Al-Ishlah: Jurnal Pendidikan
Vol.16, 3 (September, 2024), pp. 3073-3088
ISSN: 2087-9490 EISSN: 2597-940X, DOI: 10.35445/alishlah.v16i3.5072

Assessing Data Literacy Competencies in Mathematics Among Junior High School Students in Yogyakarta City

Retno Ayu Trisnawati[1], Ali Mahmudi[2]

[1] Universitas Negeri Yogyakarta, Indonesia; retnoayu101@gmail.com
[2] Universitas Negeri Yogyakarta, Indonesia; alimahmudi@uny.ac.id

ARTICLE INFO

Keywords:

data literacy;
mathematics education;
junior high school students;
Yogyakarta city;
educational assesment

Article history:

Received 2024-02-15
Revised 2024-05-17
Accepted 2024-06-20

ABSTRACT

This research aims to assess students' data literacy skills, focusing on their ability to translate data from one format to another, comprehend data-related problems, analyze strategies for data-driven problem-solving, predict outcomes based on data, draw conclusions, and construct arguments grounded in data. The study employs a survey methodology with both quantitative and qualitative approaches. A total of 362 students from various educational levels in Yogyakarta's state junior high schools were surveyed and selected through stratified and proportionate random sampling techniques. Data were collected using a test instrument comprising six questions designed to measure different aspects of data literacy, alongside interview results. The students' data literacy was categorized into high, medium, and low levels. The findings indicate that a significant majority (95%) of students exhibited low data literacy, particularly struggling with drawing conclusions and constructing arguments based on data. Students with high data literacy skills successfully met all the indicators. Those with medium-level skills performed well on most indicators but faced difficulties in choosing strategies for data problem-solving and drawing conclusions, often due to insufficient time for deeper reflection. Conversely, students with low data literacy struggled to meet any of the indicators effectively.

Corresponding Author:
Retno Ayu Trisnawati
Universitas Negeri Yogyakarta, Indonesia; retnoayu101@gmail.com

1. INTRODUCTION

Currently, there are many important statistical data sets that have not been used properly due to a lack of ability to manage the data to draw appropriate conclusions (Morales et al, 2014). Data is an important element of information because it can provide a description of an event at a certain time (D'Ignazio & Bhargava, 2015). Data can be interpreted as information that can be used to analyze something that will later be used as a consideration for making a decision. There is a lot of data that we often find in everyday life. Examples of data that are usually found are daily weather data (temperature, humidity, rainfall, and wind speed and direction), data on people exposed to COVID-19, data on child marriage rates, data on sales, data on education in Indonesia and many more (BUMN, 2019). As a result, people must have the ability to make informed decisions based on these data (OECD, 2018). However, based on Kominfo (2021), the ability of the Indonesian population to process data to be able to draw

appropriate conclusions is still not in the adequate category. This results in the Indonesian population not being able to utilize existing data properly.

When an individual's capacity to analyze data and draw conclusions falls short of established standards, the researcher seeks to conduct a more thorough examination of this ability. This aligns with the findings of Batur and Baki (2022), who emphasized the necessity of assessing each individual's data processing skills. Ultimately, such an in-depth analysis is crucial for identifying areas of improvement and enhancing overall data interpretation accuracy.

The abilities being discussed refer to the skills that enable individuals to effectively manage presented data. One crucial ability in this context is data literacy. Individuals with data literacy skills can transform raw data into meaningful information, which in turn supports informed decision-making (Herzog, 2015). For example, consider the data on the spread of COVID-19 from January 29 and 30, 2022, where the number of people affected by the virus was 0.023% and 0.024% of the total population each day, respectively, compared to the World Health Organization's (WHO) recommended figure of 0.014% per day (Ministry of Health, 2022). Individuals with data literacy skills would analyze this information to assess the stability of COVID-19 cases and use it to make informed decisions, such as seeking out strategies to avoid contracting the virus, given that the infection rates were still above WHO's recommended levels. This perspective aligns with Dibekulu (2020), who suggests that individuals who analyze data have processed and understood its purpose. Conversely, those lacking data literacy skills may fail to recognize the significance of the COVID-19 data from January 30, 2022, and therefore may not take necessary precautions to mitigate the virus's spread, even when it remains above the WHO's recommended threshold (GLN, 2017).

Furthermore, in the context of education, the 2022 regulation by the Ministry of Education, Culture, Research, and Technology (Permendikbudristek No. 5) on Graduate Competency Standards for Early Childhood Education, Basic Education, and Secondary Education, explicitly states that data literacy skills empower students to identify and investigate information relevant to the issue at hand. These skills enable students to analyze alternatives for solving problems related to the presented data, draw inferences or conclusions, and formulate responses or arguments based on the data. According to Reeves and Hiang (2019), data literacy skills are essential in the learning process, as they involve conveying information by collecting, analyzing, and interpreting data. Therefore, students with data literacy skills can effectively read, comprehend, and utilize data to make informed decisions.

In addition, with data literacy skills, students can read data, work with data, analyze data, and make arguments based on data (D'Ignazio & Bhargava, 2016). Meanwhile, according to Schield (2011) data literacy skills are related to students' ability to represent data, understand data-related problems, analyze data, predict data, make conclusions based on data, make arguments based on data. In mathematics learning, data literacy skills really help students understand material related to statistics and data representation (Khan & Mason, 2021). This explanation is a benchmark for researchers in making question indicators and an overview of what kind of questions will be used.

Research by Fitri et al. (2023) indicates that students' data literacy skills remain in the low category. Similarly, Kartika et al. (2021) reported that data literacy skills in the Prambanan area are still low. Larasati et al. (2020) also found that students in Sleman, a district in the DIY Province, exhibit low data literacy skills. Therefore, it can be concluded that data literacy skills across various regions in Indonesia are generally still at a low level, highlighting a need for focused educational interventions to improve these essential skills.

The low data literacy skills of Indonesian students are influenced by several things. To find out more deeply why this happens, an in-depth analysis of the causes of low data literacy skills is needed (Khuan & Krauss, 2015). In addition, the location of the research must be considered again. The research location is Yogyakarta City. The reason is that Yogyakarta City is also the capital city of the Special Region of Yogyakarta province and information and data spread in the capital area more quickly, requiring the ability to monitor and control themselves effectively to use the data (Kireina, 2017). Additionally, urban areas generally have more access to digital infrastructure, educational resources,

and economic opportunities, which can contribute to higher levels of data literacy than rural areas (Aljassim & Ostini, 2020).

The city of Yogyakarta offers significant considerations for understanding data literacy, particularly in the context of its educational system. The curriculum implemented in schools has fully integrated digital-based learning with traditional methods, highlighting the importance of students becoming proficient in interacting with data accessible through various digital media and files. This includes the ability to convert and create graphics from one format to another (Perdana et al., 2019). Another critical factor is the educational level targeted by this research. Junior High School (SMP) is an appropriate level for identifying these abilities, as it aligns with Piaget's theory that children aged 11 and above enter the formal operational stage. At this stage, students can think abstractly, formulate multiple hypotheses when faced with problems, and engage in combinatorial and reflective thinking (Nursalim et al., 2007). This age group corresponds to junior high school students, who are capable of developing a deeper understanding of data concepts and enhancing their data literacy skills (Blake & Pope, 2008).

Additionally, according to Permendikbudristek No. 5 of 2022 on Graduate Competency Standards in Early Childhood Education, Primary Education, and Secondary Education, junior high school students are expected to reason and apply concepts, strategies, processes, facts, and mathematical tools to solve problems related to their personal lives, immediate environment, and broader community. Batur and Baki (2022) argue that these expectations align well with aspects of data literacy. Moreover, the scope of problems faced by junior high school students is quite broad, as they often deal with data circulating within their community environment (OECD, 2018). Therefore, it is appropriate to consider junior high school students as a suitable group for initial identification in analyzing data literacy skills, given their emerging complex thinking abilities.

Research indicates that globally, skills in analyzing and processing data among junior high school students require greater attention to make statistical learning more meaningful (Hafiyusholeh, 2015). These insights have motivated researchers to investigate the data literacy skills of junior high school students in Yogyakarta City. The current study differs from previous research by offering more in-depth analysis, not only examining students' data literacy abilities in terms of specific aspects or indicators but also providing a foundation for identifying areas that require further development in the data literacy skills of junior high school students in Yogyakarta City.

2. METHODS

This type of research is a survey using quantitative and qualitative approaches. The use of a quantitative approach is important because this research involves data obtained from test scores and questionnaires, and the conclusions of the research on the selected sample can be generalized to the research population. While the use of qualitative methods aims to understand events or phenomena and explore the meaning behind these phenomena.

Quantitative data analysis describes the percentage of students' test results. Apart from that, this analysis also uses confidence intervals by paying attention to the Z score. The reason for using the Z score is to describe the relationship between a value and the average of a group of values (Samosir et al, 2022). The qualitative data analysis used is an interactive technique, which consists of three stages of data reduction, data presentation, and drawing conclusions (Utomo, 2021).

The population for this study consists of ninth-grade students, with the research specifically targeting public schools. Public schools were chosen because they have uniformly implemented the zoning policy (PSKPKemendikbudristek, 2020). The selection of schools was conducted based on predetermined strata, which were determined by ASPD scores. The classification of these school strata is presented in Tables 1 and 2.

Table 1. Formula for Determining School and Test

Interval	Strata
$x_i \geq x + 0{,}5s$	High
$x - 0{,}5s \leq x_i < x + 0{,}5s$	Middle
$x_i < x - 0{,}5s$	Low

(Garvin & Ebel, 1980)

Table 2. Criteria for Determining School Strata

Interval	Strata
$61 \leq x_i \leq 79$	High
$49 \leq x_i < 61$	Middle
$0 \leq x_i < 49$	Low

The sampling technique used in this research is multistage sampling, which involves two distinct methods: stratified random sampling and proportionate random sampling. Stratified random sampling is a technique that selects samples based on specific strata or levels, while proportionate random sampling ensures that all members of the population have an equal chance of being selected, proportionate to the predetermined sample size. Using the Slovin formula, it was determined that the minimum required sample size is 358 students, and the actual sample size used in this study is 362 students.

$$n = \frac{N}{1+Ne^2}$$

Information:
n = Sample size
N = Population Size
e = Critical value tested (error tolerance, selected e=0.05)

$$n = \frac{3393}{1+(3393)(0{,}05)^2}$$

$$\boldsymbol{n = 357,8 \approx 358}$$

These students were taken from 6 public junior high schools in Yogyakarta City. The list of samples for this study is presented in Table 3.

Table 3. Sample size based on school strata

Strata	Size
High	126
Middle	119
Low	117
Total	362

In addition to quantitative data, qualitative data was also collected for this study. This qualitative data was obtained through interviews with three subjects: one with high data literacy skills, another with moderate skills, and the third with data literacy skills in the low category. The primary instrument used in this study is a data literacy test comprising six items. The indicators utilized in this study are presented in Table 4.

Table 4. Indicators of data literacy

Indicator	Item	Score
Representing data from one form to another	4	0-2
Understanding the problem related to the data presented	1	0-2
Analyzing strategies to solve data-related problems	2	0-2
Predicting data to solve problems	5	0-3
Making conclusions based on data	6	0-3
Making arguments based on data	3	0-2

The other instrument used was an interview guide. The interview is conducted in an unstructured manner so that the researcher's questions will adjust to what will happen in the field. Even though the interview was an unstructured interview, there were main questions that had been prepared, but as the interview time progressed, additional questions would be adjusted to the student's answer to the previous question or based on the answer to the test.

The assessment criteria used to assess data literacy skills are presented in Table 5.

Table 5. Assessment criteria

Percentage	Criteria
90% ≤ N ≤ 100%	Very high
80 %≤ N < 90%	High
70% ≤ N < 80%	Middle
60% ≤ N < 70%	Low
N < 60%	Very low

The following is the formula for determining the N value (Percentage of overall value).

$\boldsymbol{N = \frac{R}{SM} \times 100\%}$

R = number of scores obtained

SM = maximum score

Quantitative data analysis techniques use confidence intervals. Inference analysis serves to generalize the results obtained from the selected sample in making predictions of a phenomenon within the scope of the population. The concept of (1-α)100% confidence interval for the population mean, and standard deviation is known, so the formula for finding the confidence interval is:

$$\boldsymbol{\bar{X} - Z_{\alpha/2} \frac{s}{\sqrt{n}} < \mu < \bar{X} + Z_{\alpha/2} \frac{s}{\sqrt{n}}}$$

Description:

$\bar{x}$ = sample mean

Z = z value at n-1 free degree

s = sample standard deviation

n = sample size

Qualitative analysis consists of the stages of reduction, presentation, and conclusion drawing based on data (Utomo, 2021). Data reduction was carried out by researchers, namely, by simplifying and sorting the results of students' data literacy skills tests, which focused on important data to produce meaningful information and make it easier to present data and draw conclusions for the population in this study. Data presentation is done by displaying the results of students' work on the data literacy skills test and then compiling a narrative text explaining the students' test results. The final stage of drawing conclusions is done by comparing the results of student answers and interviews to describe and explore students' data literacy abilities.

3. FINDINGS AND DISCUSSION

3.1 Quantitative Analysis

The results of data literacy skills analysis in various school strata are presented in Table 6.

Table 6. Analysis of data literacy skills in each school stratum

Description	Stratum			Total
	High	Middle	Low	
Average	36,79	28,69	26,13	30,68
Standard deviation	18,18	13,54	13,47	15,97
Ideal highest score	100	100	100	100
Highest score	85,71	85,71	85,71	85,71
Ideal lowest score	0	0	0	0
Lowest score	0	0	0	0
Sample size	126	119	117	362

Based on the analysis presented in Table 6, it is evident that the data literacy skills of Indonesian students remain in the low category across different school strata. This conclusion is drawn from the average performance of students on data literacy assessments, which consistently falls short, regardless of whether the students attend high, medium, or low-strata schools. These findings highlight a pervasive challenge in developing data literacy skills, underscoring the need for targeted interventions and enhanced curriculum strategies to address this critical gap in education.

The results of the analysis of data literacy of junior high school students in Yogyakarta City have an average of 30.68 with a standard deviation of 15.97. By using the interval estimation formula, it is obtained that with a significance of 95%, the data literacy ability of junior high school students in Yogyakarta City lies between 28.71 and 32.34 which is included in the "low" category. Supported by the results of the percentage analysis of data literacy skills which illustrates that 96.93% of students in Yogyakarta City have data literacy skills in the low category, 1.66% in the moderate category and 1.38% in the high category. This is in line with the research results of Fitri et al (2023), Larasati et al (2020) and Kartika et al. (2020) which state that students' data literacy skills in several regions of Indonesia are still in the "low" category.

Next, analyze students' answers to each question item.

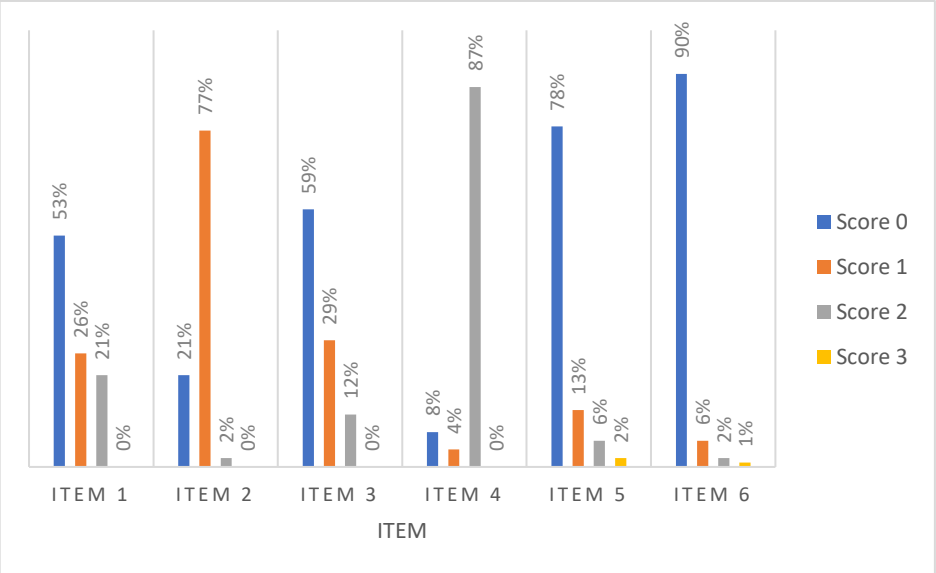


Figure 1. Representation of students' answers to each question item

Based on Table 8, we can discuss the data literacy abilities of junior high school students in Yogyakarta City, focusing on each specific indicator or item. For the first question, it was found that 53% of students were unable to fully understand the problems related to the presented data, 26% were able to grasp the issues but with some inaccuracies, and only 21% of students correctly understood the problems in their entirety. Regarding the second item, 21% of students struggled to analyze and apply the correct strategy for solving the problem, while 77% managed to analyze and attempt the correct strategy, though not completely accurately, and just 2% of students proficiently analyzed and solved the problem using the correct strategy or representation.

For the third item, 59% of students were unable to formulate an argument based on the data presented, 29% could formulate an argument but failed to substantiate it, and only 12% were able to both determine and prove their arguments based on the data. The fourth item revealed that 8% of students could not represent data in various forms, 4% could represent data but with some errors, and a significant 87% of students successfully represented data in various formats with complete accuracy. This finding contrasts with the research of Mendez-Carbajo (2020), which suggested that students' ability to visualize data requires more attention.

The items discussed here have an ideal maximum value of 2. According to the mathematics textbook by Savitri et al. (2022), data visualization skills are introduced as early as the seventh grade. Consequently, students are expected to have a good understanding of how to visualize data by this stage.

For item 5, 78% of students could not interpret or predict the data used to solve the problem, 13% of students could only use a small part of the concept to process the data, 6% of students could use most of the concepts to process the data but not yet correct as a whole and 2% of students could use predict the data used and all the concepts used to solve the problem. For item 6, 90% of students could not draw conclusions correctly, 6% of students could only determine a small part of the process used to make conclusions, 2% of students could determine most of the process but had not yet reached the conclusion stage and 2% of students could determine all the processes needed so that they could draw conclusions correctly.

3.2 Qualitative Analysis

The following are the results of the qualitative analysis of students' data literacy skills. Where there are 3 subjects in this study. Starting from the first subject. The first subject has data literacy skills in the high category. The first subject can achieve all data literacy indicators well; in other words, the first subject can answer data literacy questions well. As a sample of student work, this can be reviewed from the results of student work on questions related to indicators of predicting data to solve problems.

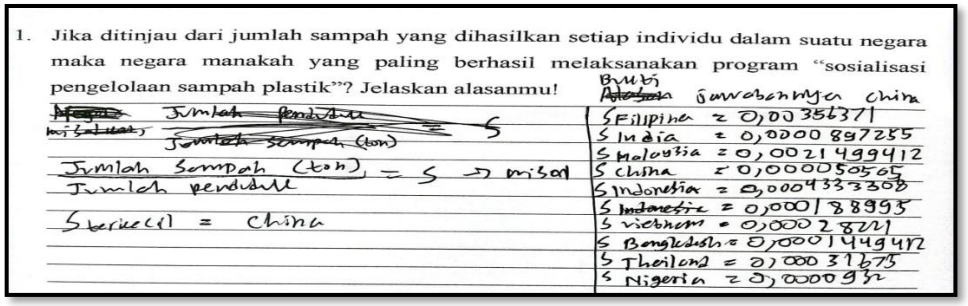


Figure 2. First Subject's Answer on Problem Item Number 1

Based on Figure 2, it is evident that the student was able to understand the problem related to the presented data and grasp its significance. By applying the concept of averages, the first subject identified China as the most successful country in implementing the "socialization of plastic waste management" program. During the interview, the subject explained their approach to solving the problem, which involved calculating the average amount of waste produced by each individual in each

country. The country with the lowest average waste per person was deemed the most successful in implementing the plastic waste management program. Additionally, the first subject did not encounter significant difficulties in solving the problem. Although initially confused when reading the question, the subject was able to find the correct answer after gaining a deeper understanding. Another notable observation was that the first subject frequently practiced literacy-related questions, which contributed to their familiarity and comfort with such tasks.

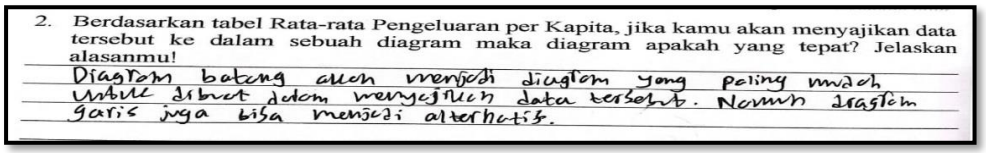

2. Berdasarkan tabel Rata-rata Pengeluaran per Kapita, jika kamu akan menyajikan data tersebut ke dalam sebuah diagram maka diagram apakah yang tepat? Jelaskan alasanmu!
Diagram batang akan menjadi diagram yang paling mudah untuk dibuat dalam menyajikan data tersebut. Namun diagram garis juga bisa menjadi alternatif.

Figure 3. First Subject's Answer on Problem Item Number 2

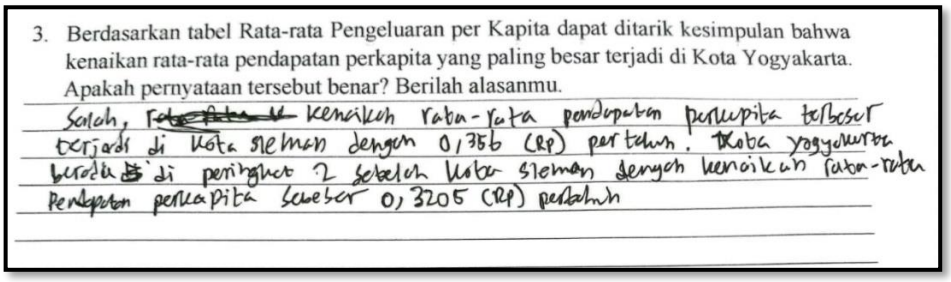
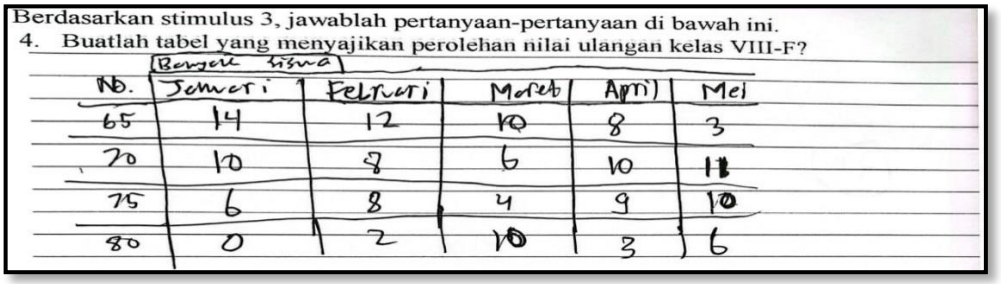

3. Berdasarkan tabel Rata-rata Pengeluaran per Kapita dapat ditarik kesimpulan bahwa kenaikan rata-rata pendapatan perkapita yang paling besar terjadi di Kota Yogyakarta. Apakah pernyataan tersebut benar? Berilah alasanmu.
Salah, ~~Rata-rata~~ kenaikan rata-rata pendapatan perkapita terbesar terjadi di Kota Sleman dengan 0,356 (Rp) pertahun. Kota yogyakarta berada di peringkat 2 setelah kota sleman dengan kenaikan rata-rata pendapatan perkapita sebesar 0,3205 (Rp) pertahun

Figure 4. First Subject's Answer on Problem Item Number 3

Based on Figure 4, it is known that the The first subject always believes that he is able to make the right decision and prove the correctness of the decision made. Strengthened by the results of the interview, it is known that the first subject first looked for the difference in average per capita expenditure each year in the five regencies/cities in DIY province. After that, find the average of these differences. The first subject answered that the statement in the question was wrong because based on the explanation from the first subject, the district with the largest increase in average per capita income was Sleman.

.


Berdasarkan stimulus 3, jawablah pertanyaan-pertanyaan di bawah ini.
4. Buatlah tabel yang menyajikan perolehan nilai ulangan kelas VIII-F?

Banyak siswa

No.	Januari	Februari	Maret	April	Mei
65	14	12	10	8	3
70	10	8	6	10	11
75	6	8	4	9	10
80	0	2	10	3	6

Figure 5. Subject's Answer on Problem Item Number 4

Based on Figure 5, it is known that the The first subject can change the representation of data presented in bar chart form into table form. This is in line with filling out the first subject's questionnaire which shows that he is confident that he can represent data in various forms of data presentation and one of them is a table. This was strengthened by the results of the interview with the first subject who stated that he was confident that he could change the data representation into various forms of data presentation by paying attention to matters related to data presentation such as the month, grades and number of students adjusted to the month and the amount of value obtained. The first subject felt very

proficient in working on question number 4 because he had encountered this question very often since grade 7, so the first subject had no problems working on this question.

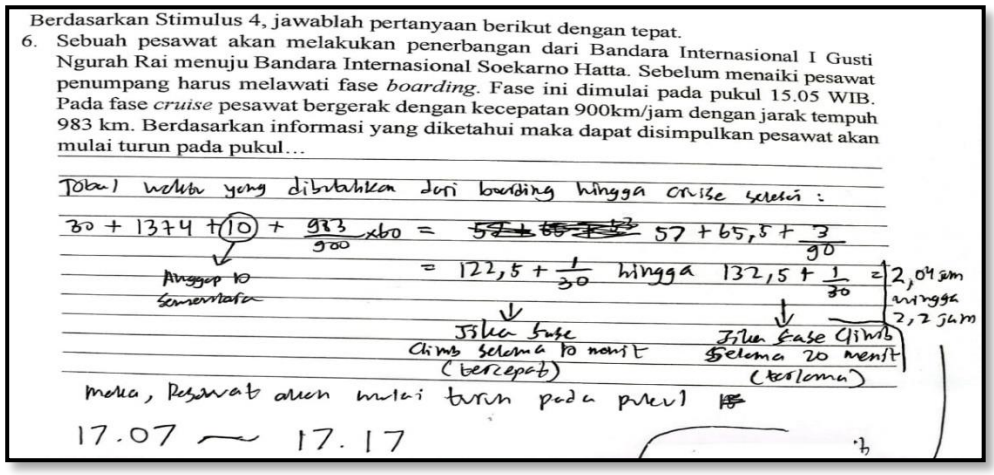


Figure 6. First Subject's Answer on Problem Item Number 6

Based on Figure 6, it is known that The first subject was very confident that he could answer the question even though it took a little longer than doing the other questions. The interview results also stated that important information that can be used to make conclusions is the parked/boarding, take off & departure, climb and cruise phases. For the time in the climb phase, the minimum and maximum times are found and then added to the cruise phase time obtained from the speed concept. Because the time in the climb phase consists of a minimum and maximum time, the final result is a time span consisting of a minimum time and a maximum time as well. Apart from that, first subject is confident that he can use non-statistical concepts to answer these questions and this is in line with the way students use the concept of speed to answer the questions presented before drawing conclusions based on the data presented.

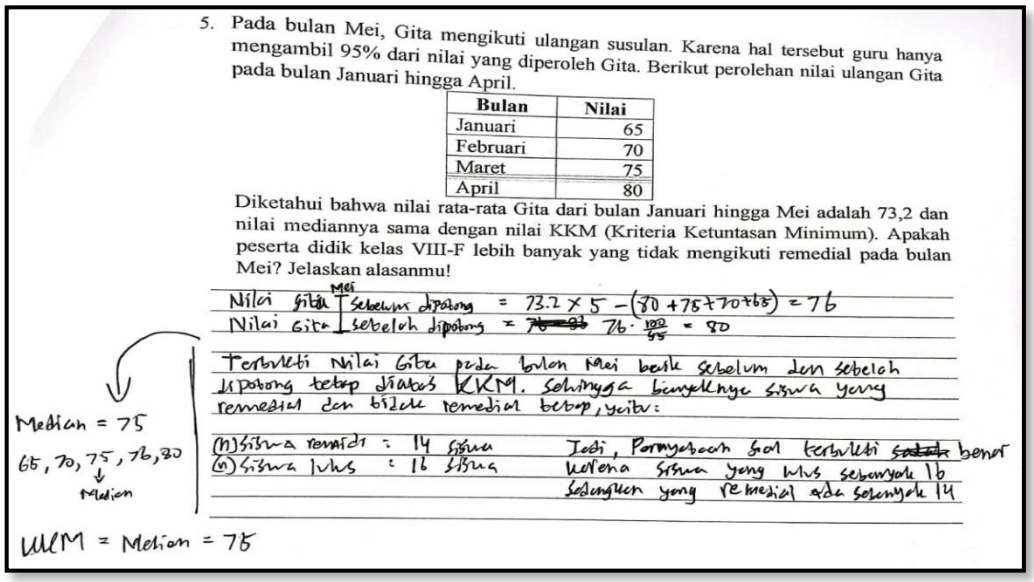
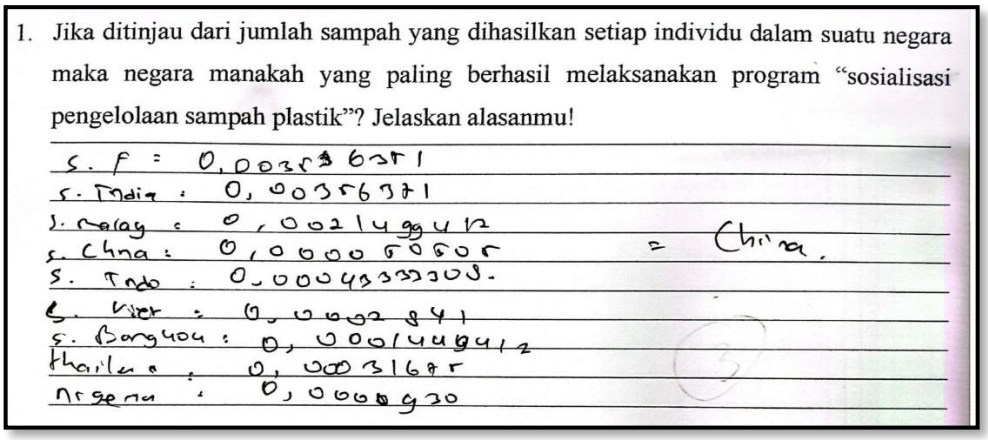


Figure 7. Results of students' answers who have high data literacy skills

Based on Figure 7, it is known that the first subject understood the right strategy to solve the problem. The strategy used to solve the problem is to apply statistical concepts in the form of mean,

median, and mode, as well as the concept of numbers in the form of percentages. Reinforced by the results of the interview with the first Subject who stated that he knew the concepts used and could apply these concepts to solve the problem. Where to determine unknown data, namely Gita's score in May using the concept of average, determining the KKM with the concept of median and using the concept, determining Gita's actual score with the concept of number (percentage) and the concept of mode to find out whether more students take remedial or not. The first subject can also connect the information listed in question number 5 and the information listed in stimulus 3 so that they can solve the problem correctly. Therefore, it can be concluded that students who have high data literacy skills can achieve all indicators of data literacy skills and can also explain the process in answering the questions presented. Although initially, the first subject had difficulty determining the average used, the first subject was able to solve this problem.

Furthermore, the second subject. The second subject is a student who has data literacy skills in the moderate category. Based on the results of the analysis of student answers, it is known that students who have sufficient data literacy skills cannot achieve some indicators well.


1. Jika ditinjau dari jumlah sampah yang dihasilkan setiap individu dalam suatu negara maka negara manakah yang paling berhasil melaksanakan program "sosialisasi pengelolaan sampah plastik"? Jelaskan alasanmu!

S. F = 0,0035[illegible]6371
S. India : 0,00356371
S. Malay : 0,0021499412
S. China : 0,0000050505 = China.
S. Indo : 0,00043333303.
S. Viet : 0,0002841
S. Bangkok : 0,0001448412
Thailand : 0,0003168
Nigeria : 0,0000920

Figure 8. Second Subject's Answer to Problem Item Number 1

Based on Figure 8, it is known that the second subject can understand the meaning of the data presented and understand the problem of question number 1. Based on the results of the interview, it is known that the answer was obtained by first finding the amount of waste obtained by each individual by dividing the amount of waste by the population of each country. Since each resident of China produces the least amount of waste, China is the most successful in implementing the plastic waste management program.

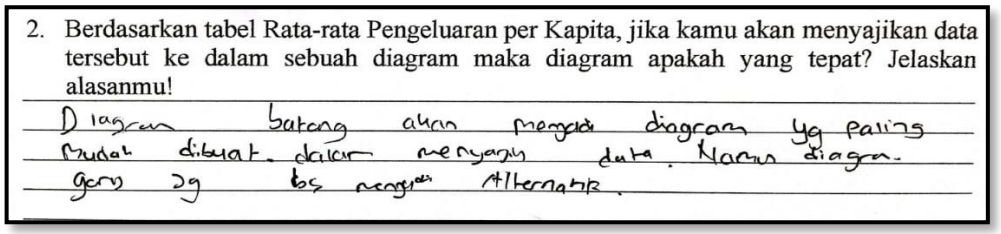

2. Berdasarkan tabel Rata-rata Pengeluaran per Kapita, jika kamu akan menyajikan data tersebut ke dalam sebuah diagram maka diagram apakah yang tepat? Jelaskan alasanmu!

Diagram batang akan menjadi diagram yg paling mudah dibuat. dalam menyajikan data. Namun diagram garis yg tdk menjadi alternatif.

Figure 9. Second Subject's Answer to Problem Item Number 2

Based on Figure 9, it is known that the second Subject can determine the right diagram to determine the data presented but has not been able to determine the exact reason why he chose the diagram. The results of the interview conducted with the second Subject stated that the reason for choosing the diagram was because the bar chart presented the data per year more clearly. This reason is still not correct because the reason intended by the question must be adjusted to the form and

condition of the data presented. In addition, there are student obstacles in answering question number two, namely difficulty in determining the reasons why they have this data representation.

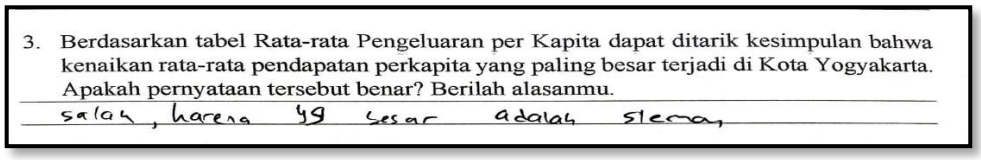


Figure 10. Second Subject's Answer on Problem Item Number 3

Based on Figure 10, it is known that the second subject can make arguments and prove his arguments correct. The results of the interview which stated that he was sure that he could answer the question or could make an argument appropriately and he could prove his argument by first finding the difference in average per capita income per year in each Regency / City in DIY Province after that he found that Sleman Regency had the biggest difference so he could prove that the argument he made was correct.

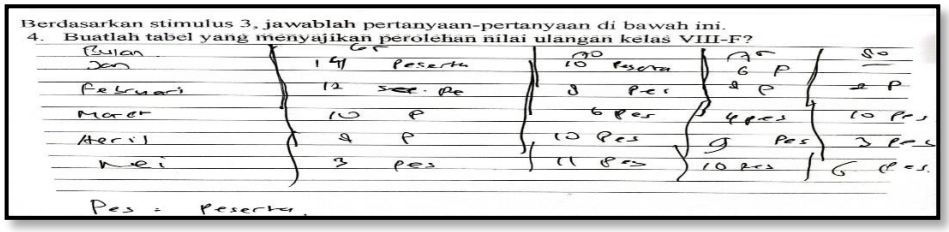


Figure 11. Second Subject's Answer on Problem Item Number 4

Based on Figure 11, it is evident that the second subject was able to convert the presented data from a bar chart into a table. During the interview, the second subject revealed that they could successfully represent the data in table form and found it relatively easy to change the data representation by identifying the colors associated with each value.

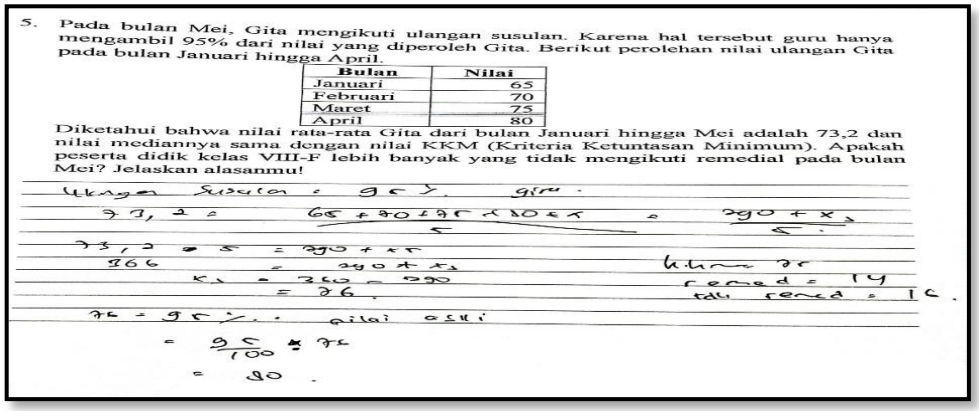


Figure 12. Second Subject's Answer to Problem Item Number 5

Based on Figure 12, it is evident that the second subject was able to determine the appropriate strategy for answering question number 5. The subject successfully applied this strategy, connecting the data from Stimulus 3 with the data in question 5, and effectively linking numerical concepts (percentages) with statistical concepts. During the interview, the second subject expressed confidence in tackling this problem, which was reflected in their explanation of how they approached the question. The subject explained that they predicted Gita's test scores for May using the concept of the average, determined the KKM (Minimum Competency Criteria) using the median, and predicted whether most

students would need remedial lessons using the mode. The subject not only understood these concepts but was also able to apply them effectively.

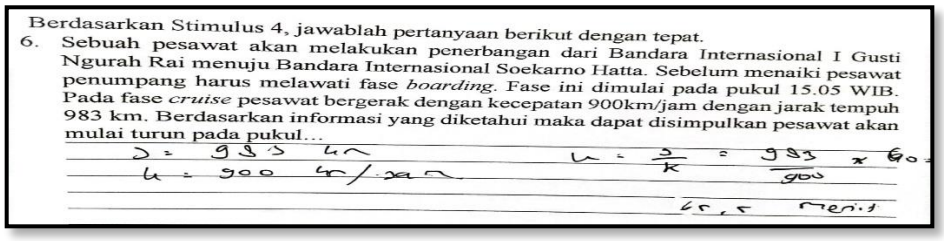
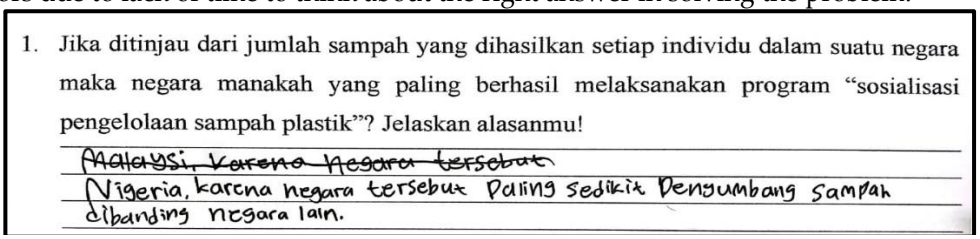

Berdasarkan Stimulus 4, jawablah pertanyaan berikut dengan tepat.
6. Sebuah pesawat akan melakukan penerbangan dari Bandara Internasional I Gusti Ngurah Rai menuju Bandara Internasional Soekarno Hatta. Sebelum menaiki pesawat penumpang harus melawati fase *boarding*. Fase ini dimulai pada pukul 15.05 WIB. Pada fase *cruise* pesawat bergerak dengan kecepatan 900km/jam dengan jarak tempuh 983 km. Berdasarkan informasi yang diketahui maka dapat disimpulkan pesawat akan mulai turun pada pukul...

Figure 13. Second Subject's Answer to Problem Item Number 6

Based on Figure 13, it is known that the second Subject can only find the time in the cruise phase. The results of the interview on the second subject showed that he did not know what phases were important in making conclusions because he considered all phases important even though this was not the case. He considered question number 6 to be in the difficult category but strongly believed that he could determine the time in the cruise phase, which was in accordance with the answers written on the second Subject's answer sheet.

It can be concluded that students who have data literacy skills in the moderate category are able to achieve almost all indicators well. There are several indicators that students cannot achieve, including choosing the right data presentation strategy and making conclusions based on the data presented. As for other factors that result in students not being able to answer questions or achieve these indicators due to lack of time to think about the right answer in solving the problem.


1. Jika ditinjau dari jumlah sampah yang dihasilkan setiap individu dalam suatu negara maka negara manakah yang paling berhasil melaksanakan program "sosialisasi pengelolaan sampah plastik"? Jelaskan alasanmu!

Nigeria, karena negara tersebut paling sedikit pengumbang sampah dibanding negara lain.

Figure 14. Third Subject's Answer on Question Number 1

Based on Figure 14, it is known that the third Subject has not been able to know the meaning of problem number one. Based on the results of the interview with the third Subject, information was obtained that he knew the meaning of the data presented but did not know how to solve the problem from problem number one. This is because the third subject never repeated the lesson at home and had no preparation to answer the questions. Apart from that, the third subject was also not used to working on data literacy questions.

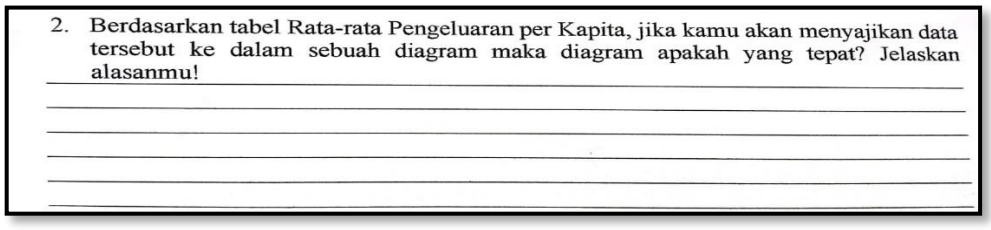
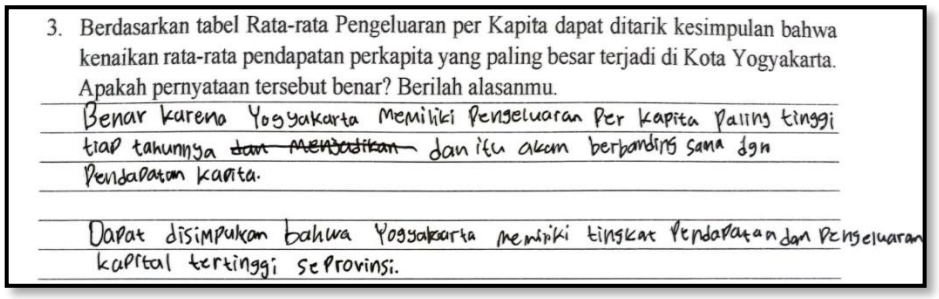

2. Berdasarkan tabel Rata-rata Pengeluaran per Kapita, jika kamu akan menyajikan data tersebut ke dalam sebuah diagram maka diagram apakah yang tepat? Jelaskan alasanmu!

Figure 15. Third Subject's Answer to Question Number 2

Based on Figure 15, it is known that students have not been able to choose the right data presentation and its reasons. Based on the results of the interview with the third Subject, it was found that he took longer to think about what representation was most appropriate to present the data

referred to in the problem. In other words, it took him longer to achieve the right indicator of analyzing or interpreting. It turns out that besides the time of working on the problem, it also affects students in answering questions. Time affects the opportunity for students to think about the right answer to a problem (Septiani & Gunawan, 2017).

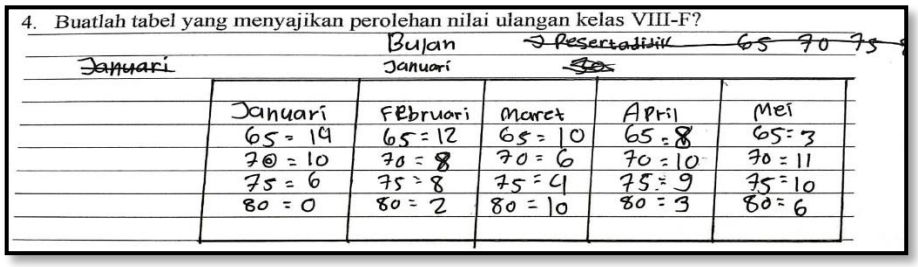

3. Berdasarkan tabel Rata-rata Pengeluaran per Kapita dapat ditarik kesimpulan bahwa kenaikan rata-rata pendapatan perkapita yang paling besar terjadi di Kota Yogyakarta. Apakah pernyataan tersebut benar? Berilah alasanmu.

Benar karena Yogyakarta memiliki pengeluaran per kapita paling tinggi tiap tahunnya ~~dan menjadikan~~ dan itu akan berbanding sama dgn pendapatan kapita.

Dapat disimpulkan bahwa Yogyakarta memiliki tingkat Pendapatan dan Pengeluaran kapital tertinggi seProvinsi.

Figure 16. Third Subject's Answer on Question Number 3

Based on Figure 16, it is known that the third subject has not been able to make arguments and prove his arguments correctly. Additional information obtained from the interview results is that the third subject made the argumentation as in Figure 16 because the nominal average per capita income in each district is the highest in the city of Yogyakarta.


4. Buatlah tabel yang menyajikan perolehan nilai ulangan kelas VIII-F?

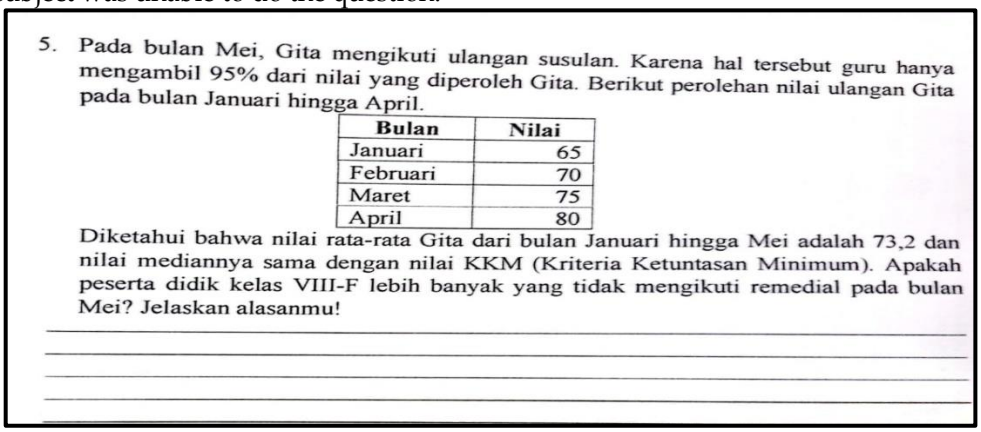
Bulan ~~Peserta didik 65 70 75~~

~~Januari~~ Januari ~~80~~

Januari	Februari	Maret	April	Mei
65 = 14	65 = 12	65 = 10	65 = 8	65 = 3
70 = 10	70 = 8	70 = 6	70 = 10	70 = 11
75 = 6	75 = 8	75 = 4	75 = 9	75 = 10
80 = 0	80 = 2	80 = 10	80 = 3	80 = 6

Figure 17. Third Subject's Answer on Question Number 4

Based on Figure 17, it is known that students have not been able to change the representation of data in the form of diagrams into tabular form appropriately. There is additional information found based on the results of the interview which states that, the third subject pays attention to the month, value and number of students but the presentation of data as in Figure 17 is not entirely correct. This was learned in eighth grade, but due to the third subject's lack of preparation in taking the test, the third subject was unable to do the question.

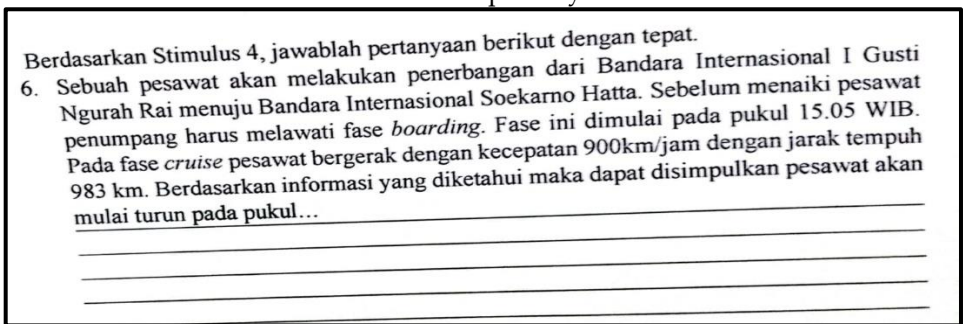

5. Pada bulan Mei, Gita mengikuti ulangan susulan. Karena hal tersebut guru hanya mengambil 95% dari nilai yang diperoleh Gita. Berikut perolehan nilai ulangan Gita pada bulan Januari hingga April.

Bulan	Nilai
Januari	65
Februari	70
Maret	75
April	80

Diketahui bahwa nilai rata-rata Gita dari bulan Januari hingga Mei adalah 73,2 dan nilai mediannya sama dengan nilai KKM (Kriteria Ketuntasan Minimum). Apakah peserta didik kelas VIII-F lebih banyak yang tidak mengikuti remedial pada bulan Mei? Jelaskan alasanmu!

Figure 18. Third Subject's Answer on Question Number 5

Based on Figure 18, it is known that the third subject has not been able to process the data to answer question number 5. Based on the results of the interview, the third subject did not know how to predict the value of Gita in May and what concepts were used to predict the right answer. This is because students do not understand the statistical concepts they have learned.


Berdasarkan Stimulus 4, jawablah pertanyaan berikut dengan tepat.

6. Sebuah pesawat akan melakukan penerbangan dari Bandara Internasional I Gusti Ngurah Rai menuju Bandara Internasional Soekarno Hatta. Sebelum menaiki pesawat penumpang harus melawati fase *boarding*. Fase ini dimulai pada pukul 15.05 WIB. Pada fase *cruise* pesawat bergerak dengan kecepatan 900km/jam dengan jarak tempuh 983 km. Berdasarkan informasi yang diketahui maka dapat disimpulkan pesawat akan mulai turun pada pukul…

Figure 19. Third Subject's Answer on Question Number 6

Based on Figure 19, it is known that the third subject has not been able to make conclusions correctly. Additional information obtained from the interview was that the third subject was unable to select important information used to make conclusions. This is because he considered all the information in the stimulus to be important.

At the end of the interview, the third subject admitted to lacking both preparation and confidence in answering the questions, leading them to respond without much thought or effort. The third subject also mentioned having forgotten the statistics concepts they had previously learned. An analysis of their answers revealed that students with low literacy skills, like the third subject, struggled to meet any of the data literacy indicators effectively. This finding contrasts with the results of Gunawan et al. (2022), who found that students with low data literacy skills were still able to achieve several data literacy indicators to some extent.

From the interviews with various research subjects, it can be concluded that several factors contribute to successful data literacy skills. These include self-confidence, self-efficacy, study habits, memory retention, and the time allocated for working on questions. According to Savitri et al. (2022), the discussion of data literacy-related questions in mathematics can help build students' foundational skills. However, consistent practice is essential for students to become more efficient in solving data literacy problems across various topics. This suggests that fostering these supportive factors and increasing practice opportunities could significantly enhance students' data literacy abilities.

4. CONCLUSION

Based on the average performance of students when answering data literacy questions, it is evident that junior high school students in Yogyakarta City generally fall into the low category in terms of data literacy skills. However, some students demonstrate data literacy skills at high and sufficient levels. Students with high data literacy skills are able to meet all the indicators of data literacy proficiency. Those with sufficient data literacy skills can nearly achieve all indicators but struggle specifically with analyzing strategies to solve data-related problems and drawing conclusions based on data.

A key factor contributing to these challenges is the insufficient time allotted for students to thoroughly think through these problems. To address this, incorporating learning methods that enhance data literacy, along with related factors such as self-confidence, self-efficacy, study habits, and memory retention, could be beneficial. Additionally, teachers can support the development of students' data literacy skills by utilizing real-world data, such as information found in newspapers, to help students practice making data-driven conclusions. Adjusting school schedules to consider students'

commute distances and modes of transportation could also aid in training students to interpret and predict data more effectively.

REFERENCES

Batur, A., & Baki, A. (2022). Examination of the relationship between statistical literacy levels and statistical literacy self-efficacy of high school students. *Egitim ve Bilim*, *47*(209), 171–205. https://doi.org/10.15390/EB.2022.9970

Blake, B & Pope, T. (2008) Developmental Psychology: Incorporating Piaget's and Vygotsky's Theories in Classrooms, *Journal of Cross-Disciplinary Perspectives in Education*, 1(1), 59-67.

BUMN. (2019). *Industri 4.0: Asuransi Harus Siap Bertransformasi*. Jakarta: PT Reasuransi Indonesia Utama.

D'Ignazio, C., & Bhargava, R. (2015). Approaches to Building Big Data Literacy. *Bloomberg Data for Good Exchange Conference*.

D'Ignazio, C., & Bhargava, R. (2016). DataBasic: Design Principles, Tools and Activities for Data Literacy Learners. *The Journal of Community Informatics*, *12*(3), 83–107. https://doi.org/10.15353/joci.v12i3.3280

Fitri, I., Setyaningrum, W., & Pulungan, D. A. (2023). Fenomena Literasi Statistik Pada Pembelajaran Matematika Siswa Sma Di Lhokseumawe Aceh. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *12*(2), 1927. https://doi.org/10.24127/ajpm.v12i2.7000

GLN. (2017). Materi Pendukung Literasi Digital. In *Kementerian Pendidikan dan Kebudayaan*. Jakarta. Retrieved from http://gln.kemdikbud.go.id/glnsite/wp-content/uploads/2017/10/literasi-DIGITAL.pdf

Gunawan, G., Asriani, N. W., Kumala, F. Z., Akhsani, L., & Rohmawati, S. (2022). Karakteristik Kemampuan Literasi Statistika Siswa Dalam Menyelesaikan Masalah Model Pisa. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *11*(3), 2282. https://doi.org/10.24127/ajpm.v11i3.5443

Hafiyusholeh, M. (2015). Literasi Statistik dan Urgensinya Bagi Siswa. *Wahana*, *64*(1), 1–8. Retrieved from https://jurnal.unipasby.ac.id/index.php/whn/article/view/531/390

Herzog, D. (2015). Data Literacy. Retrieved November 25, 2023, from SAGE Publications website: https://books.google.co.id/books?hl=id&lr=&id=rDFyBgAAQBAJ&oi=fnd&pg=PP1&ots=nFF1xO_uGi&sig=W5VBpfuT2_bbSOx_8t4jKadFw7o&redir_esc=y#v=onepage&q&f=false

Kartika, E., Ariswan, Suban, M. E., & Arafah, Z. U. (2021). Students' Data Literacy Ability in Physics Using the Physics E-Module Integrated with the Values of Pancasila During the Covid-19. *Proceedings of the 6th International Seminar on Science Education (ISSE 2020)*, *541*(Isse 2020), 329–335. https://doi.org/10.2991/assehr.k.210326.047

Kemenkes. (2022). UPDATE 30 Januari: Kasus Covid-19 di Indonesia Bertambah 12.422 Artikel ini telah tayang di Kompas.com dengan judul "UPDATE 30 Januari: Kasus Covid-19 di Indonesia Bertambah 12.422", Klik untuk baca: https://nasional.kompas.com/read/2022/01/30/18153891/up. Retrieved from Kompas website: https://nasional.kompas.com/read/2022/01/30/18153891/update-30-januari-kasus-covid-19-di-indonesia-bertambah-12422

Khan, K., & Mason, J. (2021). The M in STEM and Issues of Data Literacy. *29th International Conference on Computers in Education Conference, ICCE 2021 - Proceedings*, *1*, 632–637.

Khuan, L., & Krauss, S. E. (2015). of Clients ' Experiences in Healthcare Research ? *International Journal of Public Health and Clinical Sciences*, *2*(4), 1–6.

Kireina, N. F. (2017). Mesin parkir elektronik sebagai wujud dari smart city di kota bandung. *Jurnal Ilmu Sosial Dan Ilmu Politik*, *7*(2), 63–80.

Larasati, P. E., Supahar, & Yunanta, D. R. A. (2020). Validity and reliability estimation of assessment ability instrument for data literacy on high school physics material. *Journal of Physics: Conference Series*, *1440*(1). https://doi.org/10.1088/1742-6596/1440/1/012020

Moralez, L. G, Hsu, Y. C., Poole, J., Rae. B., & Rutherford. I. (2014).*A World that Counts Mobilishing the Data Revolution for Sustainable Development.* Admir Jahi.

OECD. (2018). Pisa 2021 Mathematics Framework (Draft). In *Angewandte Chemie International Edition, 6(11), 951–952.* OECD Publisher. Retrieved from http://www.oecd.org/pisa/pisaproducts/pisa-2021-mathematics-framework-draft.pdf

Permendikbudristek. (2022). Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Rebuplik Indonesia Nomor 5 tentang Standar Kompetensi Lulusan pada Pendidikan Anak Usia Dini, Jenjang Pendidikan dasar, dan Jenjang Pendidikan Menengah.

PSKPKemendikbudristek. (2020). *Penerimaan Peserta Didik Berdasarkan Zonasi Pendidikan*. Jakarta: Pusat Penelitian Kebijakan, Badan Penelitian dan Pengembangan dan Perbukuan, Kementerian Pendidikan dan Kebudayaan.

Reeves, T. D., & Chiang, J. L. (2019). Effects of an asynchronous online data literacy intervention on pre-service and in-service educators' beliefs, self-efficacy, and practices. *Computers and Education*, *136*, 13–33. https://doi.org/10.1016/j.compedu.2019.03.004

Samosir, P., Rajagukguk, W., & Ratnawati. (2022). *Dasar-Dasar Statistika Inferensi Dalam Penelitian*.Schield, M. (2011). Statistical literacy: A new mission for data producers. *Statistical Journal of the IAOS*, *27*(3–4), 173–183. https://doi.org/10.3233/SJI-2011-0732

Septiani, & Gunawan, D. (2017). Metode Horisontal Untuk Pembelajaran Berhitung Pembagian Pada Siswa Tunarungu. *JASSI_anakku*, *18*(1976), 57–62. Retrieved from http://uir.ac.id/?p=2499].