial, as these skills form the foundational building blocks for broader statistical understanding. Early engagement with data analysis enables students to interpret information, recognize patterns, and make informed decisions—skills that are increasingly vital in today's data-driven society (Ridgway et al., 2011). Research indicates that fostering these competencies in primary education helps students develop critical abilities for identifying patterns, drawing conclusions, and making predictions, all of which form the basis of advanced statistical reasoning and problem-solving (Franklin et al., 2005; Jones et al., 2000; Watson & Callingham, 2003). As emphasized by Gal (2002), statistical literacy should begin with foundational skills that allow students to meaningfully interact with data, such as reading, interpreting, and understanding basic data displays. These early skills provide the groundwork for the more complex statistical thinking and problem-solving abilities that students will develop as they progress in their education. By focusing on data analysis and interpretation in primary education, this study aims to address foundational gaps that, if left unaddressed, may limit students' capacity to engage with statistics in both academic and real-world contexts.

Therefore, this study investigates the statistical literacy levels of upper primary students, specifically those in grade five, with a focus on their abilities in data analysis and interpretation as outlined in the Indonesian curriculum. This includes skills related to data display, distribution, and interpretation. The primary research question guiding this study is: What are the statistical literacy levels of upper primary students in Indonesia, particularly in the areas of data analysis and interpretation?

The key indicators of statistical skills in the areas of data analysis and interpretation include the ability to describe, analyze, and interpret data (Jones et al., 2000). Describing data involves the capacity to read data displays, understand basic graphing conventions (such as titles and axis labels), and recognize when different displays represent the same dataset. It also entails evaluating the effectiveness of various displays in conveying data. Analyzing and interpreting data, on the other hand, involves comparing and combining elements within a dataset ('reading between the data') and making predictions or inferences ('reading beyond the data'). This also requires understanding the limitations of data displays and recognizing what information may be omitted or not explicitly shown.

While closely related, statistical literacy and statistical thinking are distinct concepts (Sabbag et al., 2018; Sharma, 2017). This study focuses specifically on statistical literacy rather than statistical thinking. In educational research, statistical literacy is defined as the foundational ability to understand, interpret, and critically evaluate statistical information in everyday contexts (Gal, 2002; Sharma, 2017; Watson & Callingham, 2003). In contrast to statistical thinking, which involves deeper analytical engagement with data processes (Sharma, 2017), statistical literacy emphasizes interpretive and evaluative skills that are crucial for young learners who are just beginning to engage with data in both academic and real-world settings (Gould, 2017; OECD, 2019; Sharma, 2017). This distinction is particularly relevant in primary education, where foundational skills in data description, grouping, and basic analysis provide the essential groundwork for future statistical understanding.

To assess the level of statistical literacy among primary school students, this study adapts the framework developed by Jones et al. (2000), which is particularly suited for evaluating young learners' foundational skills. Unlike the frameworks by Gal (2002) and Watson and Callingham (2003), which are designed for more advanced learners, Jones' framework specifically addresses the developmental stage of primary students, covering essential skills such as data description, grouping, representation, analysis, and critical thinking. By focusing on Jones' framework, this study investigates primary students' abilities to analyze and interpret data, with particular attention to the competencies of comparing and combining data—skills identified by Gould (2017) as critical for building statistical literacy.

Although originally intended for statistical thinking, the framework by Jones et al. (2000) is adapted here to assess the development of statistical literacy for three key reasons. First, it emphasizes fundamental data-handling processes—such as description, organization, representation, and analysis—that are the building blocks of statistical literacy. Second, the framework's four developmental levels (idiosyncratic, transitional, quantitative, and analytical) provide a structured approach to understanding how primary students' statistical literacy evolves, aligning well with educational objectives (Gal, 2002; Sharma, 2017; Watson & Callingham, 2003). Finally, the framework's focus on skills like pattern recognition, inference-making, and data interpretation aligns closely with the core components of statistical literacy, enabling a comprehensive examination of students' foundational competencies. Therefore, by adapting a framework originally designed for statistical thinking, this study offers a focused approach to measuring key interpretive skills, thereby establishing a strong foundation for students' future development of more advanced statistical understanding.

The four developmental levels outlined in Jones et al. (2000)—idiosyncratic, transitional, quantitative, and analytical—represent a progression of statistical skills aligned with cognitive development stages, illustrating how children engage with and comprehend statistical concepts. Each level builds upon the previous one, offering a framework for educators and researchers to assess and support students' growth from basic to more complex forms of data reasoning. In the context of statistical literacy, these developmental levels reflect how children's abilities to understand, interpret, and make judgments based on data evolve over time. Given that statistical literacy focuses on foundational skills required for reading, interpreting, and critically evaluating statistical information in everyday contexts (Gal, 2002; Sharma, 2017; Watson & Callingham, 2003), these levels provide a valuable lens through which to examine the development of these competencies in young learners.

By adapting Jones et al. (2000)'s framework, this study categorizes students' development of statistical literacy into four levels (see Table 1). At the idiosyncratic level, children's understanding of statistical information is limited and often subjective. In terms of statistical literacy, this means that they may not yet grasp the purpose of data displays or understand how to extract meaningful insights from them. For example, children might describe a bar chart by its color or size rather than by the quantities or categories it represents. At this stage, students are unable to engage with the data as informed consumers of information; they lack the foundational skills necessary to interpret or critically evaluate the data presented.

**Table 1**: Developmental levels in statistical literacy adapted from Jones et al. (2000)'s framework

| Level | Description | Key Characteristics | Progression from Previous Level |
|---|---|---|---|
| Idiosyncratic | Initial, unstructured response to data; little understanding of data conventions. | - Focus on unrelated or cosmetic features (e.g., colors, shapes).<br>- Limited or irrelevant interpretation of data. | First exposure to data, where understanding is intuitive and lacks structure. |
| Transitional | Basic understanding of data conventions; limited but emerging sense of structure. | - Recognizes some data conventions, such as titles or labels.<br>- Can interpret simple data points but with a narrow focus. | Begins to see data as structured information but still interprets in isolated parts. |
| Quantitative | Use of informal quantitative reasoning; more complete understanding of data conventions. | - Identifies trends, typical values, and uses basic measures (e.g., mode).<br>- Begins organizing and analyzing data logically. | Moves toward interpreting data critically, considering multiple aspects. |
| Analytical | Coherent, context-aware statistical reasoning; integration of complex data interpretations. | - Makes inferences, compares data, and sees broader context.<br>- Understands central tendency, variability, and trends. | Achieves full statistical literacy by critically evaluating and contextualizing data. |

As children progress to the transitional level, they begin to grasp basic elements of statistical literacy, such as recognizing that data displays convey information about categories or quantities. At this stage, students may be able to identify the highest or lowest categories in a chart but still lack a comprehensive understanding of the data as a whole. This level represents the early stages of statistical literacy, where students start to perceive data as structured information, although their focus tends to be on isolated aspects. They begin asking questions and making simple interpretations, yet their understanding remains fragmented, often limited to surface-level observations.

At the next stage, the quantitative level, students start to think more statistically, aligning more closely with the goals of statistical literacy. They use basic statistical measures, such as identifying the most frequent value (mode) or creating simple data representations, to make sense of data. This stage reflects a more advanced form of statistical literacy, as students are now capable of interpreting data with a developing critical perspective. They can discuss trends or typical values and begin to engage with concepts like central tendency. However, they may still struggle to apply these insights across different contexts or consider the deeper implications of the data.

Finally, at the analytical level, students exhibit the highest form of statistical literacy. At this stage, they can critically evaluate data, make informed inferences, and understand the broader context of data within real-world scenarios. For example, students might compare datasets, make predictions, or analyze trends over time, demonstrating a sophisticated understanding of what the data implies beyond its immediate values. This level embodies a mature statistical literacy, where students are not only consumers but also critical evaluators of data. They can interpret complex datasets, question the accuracy or sources of data, and understand how data informs broader decisions and contexts. These skills are essential for making informed judgments in everyday situations.

## METHODS

### Research Approach and Participants

This study adopted a descriptive research approach to assess the statistical literacy levels of primary school students, specifically focusing on their skills in data analysis and interpretation. Descriptive research was chosen for its ability to provide detailed insights into students' real-world competencies in understanding statistical information, in alignment with the study's aim of evaluating foundational skills in statistical literacy (Creswell, 2017; Fraenkel et al., 2011).

A cross-sectional design was employed to capture a snapshot of statistical literacy levels across four distinct data representations—pictograms, bar charts, tables, and line charts—providing insights into the developmental stages of statistical understanding at a specific educational level. This design enabled the identification of trends and competency levels among fifth-grade students, thus supporting the research's objective of mapping the progression of statistical literacy skills in primary education (Creswell, 2017; McMillan & Schumacher, 2010).

The study involved 50 fifth-grade students, aged 10 to 11, from a primary school in Indonesia. This age group was selected as fifth-grade students are expected to have acquired foundational knowledge in mathematics, including competencies relevant to data handling, as outlined in the Indonesian national curriculum. Informed consent was provided to all participants, and parental consent was obtained prior to the commencement of the study.

### Data Collection

Data were collected through a written descriptive test specifically designed to assess the statistical literacy of primary students, with a particular emphasis on data analysis and interpretation. This test format enabled students to engage with statistical information presented in real-life contexts, minimizing the need for additional explanations on how to interpret data displays. This approach ensured that the results accurately reflected the students' inherent literacy levels.

The selection of test item formats—pictograms, bar charts, tables, and line charts—was aligned with the current structure of the Indonesian primary school mathematics curriculum, which emphasizes these four types of data representation in teaching data presentation skills (Mendikbudristek, 2022; Setiawan, 2019). Additionally, the test items were adapted from established sources, including Australia's National Assessment Program – Literacy and Numeracy (NAPLAN), as well as instruments developed by Jones et al. (2000) and Callingham and Watson (2022), both of which are well-suited for assessing key indicators of statistical literacy in areas such as data description, analysis, and interpretation.

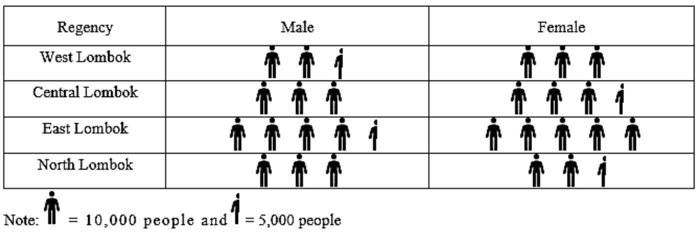
Each item was carefully designed to assess students' abilities to interpret and analyze data in formats they are likely to encounter in real-life contexts, reflecting everyday scenarios familiar to them. The test included varied visual data formats—pictograms, bar charts, tables, and line charts—to assess different aspects of data interpretation and comparison. For example, pictograms required basic interpretation and comparison skills, bar charts emphasized relational analysis, tables focused on data synthesis, and line charts assessed trend identification and comparison. A summary of the items and their respective purposes is provided in Table 2.

**Table 2.** Test items in English translation

| Item | Purpose and Source |
|---|---|

**1. Data analysis and interpretation in pictograms (A1)**

The following table shows the population of Lombok Island in four districts




**Which county has the least female population?**

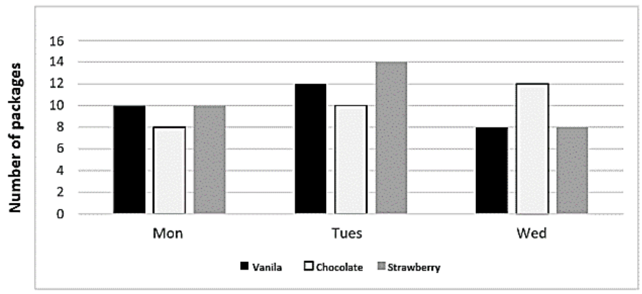
*Purpose and Source:* This test requires students to understand the data displayed in the form of pictograms and use their mathematical knowledge to make comparisons.

Note: This item is adapted from NAPLAN

---

**2. Analysis and interpretation of data in bar charts (A2)**

The following diagram shows the sale of three types of ice cream over three days.




**What was the total sales of chocolate-flavored ice cream on those three days?**

*Purpose and Source:* Students are given the task of analyzing and interpreting information from the data displayed in a bar chart. They are expected to be able to understand the relationship between data in depth, provide valid responses, and compare and integrate data comprehensively and completely.

Note: This item is adapted from Jones et al. (2000)

---

**3. Analysis and interpretation of data in table (A3)**

The following table illustrates a list of the four types of notebook packages sold in a day where each package contains a different number of books.

| Package Type | Book contents per package | Number of packages sold |
|---|---|---|
| Package A | 12 | 2 |
| Package B | 10 | 3 |
| Package C | 8 | 5 |
| Package D | 6 | 4 |

**What was the total number of books sold that day?**

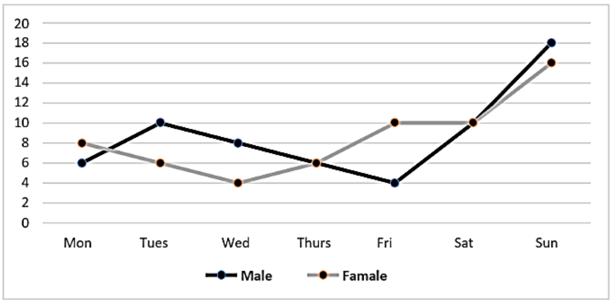
*Purpose and Source:* The students are expected to provide a valid and thorough response, drawing on their knowledge of relevant mathematical operations to answer the question. This question is designed to test students' ability to integrate information in a coherent and comprehensive manner.

Note: This item is adapted from NAPLAN

---

**4. Analysis and interpretation of data in line charts (A4)**

*Purpose and Source:* The students are expected to be able to interpret the data carefully through coherent data comparison and merging techniques, according to the

| | |
|---|---|
| The following graph shows the number of visitors to the park during the week   **What are the total visitors on weekends (Saturday and Sunday)?** | questions given. This means that they must be able to combine and compare data quantitatively, using their understanding of relevant mathematical concepts and mathematical operations to reach accurate conclusions.<br>Note: This item is adapted from Callingham and Watson (2022) |

The data collection process was conducted over a single day at the participating primary school, allowing for minimal disruption to the students' regular class schedule. The procedure followed these steps:

1. Step 1 (Preparation and Introduction): Before the data collection, the researchers provided the classroom teacher with a briefing about the study's objectives and procedures. This was followed by an introduction to the students, explaining the purpose of the study in an age-appropriate manner and reassuring them that their responses would be used solely for research purposes. Each student received a written consent form, signed by their parents or guardians, acknowledging their voluntary participation.
2. Step 2 (Test Administration): The written test was administered by their math teacher during a regular mathematics class period. Students were seated individually to prevent collaboration and to ensure that each student's responses reflected their independent understanding. Instructions were provided to the students, emphasizing the importance of carefully interpreting each question and answering to the best of their ability. The teacher clarified that the test was not graded and encouraged students to focus on understanding and analyzing the data presented.
3. Step 3 (Guidance on Test Items): The test consisted of four items, each focusing on a specific data visualization format—pictograms, bar charts, tables, and line charts. To reduce potential anxiety, students were reminded that they were familiar with similar materials from their curriculum and could use their mathematical skills as they would in their regular studies. No additional explanations of the visual data formats were given to avoid influencing the students' interpretations and to assess their natural statistical literacy skills as they would encounter in everyday contexts.
4. Step 4 (Time Allocation): Students were allocated approximately 20 minutes to complete the test. Students were allowed to work at their own pace, and any additional time needed by individual students was accommodated to ensure a comfortable testing environment.
5. Step 5 (Collection and Storage of Responses): Upon completion, students submitted their tests directly to the teacher and then to the researchers, who ensured all responses were collected and securely stored. Responses were then anonymized by assigning each student a unique identifier to maintain confidentiality. Data was securely stored in a password-protected digital format, accessible only to authorized research team members. Following the test, the teacher provided a short debriefing to the students, thanking them for their participation and briefly explaining how their responses would contribute to understanding and improving statistical literacy education.

This structured procedure ensured consistency in test administration and facilitated an environment in which students could demonstrate their statistical literacy skills authentically. The approach aimed to minimize testing stress, encouraging students to provide natural responses that accurately reflected their understanding of data analysis and interpretation.

## Data Analysis

Student responses were analyzed using a four-level statistical literacy scale adapted from Jones et al. (2000), categorizing responses as Idiosyncratic, Transitional, Quantitative, or Analytical. Each level reflects a progressive development in students' ability to interpret and reason with data, allowing for a detailed assessment of their statistical literacy in the areas of data analysis and interpretation. A more comprehensive description of these levels is provided in Table 3.

**Table 3.** Characteristics of statistical literacy levels (Adapted from Jones et al. (2000))

| Level | Characteristic | Sample Response |
|---|---|---|
| 1 (Idiosyncratic) | Invalid or incomplete response. Students do not give answers or give incorrect answers without explaining the calculation process. | On item A2 students replied, "I don't know," or "Chocolate ice cream is preferred on Tuesdays" or no response. |
| 2 (Transitional) | The response is valid but not appropriate. The student gives the correct answer but does not provide the calculation process from which the result is obtained, or the calculation process is correct but gives the wrong answer. | On item A2 students answer, "The amount of chocolate ice cream is about 30," without explanation, or "8 on Monday, 10 on Tuesday, and 13 on Wednesday." |
| 3 (Quantitative) | Students make several comparisons globally, meaning that students give correct answers, explain the calculation process in detail and can make simple comparisons even if they are not detailed or in-depth. | In item A2 the student answers, "The total is 30 (8+10+12). More chocolate ice cream was bought on Wednesday than on Monday, but less on strawberries on Tuesday." |
| 4 (Analytic) | Students can make comparisons globally with integrated and thorough analysis, meaning that students provide correct answers, explain the calculation process in detail and are able to make in-depth comparisons in detail or in depth. | On item A2 the student replied, "The total chocolate ice cream sold is 30 (8+10+12). This is less than strawberries which have a total sale of 32. Chocolate sold the most on Wednesday, which showed an increase in sales from the previous days. Chocolate sales were more stable than vanilla, which fell sharply on Wednesday." |

The data analysis process involved several systematic steps to categorize and interpret students' responses accurately, ensuring consistency and reliability in assessing statistical literacy levels. The procedure followed these steps:

1. Step 1 (Initial Data Preparation): Collected responses were anonymized and assigned unique identifiers to protect student confidentiality. Responses were then organized according to the four test items, enabling the researchers to examine students' answers within each data format—pictograms, bar charts, tables, and line charts.
2. Step 2 (Coding Scheme Development): A coding scheme based on the adapted Jones et al. (2000) framework was used to categorize responses into four levels of statistical literacy: Idiosyncratic,

Transitional, Quantitative, and Analytical. This coding scheme allowed researchers to assess each student's progression in data analysis and interpretation skills. Table 3 outlines the characteristics of each level, including criteria for categorization, which provided consistency in scoring across responses.

3. Step 3 (Independent Scoring by Assessors): Two independent assessors with expertise in statistical education including the main author of the current article were recruited to score the responses. Each assessor was trained on the coding scheme to ensure a thorough understanding of the characteristics defining each level. Assessors scored each response independently, without knowledge of the other's scores, to maintain objectivity.
4. Step 4 (Assessment and Categorization): Using the coding scheme, each student's response was categorized according to the highest level of statistical literacy demonstrated. For example, a response was rated at the "Analytical" level if the student showed integrated and detailed analysis, while a response demonstrating limited understanding with surface-level observations was categorized as "Idiosyncratic." Responses were scored on a per-item basis, allowing the researchers to assess patterns in statistical literacy across different data formats (pictograms, bar charts, tables, line charts) as well as the overall literacy level of each student.
5. Step 5 (Inter-Rater Reliability Calculation): After both assessors completed their independent scoring, inter-rater reliability was calculated using Cohen's Kappa to ensure the consistency of scoring. An agreement rate of 86% was achieved, indicating a high level of reliability between assessors. For responses with discrepant scores, assessors reviewed the responses together and discussed their rationale, reaching a consensus on the final categorization. This step ensured accuracy and objectivity in categorizing each student's statistical literacy level.
6. Step 6 (Quantitative and Qualitative Analysis): Quantitative analysis was conducted by calculating the frequency and percentage of students at each level (Idiosyncratic, Transitional, Quantitative, Analytical) across the four test items. This analysis provided insights into the overall distribution of statistical literacy levels within the sample. Meanwhile, qualitative analysis was performed to examine common patterns and specific examples within each level, providing richer insights into students' reasoning processes. Sample responses were selected to illustrate characteristic responses for each level, highlighting typical reasoning patterns among students at different levels of statistical literacy.
7. Step 7 (Data Interpretation and Reporting): Findings were synthesized to identify trends in students' statistical literacy. Results were organized according to the four developmental levels and discussed in terms of their implications for statistical literacy education. Anonymized examples from students' responses were included in the final report to provide concrete illustrations of each statistical literacy level. These examples served to clarify the differences between the levels and highlight areas for potential instructional improvement.

## RESULTS AND DISCUSSION

### Overview of Students' Statistical Literacy Levels

The analysis of statistical literacy levels among the 50 students who participated in the study offers valuable insights into their capacity to interpret and analyze data across four distinct developmental stages: Idiosyncratic, Transitional, Quantitative, and Analytical. These stages were assessed based on various data formats, including Pictograms, Bar Charts, Tables, and Line Charts. The assessment aimed to determine the distribution of students across these stages for each item type, with the findings summarized in Table 4, which presents the percentage of students at each statistical literacy level for each format.

**Table 4**. The percentage of students at each statistical literacy level for each item type

| Item Type | Idiosyncratic (%) | Transitional (%) | Quantitative (%) | Analytical (%) |
|---|---|---|---|---|
| Pictogram | 58 | 42 | 0 | 0 |
| Bar Chart | 64 | 36 | 0 | 0 |
| Table | 100 | 0 | 0 | 0 |
| Line Chart | 90 | 10 | 0 | 0 |

The results indicate that the majority of student responses fell within the Idiosyncratic level across all data formats, with particularly high proportions observed for the Table and Line Chart items. Notably, the Table item, which required students to calculate the total number of books sold from provided package data, proved to be especially challenging. In this instance, 100% of the students scored at the Idiosyncratic level, suggesting difficulty in integrating numerical data and performing basic arithmetic calculations effectively. Similarly, for the Line Chart item, which asked students to compare the visitor totals on weekends, 90% of students were placed at the Idiosyncratic level. This result underscores potential difficulties in interpreting trends over time and making quantitative comparisons between data points.

For the Pictogram item, which required students to identify the district with the lowest female population, 42% of students achieved the Transitional level. This suggests that nearly half of the participants were able to make simple comparisons based on the visual data, although their responses lacked the depth and detail expected at higher levels. Similarly, the Bar Chart item, which involved interpreting ice cream sales across three days and comparing quantities, revealed that 36% of students reached the Transitional level. This indicates a basic familiarity with bar charts, enabling students to perform elementary comparisons between data points, though they were unable to integrate the data in a comprehensive manner.

None of the students reached the Quantitative or Analytical levels across all item types. This absence suggests that the majority of students were unable to engage in higher-order analytical tasks, such as synthesizing multiple data points, drawing inferences, or identifying trends within the data. For example, in the Line Chart item, students faced difficulties in making meaningful comparisons of visitor totals over the weekend—an exercise that requires the ability to integrate information across multiple data points.

The distribution of student responses across the four item types reveals a notable gap in statistical literacy among the participants. Most responses were concentrated at the Idiosyncratic level, indicating a limited capacity to interpret, compare, and integrate data. Although some students attained the Transitional level in the Pictogram and Bar Chart items, their overall performance suggests insufficient proficiency in handling data across various formats. The absence of students at the Quantitative and Analytical levels underscores the need for targeted instructional support to develop these higher-order skills. This lack of advanced statistical reasoning highlights potential deficiencies in current educational practices, which may emphasize basic data handling skills while neglecting the cultivation of more complex analytical competencies (Callingham & Watson, 2022).

The statistical literacy gap observed in this study aligns with findings by Jones et al. (2000), particularly regarding challenges in data analysis and interpretation among young learners. These challenges are often attributed to limited exposure to statistics and the insufficient emphasis on statistical education in early schooling. This underscores the necessity of enhancing statistical instruction at the primary education level to establish a stronger foundation in statistical literacy. Strengthening the quality and depth of statistical learning will better equip students for more advanced analytical tasks and facilitate their ability to apply theoretical knowledge to real-world situations.

### Characteristics of Student Responses Across Statistical Literacy Levels

This study investigates the statistical literacy levels of primary school students, with a particular focus on their roles as data consumers, specifically in relation to data analysis and interpretation skills. The findings indicate a widespread difficulty with basic data interpretation, as the majority of students performed at the Idiosyncratic level across various data formats. A smaller proportion of students achieved the Transitional level, reflecting limited abilities in data comparison, while none demonstrated the advanced skills necessary to reach the Quantitative or Analytical levels. Given the absence of students proficient at the Quantitative or Analytical levels, the following section provides an analysis of the characteristics of student responses within the two observed statistical literacy levels: Idiosyncratic and Transitional.

#### *Level 1 (Idiosyncratic)*

The majority of students demonstrated idiosyncratic comprehension, characterized by responses that lacked depth and were often unfocused, resulting in mismatches between their answers and the information provided in the test. This level of understanding reflects significant challenges in processing and interpreting visual data, such as pictograms and graphs. These difficulties not only hinder students' ability to provide accurate answers but also highlight deficiencies in their broader skills in reading and analyzing data. The following section presents examples of student responses to the test items.

In the Pictogram item (Question A1 in Table 2), 29 out of 50 primary school students, or 58%, were classified at the Idiosyncratic level in their understanding of statistical literacy, particularly in analyzing and interpreting data presented through pictograms. Question A1 illustrates the population distribution of Lombok Island, divided into four districts, with each district represented by human figures in a pictogram. One full figure represents 10,000 people, while a half figure represents 5,000 people. The task required students to identify the district with the smallest female population. The responses varied significantly, indicating an idiosyncratic level of comprehension. For instance, a student identified as BY answered "15,000," which did not reflect an accurate understanding of the data representation (Figure 1a). Another student, NN, provided no answer, suggesting an inability to comprehend the question (Figure 1b). Conversely, SL made an attempt at calculation but still provided an inaccurate response, stating "the population of West Lombok: 2 and East Lombok: 4, so the total is 6" (Figure 1c).

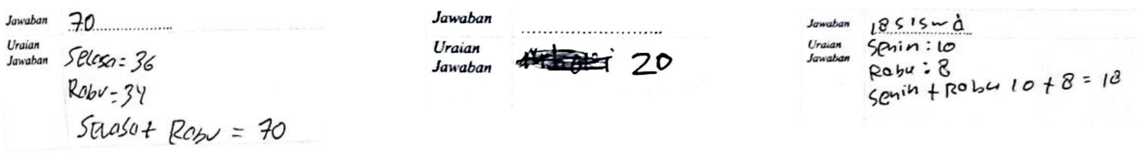

a b c

**Figure 1.** Examples of students' response on the Pictogram item

The difficulties encountered by students reflect common challenges in basic statistical literacy, where students frequently struggle to interpret data accurately and often make fundamental errors in reading visual representations (Jones et al., 2000). These errors in data interpretation often arise from an inability to recognize and process symbols or numerical information presented in non-standard formats, such as pictograms. Additionally, many students tend to overlook crucial information needed to answer questions correctly, such as failing to consider the gender specified in the question or making miscalculations based on the visual representation of figures. These issues highlight the need for more comprehensive educational approaches that strengthen students' visual interpretation and data analysis skills.

In the Bar Chart item (Question A2 in Table 2), 64% of primary school students exhibited difficulty in understanding and interpreting the data presented in a bar chart. This item revealed that students often provided inaccurate answers, reflecting a lack of depth in their understanding of analytical techniques and effective data manipulation. Many responses appeared to be based on initial impressions or misinterpretations of the data, indicating a failure to engage in critical analysis or follow a logical problem-solving process.

For instance, a student identified as SH provided a total of 70 for Tuesday and Wednesday (Tuesday = 36, Wednesday = 34), a figure that was not supported by accurate data or a correct calculation (Figure 2a). Another student, BN, responded with the irrelevant figure "20," without providing any supporting explanation or demonstrating a clear understanding of the data (Figure 2b). Meanwhile, SL mistakenly reported a total of 18 for ice cream sales on Monday and Wednesday (Monday = 10, Wednesday = 8), but incorrectly focused on vanilla ice cream, despite the question specifically asking for data on chocolate ice cream (Figure 2c). This error underscores the difficulty students experience in identifying and interpreting symbols or color codes in bar charts, which are crucial for accurately analyzing the data.

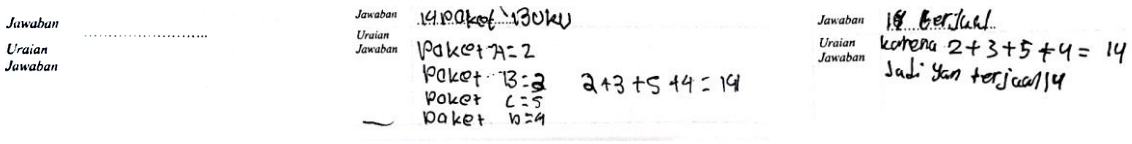



a b c

**Figure 2.** Examples of students' response on the Bar chart item

For data presented in tabular format (Question A3 in Table 2), all students failed to provide accurate and relevant responses. Tables, by nature, require a comprehensive understanding of data and the correct application of mathematical operations, which frequently presents challenges for students in reading and interpreting information accurately. Common errors included difficulties in correctly reading the data and neglecting to apply the necessary mathematical operations to perform precise and effective analysis.

For example, students such as LK, MN, and SS did not respond to the question, indicating a lack of competence or preparation for the task of table analysis (Figure 3a). In contrast, students such as FT and KM provided incorrect answers, reflecting a misunderstanding or lack of analytical ability (Figures 3b and 3c). They erroneously calculated the total sales of several packages—2 packages of A, 3 packages of B, 5 packages of C, and 4 packages of D—resulting in a total of 14 packages, without considering the number of books in each package. This error was significant, as it overlooked the actual book contents of each package—Package A containing 12 books, Package B containing 10 books, Package C containing 8 books, and Package D containing 6 books. A correct calculation would yield a total of 114 books sold, rather than just the number of packages. This mistake, which involves failing to recognize the relationship between variables and errors in mathematical application, underscores the pressing need to improve teaching strategies and evaluation methods in statistical literacy, particularly in terms of understanding and analyzing data presented in tables.

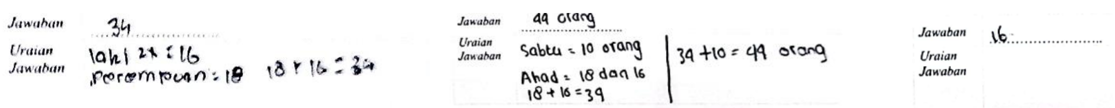



a b c

**Figure 3.** Examples of students' response on the Table item

Finally, with regard to data analysis presented through line charts (Question A4 in Table 2), most students encountered significant difficulties in providing accurate and valid responses. These challenges largely stem from a fundamental misunderstanding of both visual and numerical data, often resulting from incorrect or incomplete interpretations. Students tended to base their interpretations on subjective perceptions, which were not always supported by an adequate understanding of mathematical principles necessary for correct and thorough analysis. This suggests a gap in students' basic ability to interpret and utilize the information presented in line charts effectively.

A notable example of these difficulties can be seen in the responses of several students to this task. For instance, FT answered that the number of males was 16 and females was 18, resulting in a total of 34 people. While this answer appears logical, it fails to account for the complete dataset (Figure 4a). Conversely, RH provided a more detailed response, stating that on Saturday there were 10 people, and on Sunday there were 18 males and 16 females, totaling 44 people. However, this answer also demonstrates a misunderstanding, as it isolates the data by day without considering the broader context (Figure 4b). Meanwhile, NM gave an answer of "16," which lacked sufficient support from the provided data or a clear explanation of how the figure was derived (Figure 4c). This inappropriate response further underscores the need for focused educational interventions to improve students' ability to interpret line charts. In many instances, incorrect responses not only reflect errors in reading the data but also in connecting the information to the specific questions posed. This highlights the necessity for more intensive instruction to help students understand how to apply their mathematical knowledge in practical contexts and avoid common errors in interpreting visual and numerical data, particularly at the primary education level.




a b c

**Figure 4.** Examples of students' response on the Line chart item

In summary, the analysis of students' responses reveals significant challenges in their statistical literacy, particularly in the areas of data interpretation and analysis across various visual formats, such as pictograms, bar charts, tables, and line charts. The majority of students exhibited idiosyncratic comprehension, characterized by superficial or inaccurate responses that lacked depth and precision.

For instance, when interpreting pictograms, a large proportion of students misinterpreted population data, providing answers that demonstrated a fundamental misunderstanding of how visual data is represented. Similar difficulties were observed with bar charts, where students frequently misinterpreted quantities or failed to differentiate between categories. In the case of tables, none of the students accurately calculated totals due to errors in arithmetic and a limited understanding of table structures. The line chart task presented additional challenges, with students often making assumptions unsupported by the data or isolating elements without considering the full context.

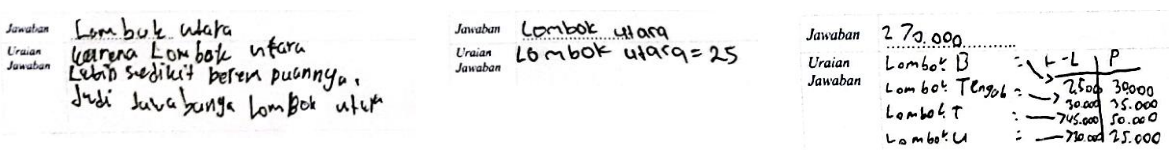
These findings emphasize the critical need to strengthen statistical literacy at the primary education level. The widespread difficulty students face in interpreting both visual and numerical data highlights the necessity for educational strategies that foster foundational skills in data literacy. Such strategies should focus on developing careful reading, contextual understanding, and analytical thinking. Incorporating these skills into the curriculum will enable students to draw meaningful inferences from data, thereby better preparing them for more advanced statistical reasoning and its practical applications in future learning stages.

***Level 2 (Transitional)***

Students at the Transitional level demonstrate a notable improvement in their ability to respond to data in a valid and detailed manner, signifying a shift from idiosyncratic thinking to more structured and analytical reasoning. At this level, students not only provide correct answers but also offer logical and evidence-based justifications for their responses. This represents a significant advancement, indicating that students are beginning to employ mathematical reasoning to effectively interpret and analyze data within various contexts. Below are examples of student responses that exemplify level 2 proficiency in data analysis and interpretation.

For instance, when presented with data in the form of pictograms (Question A1), students exhibited considerable variation in their understanding and application of statistical concepts, as illustrated by the responses of KM (Figure 5a), DD (Figure 5c), and NZ (Figure 5b).

> "*North Lombok, because in North Lombok the number of women is less*" (KM)
> "*North Lombok = 25*". (NZ)
> "*270,000, men in West Lombok are 25,000, Central Lombok 30,000, East Lombok 45,000, and North Lombok 30,000; women in West Lombok are 30,000, Central Lombok 35,000, East Lombok 50,000, and North Lombok 25,000.*" (DD)



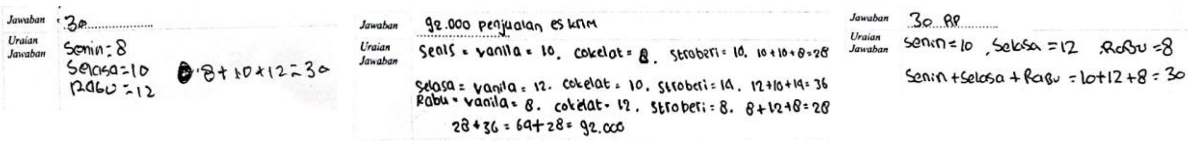

a b c

**Figure 5.** Examples of students' response on the Pictograms item

KM and NZ provided valid responses but omitted the calculation process they used, relying instead on a simple comparison based on the number of images. This suggests that while they have begun to use data to support their answers, their understanding remains at a transitional stage. They are able to make basic comparisons but have not yet fully grasped the more complex aspects of statistical reasoning. In contrast, DD, although his response was invalid in relation to the question, demonstrated a stronger grasp of mathematical operations and analytical processes through his clear explanation of the calculations. This indicates that DD possesses solid technical skills but may struggle with contextual understanding or accurate interpretation of data, which prevents him from providing a valid response.

In the case of data presented in the form of a bar chart (A2), the responses from FT, RH, and AD further illustrate their position at the transitional level (Level 2) in statistical reasoning. FT correctly stated the total ice cream sales for Monday, Tuesday, and Wednesday, providing a valid rationale for the calculations (Figure 6a). RH offered a more detailed response by calculating the total sales for all ice cream flavors—vanilla, chocolate, and strawberry—on each day. However, RH's failure to clearly differentiate between the ice cream flavors may cause confusion in the interpretation of the data (Figure 6b). AD, while providing a valid total, incorrectly referenced vanilla ice cream sales instead of chocolate, which was the focus of the question (Figure 6c). Although these students demonstrated basic competence in understanding the data, their responses highlight an ongoing struggle with more independent and complex data analysis. This indicates the need for further development of their statistical understanding and analytical reasoning skills.

"*30; Monday=8, Tuesday=10 and Wednesday=12 so 8+10+12=30*" (FT)
"*92,000; Monday=vanilla=10, chocolate=8, strawberry 10, 10+10+8=28; Tuesday=vanilla=12, chocolate=10, strawberry=14, 12+10+14=36; Wednesday=vanilla=8, chocolate=12, strawberry=8, 8+12+8=28 so the overall total is 28+36+28=92,000*" (RH)
"*30; Monday=10, Tuesday=12, Wednesday=8*" (AD)




a b c

**Figure 6.** Examples of students' response on the Bar chart item

Thirdly, a small group of 10% (5 out of 50 students), including DD, NZ, LL, AV, and KH, responded at the transitional level in the analysis of line chart data (A4). While these students demonstrated some variation in their ability to interpret and apply statistical knowledge, they were able to accurately calculate the total number of visitors (54 people). However, they struggled to make more nuanced comparisons, such as contrasting the number of male and female visitors across different days, despite having sufficient data for this task.

These limitations signal a shift from a basic understanding to a more integrated comprehension of statistical concepts. Although these students have begun to apply statistical reasoning in data analysis, they have not yet mastered the ability to draw more complex inferences or make meaningful comparisons between variables. This reflects the transitional phase in the statistical thinking framework described by Jones et al. (2000), where students start applying statistical methods but have not yet fully developed these skills for broader application. The following section presents the results of students' responses to the analysis and interpretation of data in relation to A4 in Table 2.

The responses that are valid but lack comparative analysis reflect progress in statistical literacy, yet they also highlight the need for further development in both conceptual understanding and the application of statistical principles. For instance, DD (Figure 7a) accurately provides the total number of visitors by gender. However, NZ (Figure 7b) and LL (Figure 7c) present a breakdown by day without offering a detailed analysis of the differences between male and female visitors. On the other hand, AV (Figure 7d) and KH (Figure 7e) not only present data categorized by gender but also include comparisons across days, demonstrating more advanced analytical skills. These students go beyond simple calculations to explore the broader context of the data, showing an emerging ability to integrate and compare multiple data points.

"*50, where Male = 20 and Female = 34*" (DD)
"*54 visitors, Saturday = 10, Sunday = 18 and 16 (10+18+16=54)"* (NZ)
"*54 people, male=10 +18 and female=10+16*" (LL)
"*54, female = Sunday 16, Saturday 10 = 26 and male = Sunday 18, 10 Saturday = 28 to 26+28=54*" (AV)
"*54 visitors, Saturday = male 10 + female 10 = 20 and Sunday = male 18 + female 16 = 34 so 20 + 34 = 54*" (KH)

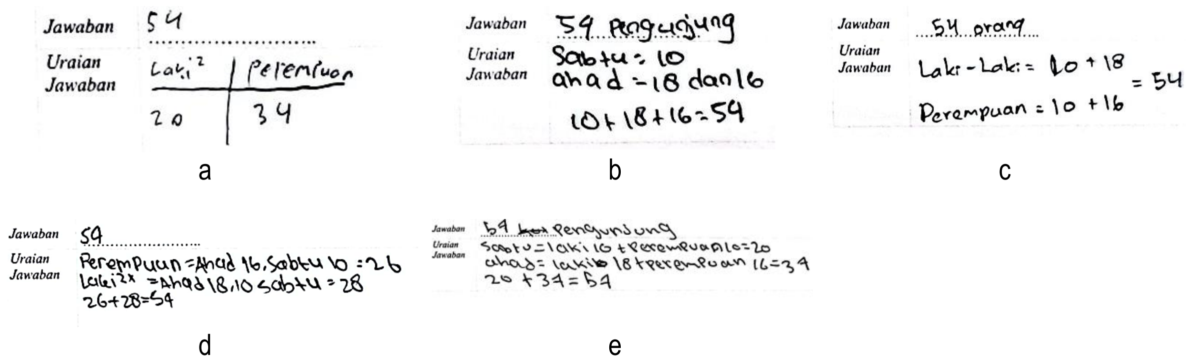



**Figure 7.** Examples of students' response on the Line chart item

The ability to delve deeper into data and understand its broader context is crucial in statistical reasoning, as emphasized in statistical literacy theory. AV and KH exemplify a shift toward a more analytical understanding, where data processing extends beyond simple numerical calculations to include deeper interpretations. This sets them apart from other students who remain largely anchored in an idiosyncratic level of understanding. Their approach underscores the importance of statistical education that not only focuses on computational skills but also fosters the ability to apply those skills in deriving meaningful and relevant inferences.

In summary, students at the Transitional level exhibit significant progress in statistical literacy, transitioning from basic, idiosyncratic responses to more structured, evidence-based answers. Their responses reveal an emerging capacity to interpret data, apply mathematical reasoning, and explain their answers logically. For example, in interpreting pictograms, some students were able to identify population differences through simple comparisons, although they often lacked the necessary calculations. When analyzing bar charts, students demonstrated a basic understanding of data manipulation, though occasional mistakes—such as confusing categories—indicated areas for improvement. Similarly, responses to line chart tasks showed developing skills in summarizing totals but were less effective in making comparative inferences.

These findings suggest that while these students are beginning to grasp fundamental statistical concepts, their skills remain in development, with notable gaps in their ability to perform comprehensive analyses and make comparisons. This transitional phase, as outlined in statistical literacy frameworks, highlights the need for targeted instructional strategies that enhance both conceptual and analytical skills, guiding students from basic data interpretation toward a more sophisticated understanding. By emphasizing both computational accuracy and contextual interpretation, educators can support these students in making meaningful, real-world inferences, ultimately fostering deeper statistical reasoning.

**Discussion and Implications**

The findings from this study underscore significant gaps in primary students' statistical literacy, particularly in their ability to interpret and analyze data presented in various formats. By examining students' responses through the framework of Idiosyncratic, Transitional, Quantitative, and Analytical levels, we uncover foundational challenges in statistical literacy that have implications for both instructional strategies and curriculum design in early education.

The predominance of Idiosyncratic responses across all data formats—Pictogram, Bar Chart, Table, and Line Chart—reveals students limited foundational skills in data comprehension and representation. This aligns with the work of Jones et al. (2000), who noted that early learners often struggle to interpret visual data, leading to surface-level or incorrect responses. Similarly, Callingham and Watson (2022) emphasize that without a solid grasp of basic data handling and representational skills, students are likely to falter when faced with tasks that demand deeper statistical understanding. The high frequency of Idiosyncratic responses—particularly the 100% at the Idiosyncratic level for Table items—illustrates the difficulties many primary students face in accurately extracting, interpreting, and responding to numerical data, especially when tasks require mathematical calculations. These findings are consistent with previous studies, such as those by Emilia and Amir (2022) and Purwati et al. (2022), which highlight that primary school students generally exhibit low levels of statistical literacy. This reinforces the need for the development of learning methods aimed at enhancing these skills, with the hope of addressing existing gaps in statistical literacy.

Additionally, the findings emphasize how limited exposure to data-handling tasks in early education can hinder students' ability to effectively interpret and synthesize data. Research by Memisevic et al. (2017) and Ow-Yeong et al. (2023) suggests that insufficient focus on data literacy in primary education contributes to students entering secondary school without essential data-handling skills. This study's results reinforce this perspective, as evidenced by students' struggles with basic tasks, such as integrating visual and numerical data in the Table and Line Chart items. In these tasks, students had particular difficulty calculating totals and interpreting data trends. As Jones et al. (2000) argue, building these skills early through consistent scaffolding and practice is crucial to helping students progress toward more advanced statistical thinking. This perspective is further supported by researchers like Khaerunnisa and Pamungkas (2017), Setiawan and Sukoco (2021), and Yotongyos et al. (2015), who stress the importance of understanding and analyzing statistical data from various sources as a foundation for more complex analyses later in education.

The progression observed among some students who achieved Transitional-level responses—42% in the Pictogram and 36% in the Bar Chart items—demonstrates initial advances in their statistical comprehension. These students exhibit a nascent ability to compare and draw inferences from data, marking an important developmental step. This aligns with the statistical literacy framework proposed by Jones et al. (2000), which describes students at the Transitional stage as beginning to use structured reasoning, even if their analytical depth remains limited. The ability of some students to make simple comparisons, albeit without comprehensive data integration, suggests they are gradually transitioning from Idiosyncratic to more structured thinking. However, as evidenced by their continued challenges in fully synthesizing or contextualizing information, there is a clear need for targeted support to help these students develop more nuanced statistical reasoning skills.

The complete absence of student responses at the Quantitative and Analytical levels across all item types signals an urgent need for instructional strategies that foster higher-order data interpretation and analysis skills. As Gal (2002) and Ridgway et al. (2011) emphasize, statistical literacy education should encourage students not only to perform computations but also to critically analyze data within its context. The limitations observed in this study indicate that students are primarily equipped to perform basic data manipulation but lack the training necessary to interpret data contextually, identify patterns, or make inferences. This highlights the importance of integrating real-world data scenarios into the curriculum—scenarios that challenge students to apply their understanding in meaningful ways. For instance, incorporating open-ended questions that encourage students to discuss data trends, make predictions, or draw connections between data points could promote deeper engagement and critical thinking, helping to build a stronger foundation for advanced statistical reasoning.

These findings carry significant implications for curriculum design, especially within the Indonesian educational context, where recent reforms in the national curriculum emphasize data literacy. Research by Setiawan and Sukoco (2021) and the OECD (2019) underlines the growing need for statistical literacy at all educational levels, particularly as data-driven decision-making becomes increasingly integral to both personal and professional domains. However, the gaps identified in this study suggest that current curriculum changes may not yet sufficiently address foundational statistical literacy skills. Despite intentions to incorporate data handling and probability at the primary level, students in this study demonstrated significant gaps in basic data interpretation. This indicates that further curricular adjustments are needed, particularly with regard to strengthening students' ability to contextualize and reason with data. By emphasizing these competencies from an early age, the curriculum can better equip students to engage with data effectively and make informed decisions as they advance in their studies.

## CONCLUSION

In response to the research question, this study offers critical insights into the statistical literacy levels of Indonesian primary school students, identifying substantial gaps in foundational knowledge and developmental needs in data interpretation and analysis. Employing a descriptive research design, the study assessed the statistical competencies of 50 fifth-grade students through a cross-sectional approach and a carefully designed test that included pictograms, bar charts, tables, and line charts. This methodology facilitated a comprehensive evaluation of students' real-world statistical literacy, revealing patterns in their understanding across various data formats.

Guided by the framework of statistical skill development proposed by Jones et al. (2000), the findings indicate that the majority of students demonstrated skills at the Idiosyncratic level, characterized by superficial responses and limited comprehension of visual data representations. A small proportion of students reached the Transitional level, exhibiting the ability to make basic comparisons but still facing challenges in interpreting and synthesizing data effectively. Notably, no students achieved the Quantitative or Analytical levels, highlighting a significant gap in higher-order analytical skills required for advanced data interpretation and informed decision-making.

These results underscore the urgent need for educational interventions that extend beyond basic data representation skills, incorporating scaffolding strategies that promote the development of more advanced analytical capabilities. This finding aligns with the broader goals of Indonesia's curriculum for data literacy yet highlights a disconnect between these objectives and the actual competencies of students. To address this gap, curricular reforms should prioritize the integration of both computational and critical reasoning skills, fostering a deeper engagement with data in real-world contexts.

The study, however, has certain limitations. It focused exclusively on one aspect of statistical literacy—data analysis and interpretation—while omitting other critical dimensions, such as data representation and description. Future research should incorporate a more comprehensive set of statistical literacy constructs to provide a fuller picture of students' capabilities. Furthermore, the study's narrow focus on a single age group and educational setting in Indonesia may limit the generalizability of its findings. The use of a written test format, while practical, may not fully capture the breadth of students' statistical reasoning abilities, as some students may benefit from oral or interactive assessment methods.

Despite these limitations, the study contributes valuable evidence highlighting the need for targeted curriculum adjustments and instructional support to strengthen the foundational elements of statistical literacy. Addressing these gaps through early and robust education in data interpretation and critical reasoning will better equip students to navigate an increasingly data-driven world and make informed decisions.

**Acknowledgments**

The authors wish to express their sincere gratitude to the students and teachers who participated in the data collection process. Special thanks are extended to our colleagues at Yogyakarta State University and the University of Canberra for their invaluable suggestions and ideas, which significantly contributed to the motivation and direction of this research. We also acknowledge the financial support provided by the Lembaga Pengelola Dana Pendidikan (LPDP) Indonesia, which made this study possible.

**Declarations**

| | | |
|---|---|---|
| Author Contribution | : | UH: Conceptualization, Writing - Original Draft, Editing and Visualization.<br>SP: Writing - Review & Editing, Formal analysis, and Methodology.<br>FAS: Methodology, Validation and Supervision.<br>EP: Validation and Supervision. |
| Funding Statement | : | This research was funded by The Indonesia Endowment Funds for Education (LPDP). |
| Conflict of Interest | : | The authors declare no conflict of interest. |
| Additional Information | : | Additional information is available for this paper. |


## REFERENCES

Callingham, R., & Watson, J. (2022). The development of statistical literacy at school. *Statistics Education Research Journal*, *16*(1), 181-201. https://doi.org/10.52041/serj.v16i1.223

Callingham, R., & Watson, J. (2024). Statistics education research at the school level in Australia and New Zealand: A 30-year journey. *Mathematics Education Research Journal*, *36*(1), 91-122. https://doi.org/10.1007/s13394-023-00470-0

Creswell, J. W. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE Publications, Inc.

Emilia, N. R., & Amir, M. F. (2022). Treffinger learning to enhance statistical literacy primary student. *JPI (Jurnal Pendidikan Indonesia), 11*(3), 522-532. https://doi.org/10.23887/jpiundiksha.v11i3.46115

Fraenkel, J., Wallen, N., & Hyun, H. (2011). *How to design and evaluate research in education* (Vol. 60). McGraw-Hill Education.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) report*. VA: American Statistical Association.

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review / Revue Internationale de Statistique*, *70*(1), 1-25. https://doi.org/10.2307/1403713

Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal*, *16*(1), 22-25. https://doi.org/10.52041/serj.v16i1.209

Jones, G. A., Thornton, C. A., Langrall, C. W., Mooney, E. S., Perry, B., & Putt, I. J. (2000). A framework for characterizing children's statistical thinking. *Mathematical Thinking and Learning*, *2*(4), 269-307. https://doi.org/10.1207/S15327833MTL0204_3

Khaerunnisa, E., & Pamungkas, A. S. (2017). Profil kemampuan literasi statistis mahasiswa jurusan pendidikan matematika Universitas Sultan Ageng Tirtayasa [Profile of statistical literacy skills of students majoring in mathematics education at Sultan Ageng Tirtayasa University]. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika,* *6*(2), 246-255. https://doi.org/10.24127/ajpm.v6i2.970

Kurnia, A. B., Lowrie, T., & Patahuddin, S. M. (2024). The development of high school students’ statistical literacy across grade level. *Mathematics Education Research Journal, 36*(Suppl 1), 7-35. https://doi.org/10.1007/s13394-023-00449-x

Maryati, I., & Priatna, N. (2018). Analisis kemampuan literasi statistis siswa Madrasah Tsanawiyah dalam materi statistika [Analysis of statistical literacy skills of Madrasah Tsanawiyah students in statistics material]. *Journal of Medives: Journal of Mathematics Education IKIP Veteran Semarang, 2*(2), 205-212. https://doi.org/10.31331/medives.v2i2.640

McMillan, J. H., & Schumacher, S. (2010). *Research in education: Evidence-based inquiry*, 7th Edition. MyEducationLab Series.

Memisevic, H., Biscevic, I., & Pasalic, A. (2017). Developmental trends in semantic fluency in preschool children. *Cogent Psychology*, *4*(1), 1403064. https://doi.org/10.1080/23311908.2017.1403064

Mendikbudristek. (2022). *Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia Nomor 16 Tahun 2022 Tentang Standar Proses pada Pendidikan Anak Usia Dini, Jenjang Pendidikan Dasar, dan Jenjang Pendidikan Menengah [Regulation of the Minister of Education, Culture, Research and Technology of the Republic of Indonesia Number 16 of 2022 on Process Standards in Early Childhood Education, Primary Education, and Secondary Education Levels]*. Kementerian Pendidikan dan Kebudayaan RI.

Mooney, E. S. (2002). A framework for characterizing middle school students' statistical thinking. *Mathematical Thinking and Learning*, *4*(1), 23-63. https://doi.org/10.1207/S15327833MTL0401_2

OECD. (2019). *PISA 2018 assessment and analytical framework*. https://doi.org/10.1787/b25efab8-en

Ow-Yeong, Y. K., Yeter, I. H., & Ali, F. (2023). Learning data science in elementary school mathematics: A comparative curriculum analysis. *International Journal of STEM Education*, *10*(1), 8. https://doi.org/10.1186/s40594-023-00397-9

Prihastari, E., Sukestiyarno, S., & Kartono, K. (2022). Kajian literasi statistik pada jenjang pendidikan di Indonesia [Study of statistical literacy at the education level in Indonesia]. *MENDIDIK: Jurnal Kajian Pendidikan dan Pengajaran*, *8*(2), 290-299. https://doi.org/10.30653/003.202282.250

Purwati, I., Murtiyasa, B., Kontesa, D. A., & Hakim, L. (2022). Analysis of learning difficulties of mathematics statistics materials in class V elementary school students. *Math Didactic: Jurnal Pendidikan Matematika, 8*(3), 272-280. http://dx.doi.org/10.33654/math.v8i3.1929

Ridgway, J., Nicholson, J., & McCusker, S. (2011). Developing statistical literacy in students and teachers. In C. Batanero, & Díaz, C. (Ed.), *Teaching statistics in school mathematics-challenges for teaching and teacher education* (pp. 311-322). Springer. https://doi.org/10.1007/978-94-007-1131-0

Sabbag, A., Garfield, J., & Zieffler, A. (2018). Assessing statistical literacy and statistical reasoning: The REALI instrument. *Statistics Education Research Journal*, *17*(2), 141-160. https://doi.org/10.52041/serj.v17i2.163

Setiawan, E. P. (2019). Analisis muatan literasi statistika dalam buku teks matematika Kurikulum 2013 [Analysis of statistical literacy content in mathematics textbooks Curriculum 2013]. *Pythagoras: Jurnal Matematika dan Pendidikan Matematika*, *14*(2), 163-177. https://doi.org/10.21831/pg.v14i2.28558

Setiawan, E. P. (2021). Literasi statistika dalam kurikulum matematika Sekolah Dasar (SD) 2004-2020: Tinjauan historis dan pengembangannya [Statistical literacy in the 2004-2020 primary school mathematics curriculum: A historical review and its development]. *Jurnal Pendidikan dan Kebudayaan*, *6*(1), 1-20. https://doi.org/10.24832/jpnk.v6i1.1915

Setiawan, E. P., & Sukoco, H. (2021). Exploring first year university students' statistical literacy: A case on describing and visualizing data. *Journal on Mathematics Education 12*(3), 427-448. https://doi.org/10.22342/jme.12.3.13202.427-448

Sharma, S. (2017). Definitions and models of statistical literacy: A literature review. *Open Review of Educational Research*, *4*(1), 118-133. https://doi.org/10.1080/23265507.2017.1354313

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, *23*(7), 691-697. https://doi.org/10.1177/0956797612440101

Watson, J., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, *2*(2), 3-46. https://doi.org/10.52041/serj.v2i2.553

Wildani, J., Triyana, I. W., & Mahmudah, W. (2019). Literasi statistik siswa kelas VIII SMP Negeri 1 Gresik [Statistical literacy of students in class VIII SMP Negeri 1 Gresik]. *Vygotsky: Jurnal Pendidikan Matematika dan Matematika, 1*(2), 99-110. https://doi.org/10.30736/vj.v1i2.137

Yotongyos, M., Traiwichitkhun, D., & Kaemkate, W. (2015). Undergraduate Students' Statistical Literacy: A Survey Study. *Procedia - Social and Behavioral Sciences*, *191*, 2731-2734. https://doi.org/10.1016/j.sbspro.2015.04.328