

Statistical Test Proficiency and Data Analysis Literacy through Statistical Package for the Social Sciences (SPSS) Workshop

Jay A. Delideli

Passi City College, Passi City, Iloilo, Philippines

Author Email: jaydelideli@gmail.com

Date received: August 26, 2024

Date revised: September 9, 2024

Date accepted: September 14, 2024

Originality: 94%

Grammarly Score: 99%

Similarity: 6%

Recommended citation:

Delideli, J. (2024). Statistical test proficiency and data analysis literacy through Statistical Package for the Social Sciences (SPSS) workshop. *Journal of Interdisciplinary Perspectives*, 2(10), 254-264.

<https://doi.org/10.69569/jip.2024.0438>

Abstract. This quasi-experimental study addressed the need for practical training in statistical analysis by examining the impact of SPSS workshops on statistical test proficiency and data analysis literacy among seventy Bachelor of Secondary Education in Mathematics (BSED-Math) students. The research evaluated whether SPSS workshops could significantly enhance these critical skills. Using a pre-test and post-test one-group design, the study measured statistical proficiency with a performance test and data literacy with a researcher-developed questionnaire. The results showed significant improvements, with students advancing from limited proficiency to proficiency in statistical testing and from introductory understanding to analytical competence in data literacy. Additionally, a strong positive relationship was found between statistical test proficiency and data analysis literacy, meaning that as students improved in statistical test proficiency, their data analysis literacy also improved. These findings suggest that SPSS workshops are an effective tool for enhancing students' analytical capabilities. Future research should explore how well students retain these skills over time and assess the effectiveness of similar interventions across different academic disciplines. Additionally, investigating how SPSS training influences students' ability to apply statistical methods in real-world research could provide valuable insights into its broader applicability.

Keywords: Statistical test proficiency; Data analysis literacy; Statistical Package for Social Sciences; Workshop; Educational training.

1.0 Introduction

Proficiency in using statistical tools and skills in data interpretation is indispensable for conducting rigorous research, formulating evidence-based policies, and making informed decisions. Despite the critical role of statistical literacy, a significant gap persists in the proficiency levels of statistical test implementation and data analysis among students and professionals, particularly those outside statistics and data science. Research highlights that even with targeted pedagogical interventions, such as those implemented by Gonda et al. (2022), improvements in statistical skills can be achieved, yet challenges remain. Their study emphasizes the importance of developing statistical literacy at various educational levels. It highlights that integrating real-world research into statistics courses can significantly enhance students' abilities to critically evaluate and apply statistical information. This underscores the need for a dual approach that focuses on improving statistical skills and addresses and reshapes students' perceptions of the field to foster long-term engagement and proficiency.

Students often encounter difficulties in understanding and applying statistical concepts in academic settings, which can impede their research capabilities and academic performance. This challenge is further compounded by the limited integration of technology-based tools in the classroom, which could otherwise enhance students' proficiency in statistical analysis. A study by Ramos et al. (2023) demonstrated that technology-based supplementary materials can significantly improve students' performance in statistics compared to traditional manual computation methods, playing a crucial role in overcoming the difficulties students face in mastering statistical concepts. Additionally, research on the effectiveness of structured lessons in teaching high school students to interpret and evaluate statistical reports has shown that a systematic approach grounded in critical thinking can significantly enhance students' statistical literacy skills (Koga, 2024).

Similarly, in professional environments, individuals may struggle with data analysis tasks, leading to suboptimal decision-making and inefficiencies. This issue often reflects broader confidence gaps in statistical and mathematical competencies. Studies demonstrate that teachers exhibit significantly lower confidence in teaching statistics compared to other mathematical concepts, with deficiencies particularly pronounced among those lacking additional professional certification and experience, a situation that mirrors the challenges faced in professional environments where gaps in statistical knowledge can hinder decision-making and operational efficiency, thus emphasizing the critical need for targeted professional development to boost proficiency and confidence in statistical tasks, which is essential for enhancing both educational outcomes and professional decision-making effectiveness (Umugiraneza et al., 2022). Henderson and Corry (2021) further highlight the critical need for comprehensive data literacy training for educational professionals, emphasizing that effective data use is essential for informed decision-making. This proficiency gap underscores the need for effective educational interventions to enhance statistical literacy and equip individuals with the necessary skills to navigate complex data landscapes.

Teaching analytics has shown promise in enhancing teaching practices and learning outcomes through effective data use. However, there are challenges related to data collection and the usability of analytical tools, which underscore the need for improved data literacy among educators (Ndukwe & Daniel, 2020). In parallel, there is a recognized gap in empirical research and consensus regarding data literacy competencies and effective teaching methods. This highlights the need for more research to define effective teaching methods and competencies at the university level (Ghodoosi et al., 2023). Integrating these perspectives could help address the practical challenges of implementing data-driven teaching practices and the need for well-defined, evidence-based data literacy competencies.

SPSS (Statistical Package for the Social Sciences) is a widely utilized software for data management and statistical analysis, particularly within the social sciences. Its user-friendly interface and robust analytical capabilities make it a valuable tool for researchers. However, the potential of SPSS is often underutilized due to insufficient training and a lack of comprehensive understanding of its functionalities. Users generally prefer using SPSS software for data analysis and agree that it is essential for research to produce accurate results and enhance the effectiveness and authenticity of research work. Despite this, some respondents also express the need for expert guidance to understand better SPSS results (Jain & Sengar, 2024). While SPSS is considered user-friendly, its effectiveness can be compromised without proper training. Studies suggest that those with adequate training navigate the software more effectively, indicating that enhanced training programs could significantly improve user proficiency and the quality of their analyses (Huang et al., 2024). This proficiency gap hampers the effective application of statistical methods and the accurate interpretation of data, thus impacting the quality of research and decision-making processes. To address this, SPSS workshops offer practical experience with essential statistical tests and data analysis techniques through exercises that simulate real-world scenarios. This approach aims to enhance technical skills data management, and interpretation confidence. Nwagwu (2024) highlights data literacy's significant growth and interdisciplinary nature, underscoring the need for clear educational frameworks and library support to advance the field.

This study addresses the proficiency gap in statistical test proficiency and data analysis literacy by systematically evaluating the impact of SPSS workshops. Using the pretest-posttest design, the study assesses how these workshops affect participants' skills, focusing on the hands-on application of statistical concepts in a user-friendly

software environment. The workshops are designed to integrate real-world data scenarios, reinforcing theoretical knowledge through experiential learning and aiming to build technical proficiency and confidence in data management and interpretation.

2.0 Methodology

2.1 Research Design

Experimental research is a method specifically designed to test hypotheses concerning cause-and-effect relationships. It is considered one of the most rigorous methods for identifying causal links between variables, providing a valid approach to addressing educational problems in practice and theory (Creswell, 2015). This type of research is essential for determining whether an idea, practice, or procedure significantly impacts an outcome, known as the dependent variable. When the goal is to establish the possible causes and effects between independent variables and the dependent variable, experimental research becomes an indispensable tool. In this study, a one-group pretest-posttest design was employed to investigate the effectiveness of the intervention on participants' proficiency levels. This design involves selecting a single group of participants assessed before and after the intervention. Initially, the participants underwent a pretest to establish a baseline measurement of the dependent variable, reflecting their proficiency levels before introducing any intervention. Subsequently, the intervention aimed at enhancing their proficiency was administered. After the intervention, a posttest was conducted to evaluate any changes in the dependent variable. The difference between the pretest and posttest scores provided insights into the impact of the intervention. While this design does not include a control group, which may limit the ability to definitively attribute changes to the intervention, it remains valuable for preliminary investigations. Focusing on changes within the same group of participants, the one-group pretest-posttest design offers a practical way of assessing educational interventions' potential effectiveness in real-world settings.

2.2 Research Participants

The study subjects were 70 BSED - Math students. These students were administered a pretest to determine their baseline proficiency levels before the intervention. After the intervention, the same group was given a posttest to evaluate any changes in their proficiency. A purposive sampling technique was employed to select the participants. Although purposive sampling does not offer the statistical representativeness of random sampling, it is useful for targeting specific populations to understand the effects of interventions within that group (Creswell, 2015). This approach allows for a more focused investigation into the impacts of the SPSS workshops on students directly engaged in the relevant field.

2.3 Research Instrument

This study employed researcher-developed instruments as the main tools for assessing statistical test proficiency and data analysis literacy. Adhering to ethical guidelines, participation was voluntary, and the validation of these instruments followed a structured approach. Initially, educational and assessment experts scrutinized and refined the test items to align them with the study's objectives. A pilot test was carried out with a small group of participants following expert reviews. This step aimed to identify and address any ambiguities in the questions and assess the instruments' overall clarity and effectiveness.

The statistical test proficiency instrument began with 30 questions, but item analysis during the pilot phase reduced it to 25. Items that were ambiguous, redundant, or did not effectively differentiate between high and low performers were removed. Content validity was ensured through expert reviews, while construct validity was verified using factor analysis to confirm that the items accurately measured various aspects of statistical proficiency.

The data analysis literacy instrument, a 30-item Likert-scale questionnaire, was developed to cover essential aspects such as data understanding, method selection, result interpretation, and application. This instrument also underwent a pilot test, with feedback used to enhance item clarity and relevance. The high Cronbach's alpha value of 0.918 demonstrated excellent internal consistency, reflecting that the items reliably measured the same construct.

2.4 Data Gathering Procedure

This study followed a systematic procedure to gather data. First, participants were selected from BSED - Math college students enrolled at Passi City College, ensuring that the sample was representative of the target population. The selection process aimed to capture a broad range of proficiency levels in statistical tests and data analysis literacy, which was crucial for evaluating the effectiveness of the intervention. Second, participants were randomly assigned to ensure a representative sample and to minimize selection bias, thereby enhancing the rigor of the experimental design. This random assignment helped ensure that any differences observed in the outcomes could be attributed to the intervention rather than pre-existing differences between participants.

Third, participants completed a pretest, which included an objective test to assess their proficiency in statistical tests and a Likert-scale questionnaire to measure their data analysis literacy. This baseline data provided a clear picture of their skills and knowledge before the intervention, serving as a reference point for subsequent comparisons. Next, the 5-day intervention consisted of a series of SPSS workshops to enhance participants' statistical and data analysis skills. These workshops combined theoretical instruction with practical experience using SPSS software, ensuring participants could effectively apply what they learned to real-world data analysis tasks. The workshops were carefully structured to build upon participants' existing knowledge and progressively develop their competencies.

Following the workshops, participants were administered the posttest, which mirrored the pretest in content and format. This approach allowed for a direct comparison of their performance and knowledge before and after the intervention, providing a clear measure of the workshops' impact. Finally, the pretest and posttest data were analyzed using statistical methods such as paired t-tests, analysis of variance (ANOVA), and Pearson's r. These analyses were conducted to determine if the intervention significantly improved participants' proficiency and data analysis literacy and to examine the relationship between statistical test proficiency and data analysis literacy skills. With its pretest-posttest model, this experimental design enabled a robust evaluation of the workshops' effectiveness.

Strict controls were maintained throughout the data collection to ensure consistency and reliability. Ethical guidelines regarding participant consent, confidentiality, and fair treatment were rigorously followed. Participants were assured of the confidentiality of their responses and informed that the study's findings would be used solely for academic and educational purposes.

2.5 Ethical Considerations

This research study was conducted in strict accordance with ethical standards to protect participants' rights and ensure the credibility of the research process. Participation in the study was entirely voluntary, with participants being free to withdraw at any stage without any repercussions. Comprehensive measures were taken to minimize potential risks, including physical, psychological, and social harm, thereby ensuring the safety and well-being of the BSED-Math students involved in the study.

Confidentiality was rigorously maintained throughout the research process. All data collected were treated with the highest level of confidentiality and used exclusively for this study. The participants' identities were safeguarded by anonymizing all responses during data analysis and reporting, ensuring no personal information could be traced back to individual participants.

The study also upheld the dignity and rights of all participants by adhering to ethical principles of fairness, respect, and sensitivity. Informed consent was obtained from all participants before their involvement in the study. The study's nature, objectives, and procedures were clearly explained to the participants, and any questions or concerns were addressed thoroughly. This approach ensured that participants were fully informed and voluntarily consented to participate in the research.

3.0 Results and Discussion

3.1 Statistical Test Proficiency Level of Students

Statistical test proficiency refers to students' ability to correctly select, apply, and execute appropriate statistical tests for different types of data and research questions. Table 1 presents the participants' pre-test and post-test

mean scores, providing both overall results and results categorized by year level. It highlights the proficiency levels before and after the intervention, specifically focusing on the impact of the SPSS workshops.

Table 1. Statistical test proficiency levels in the pre-test and post-test

Group	Pre-test				Post-test		
	N	Mean	SD	Description	Mean	SD	Description
As a Whole	70	5.76	2.28	Limited	15.10	2.07	Proficient
First Year	23	5.57	2.52	Limited	14.96	2.33	Proficient
Second Year	17	6.18	2.00	Limited	15.59	1.77	Proficient
Third Year	14	5.64	1.34	Limited	14.79	1.37	Basic
Fourth Year	16	5.69	2.87	Limited	15.06	2.54	Proficient

Note: Advanced Proficiency (20.00 - 25.00 points); Proficient (15.00 - 19.99 points); Basic Proficiency (10.00 - 14.99 points); Limited Proficiency (5.00 - 9.99 points); Minimal Proficiency (0 - 4.99 points)

The pre-test outcomes revealed that participants across all year levels generally demonstrated "Limited Proficiency" in statistical testing. Specifically, the mean scores were 5.57 for first-year students, 6.18 for second-year students, 5.64 for third-year students, and 5.69 for fourth-year students. The overall mean score across all participants was 5.76, confirming that, on average, students were within the "Limited Proficiency" range before the intervention.

Following the intervention, which involved a series of SPSS workshops to enhance statistical test proficiency, a marked improvement was observed across all year levels. Post-test mean scores demonstrated significant gains, with first-year students reaching 14.96, second-year students 15.59, third-year students 14.79, and fourth-year students 15.06. The overall post-test mean score for all participants increased to 15.10. This score falls within the "Proficient" category, indicating that, as a whole, the participants had developed a more advanced level of proficiency in statistical testing following the intervention.

These findings suggest that the SPSS workshops were effective in significantly improving students' statistical test proficiency. The transition from "Limited Proficiency" in the pre-test to "Proficient" in the post-test underscores the intervention's efficacy in enhancing the participants' understanding and application of statistical tests. The results also highlight the importance of structured, targeted educational interventions in developing essential analytical skills among students, particularly in statistical data analysis.

The significant increase in post-test scores across all year levels further supports the conclusion that the intervention had a broad and positive impact, improving the student's ability to perform statistical tests accurately and effectively. This outcome is particularly noteworthy, given that the initial proficiency levels were relatively low across all groups. The improvement to a "Proficient" level suggests that the SPSS workshops not only met but exceeded the expectations regarding their impact on students' statistical literacy and analytical capabilities.

The findings of this study align with Castillo's (2024) study, which also emphasizes the positive impact of evidence-based strategies on statistical literacy. Castillo's research demonstrated that implementing Share and Model Concepts and Nurturing Metacognition strategies significantly improved students' understanding of statistical concepts, ability to visualize and communicate statistical information, and overall statistical literacy. Castillo's study highlights that evidence-based pedagogical strategies can substantially elevate students' proficiency in various aspects of statistical analysis. This alignment reinforces the value of employing targeted instructional approaches to foster significant gains in statistical capabilities. This study collectively underscores the importance of structured interventions in developing and improving students' analytical skills and understanding of statistical concepts.

3.2 Data Analysis Literacy Level of Students

Data analysis literacy skills refer to students' ability to accurately interpret, critically evaluate, and effectively communicate the results of statistical analyses. Table 2 presents the participants' pre-test and post-test mean scores. It highlights the literacy skill levels before and after the intervention, specifically examining the impact of the SPSS workshops on the student's ability to analyze data.

Table 2. Data analysis literacy levels in the pre-test and post-test

Group	Pre-test				Post-test		
	N	Mean	SD	Description	Mean	SD	Description
As a Whole	70	2.80	.53	Introductory Understanding	4.13	.49	Analytical Competence
First Year	23	2.78	.67	Introductory Understanding	4.13	.49	Analytical Competence
Second Year	17	2.69	.43	Introductory Understanding	3.98	.49	Applied Knowledge
Third Year	14	2.88	.36	Introductory Understanding	4.19	.40	Analytical Competence
Fourth Year	16	2.90	.57	Introductory Understanding	4.25	.54	

Note: *Advanced Application* (4.50 - 5.00); *Analytical Competence* (4.00 - 4.49); *Applied Knowledge* (3.00 - 3.99); *Introductory Understanding* (2.00 - 2.99); *Basic Awareness* (1.00 - 1.99)

Table 2 presents participants' data analysis literacy levels, assessed through pre-test and post-test evaluations and categorized by year level. The pre-test results indicated that participants across all year levels predominantly demonstrated an "Introductory Understanding" of data analysis concepts. Specifically, the mean scores were 2.78 for first-year students, 2.69 for second-year students, 2.88 for third-year students, and 2.90 for fourth-year students. The overall pre-test mean score was 2.80, confirming that, on average, the participants' data analysis literacy was within the "Introductory Understanding" range before the intervention.

Post-test assessments revealed a significant improvement in literacy levels across all groups. The mean scores increased to 4.13 for first-year students, 3.98 for second-year students, 4.19 for third-year students, and 4.25 for fourth-year students. Collectively, the overall post-test mean score rose to 4.13, reflecting a transition to the "Analytical Competence" level. These results suggest that the intervention effectively enhanced the participants' data analysis literacy. The shift from "Introductory Understanding" in the pre-test to "Analytical Competence" in the post-test demonstrates the substantial impact of the intervention in improving students' analytical skills, equipping them with the knowledge necessary for more advanced data analysis tasks.

The findings of this study are supported by the research conducted by Sheriff and Sevukan (2023), which underscores the critical importance of developing robust data literacy skills. Sheriff and Sevukan's study highlights that advancements in data literacy are essential for effective data management and analysis, and they advocate for a focused approach to improving data literacy through targeted educational strategies. This support reinforces the significance of structured interventions in enhancing data literacy. This study emphasizes the positive outcomes of targeted educational programs on students' abilities to analyze data, suggesting that such interventions can significantly advance students' data literacy levels from foundational understanding to more sophisticated analytical competence. This collective evidence underscores the effectiveness of well-designed educational interventions in improving key data analysis skills and preparing students for more complex analytical tasks.

3.3 Difference in the Statistical Test Proficiency Level

Table 3 presents the differences in pre-test and post-test scores in statistical test proficiency across different year levels. The analysis focuses on comparing scores within each year level, highlighting the changes in proficiency before and after the intervention.

Table 3. Difference in the statistical proficiency test level across year-level					
Group	Mean	Mean Difference	df	t-value	sig-value
First Year					
Pre-test	5.57	9.39	22	57.54	0.000*
Post-test	14.96				
Second Year					
Pre-test	6.18	9.41	16	48.80	.000*
Post-test	15.59				
Third Year					
Pre-test	5.64	9.14	13	64.00	0.000*
Post-test	14.79				
Fourth Year					
Pre-test	5.69	9.38	15	42.37	.000*
Post-test	15.06				

*significant at $p < 0.001$

Table 3 presents the differences in pre-test and post-test scores across various year levels. The results indicate a significant difference in all year levels' pre-test and post-test scores. Specifically, the first-year students showed a significant improvement in their scores, with a p-value of 0.000 ($p < 0.001$), indicating a notable increase in statistical test proficiency after the intervention. Similarly, the second-year students also exhibited a significant difference between their pre-test and post-test scores, with a p-value of 0.000 ($p < 0.001$). The third-year and fourth-year students followed the same trend, demonstrating significant improvements with p-values of 0.000 ($p < 0.001$) for each group.

These results suggest that across all year levels, the intervention had a profound impact on enhancing students' proficiency in statistical testing. The consistent and significant p-values across all groups indicate that the improvements observed are statistically meaningful, reinforcing the effectiveness of the intervention in improving students' statistical competencies. This analysis highlights the value of educational intervention in boosting the statistical test proficiency of students at different academic stages, confirming its role in advancing their analytical capabilities.

The findings of this study align with the research conducted by Umar and Yakubu (2023), emphasizing the positive impact of evidence-based strategies on statistical literacy among students in Northwestern Nigerian Polytechnics. Umar and Yakubu demonstrated that implementing targeted pedagogical approaches significantly improved students' understanding of statistical concepts, ability to visualize and communicate statistical information, and overall statistical literacy. This alignment reinforces the value of employing structured instructional approaches to foster significant gains in students' statistical capabilities. This study underscores the importance of well-designed interventions in developing and enhancing students' analytical skills and understanding of statistical concepts.

3.4 Difference in the Data Analysis Literacy Level

Table 4 presents the differences in pre-test and post-test scores in data analysis literacy across different year levels. The analysis focuses on comparing scores within each year level, highlighting the changes in literacy before and after the intervention.

Table 4. Difference in the data analysis literacy level across year-level

Group	Mean	Mean Difference	df	t-value	sig-value
First Year					
Pre-test	2.78				
Post-test	4.13	1.35	22	23.88	0.000*
Second Year					
Pre-test	2.68				
Post-test	3.98	1.29	16	27.48	0.000*
Third Year					
Pre-test	2.88				
Post-test	4.19	1.32	13	54.18	0.000*
Fourth Year					
Pre-test	2.90				
Post-test	4.25	1.35	15	102.61	0.000*

*significant at $p < 0.001$

Table 4 presents the differences in pre-test and post-test scores regarding data analysis literacy levels across various year levels. The results indicate a significant improvement in data analysis literacy for all year levels following the intervention. Specifically, first-year students demonstrated a significant difference in their pre-test and post-test scores, with a p-value of 0.000 ($p < 0.001$), reflecting a considerable enhancement in their literacy skills. Similarly, second-year students showed a significant improvement, as indicated by a p-value of 0.000 ($p < 0.001$). The third-year and fourth-year students also experienced similar significant advancements, with p-values of 0.000 ($p < 0.001$) for both groups.

These findings suggest that the intervention effectively elevated students' data analysis literacy levels across all year levels. The consistently significant p-values across each group imply that the improvements are statistically robust, affirming the intervention's positive impact. This analysis underscores the educational program's

importance in fostering advanced data analysis skills among students at different academic stages, highlighting its effectiveness in cultivating their analytical competencies. The findings of this study can be supported by the research conducted by Filderman et al. (2022), which provides a meta-analysis of data literacy training effects. The study found significant positive outcomes in data literacy, particularly when collaborative formats were used, leading to enhanced knowledge and skills. These results parallel the improvements in students' data analysis literacy levels in this study, suggesting that structured and collaborative interventions effectively promote data literacy across different educational stages.

The research by Filderman et al. also highlights the importance of incorporating multiple data literacy skills and collaborative elements in educational interventions. These features likely contributed to the robust improvements in data analysis literacy observed in students across all year levels. The alignment of these findings underscores the broader applicability of data literacy interventions, emphasizing the significance of well-designed programs in enhancing data literacy and analytical competencies among learners.

3.5 Difference in the Statistical Test Proficiency Level Mean Gain Scores as to Year Level

Table 5 presents the mean gain scores in statistical test proficiency across different year levels. The analysis examines whether the mean gain scores of first-year, second-year, third-year, and fourth-year students differ significantly. This comparison aims to determine if there are statistically significant variations in the improvements in statistical test proficiency among students at different academic stages.

Table 5. Difference in the statistical test proficiency mean gain scores as to year level

Source of Variation	SS	df	MS	F	p-value	Decision	Interpretation
Between Groups	0.711	3	0.237	0.401	0.753	Accept	Not Significant
Within Groups	39.71	66	0.592				

Table 5 presents the analysis of variance (ANOVA) results for the mean gain scores in statistical test proficiency across different year levels. The ANOVA results indicate no significant difference in the mean gain scores among the year levels, suggesting that the intervention had a relatively uniform impact across all academic stages. This lack of variation in mean gain scores implies that first-year and fourth-year students benefited equally from the intervention, with no particular year level showing a statistically significant advantage over the others.

These findings suggest that the educational program designed to enhance statistical test proficiency was equally effective across the different year levels, leading to consistent improvements in proficiency regardless of the student's academic progression. This uniformity in results across the year levels underscores the versatility and broad applicability of the intervention in improving statistical test proficiency among students at varying stages of their academic careers. The findings of this study align with the research conducted by Batur and Baki (2022), emphasizing the positive impact of self-efficacy on statistical literacy among high school students. Batur and Baki demonstrated that the development of statistical literacy self-efficacy significantly predicted students' ability to grasp and apply statistical concepts, leading to improved overall statistical literacy.

This alignment is further reinforced by the work of Callingham and Watson (2022), who explored the development of statistical literacy in students over time. Their research confirmed that statistical literacy develops steadily across schooling stages, with a stable hierarchy of skills. However, they also noted that progress tends to plateau in later years, underscoring the importance of continuous and consistent educational interventions. These findings suggest that interventions like the one studied can have lasting impacts across different educational stages, supporting the idea that students can consistently develop their statistical proficiency when given structured support. Both studies highlight the importance of fostering students' confidence in their statistical abilities and providing continuous support through well-structured educational programs. This study underscores the value of educational interventions in developing and advancing students' understanding of statistical concepts across various stages of their academic journey.

3.6 Difference in the Data Analysis Literacy Level Mean Gain Scores as to Year Level

Table 6 presents the mean gain scores in data analysis literacy across different year levels. The analysis investigates whether the mean gain scores of first-year, second-year, third-year, and fourth-year students differ significantly.

This comparison aims to assess if there are statistically significant variations in the improvements in data analysis literacy among students at various academic stages.

Table 6. Difference in the data analysis literacy mean gain scores as to year level

Source of Variation	SS	df	MS	F	p-value	Decision	Interpretation
Between Groups	0.039	3	0.013	0.368	0.776	Accept	Not Significant
Within Groups	2.357	66	0.036				

Table 6 presents the analysis of variance (ANOVA) results for the mean gain scores in data analysis literacy across different year levels. The results reveal no significant difference in the mean gain scores among the year levels. This indicates that the intervention's improvement in data analysis literacy was consistent across first-year to fourth-year students. The absence of significant differences suggests that the intervention was equally effective in enhancing data analysis literacy for students at all year levels. This uniformity in the mean gain scores implies that the educational program provided similar benefits to students, irrespective of their academic standing. The findings highlight the intervention's broad effectiveness in advancing data analysis skills across different stages of academic progression.

This result aligns with the research study of Knobel, Kalman, and Lankshear (2020), emphasizing the importance of well-structured qualitative research approaches in understanding complex literacy practices. Their work underscores the significance of consistency in educational interventions, highlighting that carefully designed programs can yield uniform outcomes across diverse groups. This consistency ensures that educational initiatives effectively enhance literacy skills across various stages of academic progression. The research further illustrates how such interventions can provide significant and widespread benefits, reinforcing the value of strategic educational programs in fostering literacy development.

Similarly, Lankshear and Knobel (2018) discuss the importance of ensuring quality in research processes, particularly in data collection and analysis. Their study found that structured approaches in educational programs can effectively guide participants in achieving consistent and reliable outcomes. This further supports the notion that well-designed interventions can lead to uniform improvements in literacy skills across various academic levels, as evidenced in the present study.

3.7 Relationship Between Statistical Test Proficiency Level and Data Analysis Literacy Level

Table 7 presents the correlation between statistical test proficiency and data analysis literacy levels. This analysis reveals the strength and direction of the relationship between participants' abilities in conducting statistical tests and their proficiency in analyzing data. Examining these correlations provides insights into whether improvements in statistical test proficiency are associated with enhanced data analysis skills. This information is valuable for evaluating the effectiveness of educational interventions to improve statistical and data analysis competencies, which are critical for informed research and decision-making.

Table 7. Correlation analysis between statistical test proficiency level and data analysis literacy level

	Data Analysis Literacy	
	r - value	p-value
Statistical Test Proficiency	0.855	0.000*

significant at $p < 0.001$

Table 7 presents the correlation between statistical test proficiency and data analysis literacy levels. The analysis reveals a strong positive correlation with a correlation coefficient of 0.855 and a p-value of 0.000. The high correlation coefficient of 0.855 indicates a strong positive relationship between statistical test proficiency and data analysis literacy. This suggests that students with higher proficiency in statistical tests tend to demonstrate higher levels of data analysis literacy. The statistically significant p-value of 0.000 confirms that this correlation is robust and unlikely due to random chance.

These results illustrate a significant and positive association between the two variables, indicating that enhancements in one area are closely related to improvements in the other. This strong correlation highlights the integral role of statistical test proficiency and data analysis literacy in students' analytical capabilities. The findings of this study align with the research conducted by Dani and Al Quraan (2023), which highlights the importance

of students' attitudes toward statistics in shaping their research approach. Dani and Al Quraan found that students with a more positive understanding of statistical concepts are likelier to engage with quantitative research methods, thus improving their overall proficiency in data analysis.

This alignment is further reinforced by the work of Schreiter et al. (2023), who emphasize the critical role that statistical and data literacy play in STEM education. Their systematic review underscores the necessity of equipping educators with the competencies to effectively teach these skills, profoundly impacting students' learning outcomes. The review highlights the positive effects of well-designed pedagogical approaches on teachers' and students' understanding of statistical concepts, supporting the idea that a strong foundation in statistical literacy is essential for developing students' analytical skills. Both studies underscore the importance of fostering statistical and data literacy through targeted educational strategies. The strong correlation observed in the current study between statistical test proficiency and data analysis literacy further validates the need for integrated approaches that enhance these competencies simultaneously, thereby equipping students with the skills necessary to excel in data-driven environments.

4.0 Conclusion

This study assessed the effectiveness of SPSS workshops in enhancing statistical test proficiency and data analysis literacy among college students. After the workshops, the results demonstrated significant improvements in both areas, with post-test data revealing substantial mean gain scores across all year levels. This suggests that the SPSS workshops improved students' abilities in these domains. Additionally, the strong positive correlation between statistical test proficiency and data analysis literacy underscores the success of the intervention, indicating that advancements in one skill are closely associated with improvements in the other.

The practical implications of these findings emphasize the need to integrate SPSS workshops into academic programs. Such integration would give students essential tools for effective data handling and analysis, which is vital for academic and professional success. Theoretically, the study contributes to understanding the relationship between statistical test proficiency and data analysis literacy. The observed correlation suggests that enhancing statistical skills through SPSS workshops can improve data analysis capabilities, reinforcing the theoretical connection between these competencies.

However, the study has limitations, including a relatively small sample size and a focus on a single educational setting, which may affect the generalizability of the results. Moreover, the short duration of the SPSS workshop may not fully capture the long-term effects on students' skills and performance. Future research should aim to replicate these findings with larger and more diverse samples to validate and extend the results. Longitudinal studies could provide insights into the sustained impact of SPSS workshops on students' skills, and exploring the implementation of similar workshops in various academic contexts could offer further understanding of their effectiveness in enhancing overall academic performance.

5.0 Contributions of Authors

I confirm that I am the sole author of this work and have contributed to all sections. I have reviewed and approved the final manuscript.

6.0 Funding

This work received no specific grant from any funding agency.

7.0 Conflict of Interests

The author declares no conflicts of interest about the publication of this paper.

8.0 Acknowledgment

I want to extend my sincere gratitude to the following individuals and groups for their invaluable contributions to the completion of this study. First and foremost, I thank God for providing me with strength, guidance, and perseverance throughout this research journey. To my family, especially my parents, Cecile A. Delideli and Joel B. Delideli Jr., your unwavering support, encouragement, and understanding have been invaluable. Your belief in my abilities has been a constant source of motivation. To my dear friends, especially Lee Ann Labares and Jiselle Kristine Bicodo, your companionship, encouragement, and understanding have been essential during this journey. Thank you for always being there, offering advice, and cheering me on. I am deeply grateful to the BSED Math students who participated in this study. Your willingness to engage with the research and contribute to its success is greatly appreciated. Special thanks to Dr. Mario P. Mogote, Dean of the School of Teacher Education, for permitting the workshop. I also sincerely thank Dr. Janice Jennifer P. Palmares, College President, for her support of the workshop. Her endorsement and backing were crucial to successfully executing the research, and her role in supporting academic initiatives is highly valued.

9.0 References

- Batur, A., & Baki, A. (2022). Examination of the relationship between statistical literacy levels and statistical literacy self-efficacy of high school students. *Education and Science*, 47(209), 171-205. <https://doi.org/10.15390/EB.2022.9970>
- Callingham, R., & Watson, J. M. (2022). The development of statistical literacy at school. *Statistics Education Research Journal*, 16(1), 181-201. <https://doi.org/10.52041/serj.v16i1.223>
- Castillo, I. M. (2024). Improving statistical literacy through evidence-based strategies among first-year education students in a state university. *Journal of Contemporary Educational Research*, 8(1), 246-259. <https://doi.org/10.26689/jcer.v8i1.6050>
- Creswell, W. J. (2015). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (2nd ed.). Prentice-Hall.
- Creswell, J. W. (2015). *A Concise Introduction to Mixed Methods Research*. Sage Publications.
- Dani, A., & Al Quraan, E. (2023). Investigating research students' perceptions about statistics and its impact on their choice of research approach. *Heliyon*, 9(10), e20423. <https://doi.org/10.1016/j.heliyon.2023.e20423>
- Filderman, M. J., Toste, J. R., Didion, L., & Peng, P. (2022). Data literacy training for K-12 teachers: A meta-analysis of the effects on teacher outcomes. *Remedial and Special Education*, 43(5), 328-343. <https://doi.org/10.1177/07419325211054208>
- Ghodoosi, B., West, T., Li, Q., Torrisi-Steele, G., & Dey, S. (2023). A systematic literature review of data literacy education. *Journal of Business & Finance Librarianship*, 28(2), 112-127. <https://doi.org/10.1080/08963568.2023.2171552>
- Gonda, D., Pavlovičová, G., Duriš, V., & Tírpáková, A. (2022). Implementation of pedagogical research into statistical courses to develop students' statistical literacy. *Mathematics*, 10(11), 1793. <https://doi.org/10.3390/math10111793>
- Henderson, J., & Corry, M. (2021). Data literacy training and use for educational professionals. *Journal of Research in Innovative Teaching & Learning*, 14(2), 232-244. <https://doi.org/10.1108/JRIT-11-2019-0074>
- Huang, Y., Wu, R., He, J., & Xiang, Y. (2024). Evaluating ChatGPT-4.0's data analytic proficiency in epidemiological studies: A comparative analysis with SAS, SPSS, and R. *Journal of Global Health*, 14, 04070. <https://doi.org/10.7189/jogh.14.04070>
- Jain, P., Sengar, S. (2024). Unraveling The Role Of IBM SPSS: A Comprehensive Examination Of Usage Patterns, Perceived Benefits, And Challenges In Research Practice. *Educational Administration: Theory and Practice*, 30(5), 9523-9530. <https://doi.org/10.53555/kuey.v30i5.4609>
- Knobel, M., Kalman, J., & Lankshear, C. (2020). Data analysis, interpretation, and theory in literacy studies research. Myers Education Press
- Koga, S. (2024). Lessons to Demonstrate Statistical Literacy Skills: A Case Study of Japanese High School Students on Reading Statistical Reports. *Journal of Statistics and Data Science Education*, 0(0), 1-13. <https://doi.org/10.1080/26939169.2024.2334903>
- Lankshear, C., Knobel, M. (2018). Education and "new literacies" in the middle years. *Literacy Learning: The Middle Years*, 26 (2), 7-16. <https://rb.gy/vh7b1p>
- Ndukwe, I. G., & Daniel, B. K. (2020). Teaching analytics, value and tools for teacher data literacy: A systematic and tripartite approach. *International Journal of Educational Technology in Higher Education*, 17(1), 22. <https://doi.org/10.1186/s41239-020-00201-6>
- Nwagwu, W. E. (2024). Mapping the field of global research on data literacy: Key and emerging issues and the library connection. *IFLA Journal*, 0(0). <https://doi.org/10.1177/03400352241257669>
- Ramos, K. J. S., Sapin, S. B., Yazon, A. D., Tamban, V. E., & Cortezano, G. P. (2023). Self-Guided Analysis Tool Pack (S-GAT) improved students' performance in statistics: Experimental approach. *Quantitative Research Journal*, 10(2), 123-135. <https://doi.org/10.31098/quant.1602>
- Schreiter, S., Friedrich, A., Fuhr, H., Malone, S., Brünken, R., Kuhn, J., & Vogel, M. (2024). Teaching for statistical and data literacy in K-12 STEM education: A systematic review on teacher variables, teacher education, and impacts on classroom practice. *ZDM - Mathematics Education*, 56(1), 31-45. <https://doi.org/10.1007/s11858-023-01531-1>
- Sheriff, N., & Sevukan, R. (2023). Exploration of data literacy research using a network of cluster mapping approach. *Journal of Scientometric Research*, 12(1), 130-143. <https://doi.org/10.5530/jscires.12.1.002>
- Umar, S., & Yakubu, K. M. (2023). Assessment of students' statistical literacy in Northwestern Nigerian polytechnics. *International Journal of Novel Research in Education and Learning*, 10(4), 92-96. <https://doi.org/10.5281/zenodo.8265293>
- Umugiraneza, O., Bansilal, S., & North, D. (2022). Analysis of teachers' confidence in teaching mathematics and statistics. *Statistics Education Research Journal*, 21(3), 1-18. <https://doi.org/10.52041/serj.v21i3.422>