ISSN Online: 3006-4708
SOCIAL SCIENCE REVIEW ARCHIVES ISSN Print: 3006-4694
https://policyjournalofms.com

Data Literacy Competencies: A Comparison between Applied and Social Science Students at the BS Level

Mr. Nasrullah [1], Dr. Anisul Hassan [2], Dr.Hafiz Muhammad Inam Ullah [3]

[1] PhD Scholar
[2] PhD Education Govt Higher Secondary School Dagai District Swabi
[3] Director IER University of Peshawar

DOI: https://doi.org/10.70670/sra.v3i3.963

Abstract

In today's society, workplace, and scientific environment, data is increasingly regarded as a highly valuable often the most valuable resource. Decisions are progressively being made on the basis of data. Digitalisation and datafication are reshaping life and work in the 21st century: artificial intelligence, networked production, communicating machines, and autonomous vehicles all rely on data while simultaneously generating vast amounts of it. Data thus forms the foundation for knowledge creation, value generation, and better decision-making. Therefore, data literacy must be cultivated from the start of higher education and across all disciplines. The present study is aimed to find out the data literacy of the university students in Pakistan furthermore to compare the data literacy of applied and social science students at BS level. For the purpose, a total six academic departments three from the applied sciences (Biotechnology, Chemistry, and Zoology) and three from the social sciences (Education, Psychology, and Sociology). A sample size of 360 (180 students from each applied and social science) were selected. IDL was used as data collection tool. Results showed that overall data literacy levels are low across both groups, with social science students demonstrating relatively higher proficiency. From inferential data analysis no significant difference in data literacy between applied science and social science students at the BS level was found. It is recommended that universities should organize workshops and training sessions for teachers; facilitated Enhancing faculty proficiency in statistical tools will directly contribute to improving students' understanding and use of data literacy software such as SPSS, STATA, and R. The Higher Education Commission of Pakistan (HEC) is encouraged to integrate a dedicated unit on data analysis software including SPSS, STATA, and R into relevant undergraduate programs.

Key Points; Data Literacy, IDL (Instrument for data literacy), SPSS (Statistical Package for Social Science)

Introduction

In our modern, data-driven society, we rely on data every day—whether it's to choose medications, evaluate health practices, decide where to live, or make informed judgments about education policy. News media frequently present data about nutrition, side effects of drugs, and political polls, all of which influence public opinion and individual choices. Clearly, developing data literacy is a crucial civic skill, especially for students preparing to participate actively in an increasingly complex world. At the heart of data literacy is the concept of data itself. But what exactly is data? Definitions vary across disciplines. According to Bhargava et al. (2015) and Koltay (2017), data can be understood as any object, variable, or piece of information that can be collected, stored, and identified. Similarly,

the National Science Board (2005) and the Department of National Defense and Canadian Armed Forces Data Strategy (2019) broaden this definition by including various forms of data—text, numbers, algorithms, animations, images, audio, software, movies, and simulations.

The term literacy, on the other hand, has evolved over time. While it was once considered solely the ability to read, more recent definitions encompass reading, writing, speaking, and listening. Scholars across disciplines have struggled to agree on a single definition of literacy, due in part to differing cultural, historical, economic, and political contexts (Keefe & Copeland, 2011). For example, in the middle Ages, literacy was associated with the ability to read, write, and speak Latin—a skill reserved for an elitefew with access to formal education (Soares, 1992). The invention of the printing press in the 16th century shifted the focus to literacy in vernacular languages, making it a broader societal need (Heath, 1996).

Modern perspectives on literacy go beyond traditional definitions. Schield (2002a) defines literacy as the ability to review, interpret, analyze, and evaluate written material. Fransman (2005) emphasizes the importance of context in shaping our understanding of literacy, suggesting that definitions shift depending on the setting. In a cognitive context, literacy includes the skills of reading, writing, and oral communication (Adams, 1993; Goodman, 1996; NEP, 2017; Street, 2004). Further, Gee (1999), Barton and Hamilton (1999), and Street (1998) argue that literacy also involves applying these skills in real-life situations.

This view is echoed in UNESCO's (2013) widely accepted definition—cited by Coles (2013)—which describes literacy as the ability to read, write, and understand simple statements related to daily life (Research Brief Series, 2015). and write simple statements related to his/her daily life (Research Brief Series, 2015).

Data Literacy

Ridsdale et al. (2015) define data literacy in terms of skills and competencies, describing it as the ability to collect, manage, evaluate, and apply data in a critical manner. Similarly, Wolff et al. (2016) define data literacy in terms of skills and outcomes, emphasizing the ability to ask and answer real-world questions using both large and small data sets through an inquiry-based process, while also considering the ethical use of data. Bonikowska, Sanmartin, and Frenette (2019) view data literacy through the lens of digital competencies, defining it as the ability to derive meaningful insights from data.

Likewise, Canada's Department of National Defense (2019) outlines a broad definition of data literacy, including skills such as discovering and accessing data, manipulating data, evaluating data quality, analyzing data, interpreting results, and understanding the ethical implications of data use. Sperry (2018) further explains that a data-literate individual should at minimum be able to interpret basic statistical summaries, perform simple calculations, and use these statistics to inform decision-making.

According to Hassan A. (2019), students need various tools to effectively use data, including statistical software like SPSS, STATA, Minitab, and Microsoft Excel; data presentation tools like MS Excel and PowerPoint; relational databases such as MS Access; and knowledge of Structured Query Language (SQL). These tools help students to collect, analyze, and represent data through statistical techniques, graphs, and tables.

At the university level, students often use data for research purposes. They employ approved data collection tools; analyze data using software, and present results using appropriate visualizations. In today's technology-driven society, data analysis has become more accessible due to software like SPSS. University students frequently conduct research in their areas of interest; however, without a solid foundation in data literacy—especially in quantitative research—such work cannot be effectively accomplished.

Therefore, data literacy is fundamental to academic research. This paper focuses on the comparison of data literacy levels between students in applied sciences and social sciences at the BS level, which

in Pakistan serves as the entry point to research at the university level. At this stage, students are expected to extract information from data, summarize it into simple statistics, and use the results to make informed decisions. Hence the present study is aimed to find out the data literacy of the university students and to compare the data literacy of applied and social science.

Research Questions and Hypothesis

The following three research questions and one null hypothesis were tested;

1. What is the data literacy of applied science students?
2. What is level the data literacy of Social science students?
3. What is the data literacy of university students?

H_1: There is no significant difference in the data literacy of applied and social science students.

Population

According to the Higher Education Commission of Pakistan (2023), there are 40 universities and degree-awarding institutions (DAIs) in Khyber Pakhtunkhwa province-29 in the public sector and 11 in the private sector. For the purpose of this study, six academic departments were selected: three from the applied sciences (Biotechnology, Chemistry, and Zoology) and three from the social sciences (Education, Psychology, and Sociology).

Among the 40 universities and DAIs, only nine public sector universities offered all six of the selected departments. Within these institutions, 3,115 students were enrolled in the applied science departments, while 2,997 were enrolled in the social science departments. Therefore, the total population for this study consisted of 6,112 BS-level students from the six departments.

Sample

From the nine eligible public sector universities, six were randomly selected for inclusion in the study. Using the Raosoft online sampling calculator with a 95% confidence level, the recommended sample size was 307. However, to ensure adequate representation, a total of 360 students were selected—180 from applied sciences and 180 from social sciences, with 60 students from each department. This sample size is appropriate for a survey-based study, as Creswell (2012) suggests that a sample of approximately 350 participants is sufficient for such research. Data were collected using the IDL (Instrument for Data Literacy), which was specifically designed to measure students' competencies in data literacy

Results

Table 1:

Responses on IDL

	Applied Science (180)				Social Science (180)			
Item #	Minimum	Maximum	Mean	Std. Deviation	Minimum	Maximum	Mean	Std. Deviation
Category 1	0	1	0.19	0.397	0	1	0.36	0.480
Category 2	0	1	0.18	0.383	0	1	0.17	0.374
Category 3	0	1	0.13	0.335	0	1	0.10	0.301
Category 4	0	1	0.13	0.341	0	1	0.28	0.452
Cumulative	0	3	0.63	0.95	0	3	0.90	0.93

Table 1 presents the comparative responses of students from applied sciences and social sciences on

the Instrument for Data Literacy (IDL).

Category 1 assessed familiarity with statistical software used for data analysis. Social science students had a higher mean score (M = 0.36) compared to applied science students (M = 0.19), indicating greater familiarity with such tools. The standard deviation was also higher for social science students (SD = 0.480), suggesting more variability in their responses.

Category 2 focused on understanding the purpose of columns in SPSS. Applied science students scored slightly higher (M = 0.18) than social science students (M = 0.17), indicating a marginally better understanding. However, the variation was slightly greater among applied science students (SD = 0.383) compared to their counterparts (SD = 0.374).

Category 3 evaluated knowledge of row representation in SPSS. Again, applied science students had a slightly higher mean score (M = 0.13) than social science students (M = 0.10), reflecting a minor advantage. The standard deviation was higher in the applied science group (SD = 0.335) than in the social science group (SD = 0.301).

Category 4 examined the ability to calculate mean values in SPSS. Social science students performed notably better (M = 0.28) than applied science students (M = 0.13), indicating stronger competency in statistical calculations using software.

The cumulative mean scores—0.63 for applied science students and 0.90 for social science students suggest that overall data literacy levels are low across both groups, with social science students demonstrating relatively higher proficiency.The data indicate that social science students are more data literate than their applied science counterparts, particularly in the use and operation of statistical tools such as SPSS. However, the overall low mean scores across all categories and both groups highlight a general deficiency in data literacy skills at the BS level. This finding underscores the need to strengthen data literacy education across disciplines, especially in quantitative research contexts.

Table 2:

Comparison of the responses on IDL

Groups	N	Mean	Std Deviation	Independent sample t-test α=0.05		
				Df	t	Sig
Applied Sciences	180	0.63	0.95	358	-.565	0.596
Social Sciences	180	0.90	0.93			

Table 2 compares the data literacy scores of students in applied sciences and social sciences using an independent samples t-test at a 0.05 level of significance. The mean score for social science students (M = 0.90, SD = 0.93) was higher than that of applied science students (M = 0.63, SD = 0.95), indicating that social science students performed better overall in data literacy as measured by the IDL.

However, the p-value (Sig. = 0.596) is greater than the significance level of 0.05, indicating that the difference in mean scores between the two groups is not statistically significant. Therefore, it can be inferred that there is no significant difference in data literacy between applied science and social science students at the BS level.

Discussions

This study examined BS-level students' responses regarding their ability to use and operate data literacy tools such as SPSS. The findings revealed that the majority of students—regardless of

discipline—demonstrated below-average competency in data literacy, particularly in using SPSS for data analysis. These results align with the findings of Craig, Elsa, and Evelyn (2012), who reported that only 19% of graduate students in the Philippines had basic knowledge of SPSS. In contrast, the study by Gonulal, Loewen, and Plonsky (2017) found that approximately 83% of university students in North America could understand, use, and operate SPSS for statistical analysis. This contrast suggests that regional differences and curriculum emphasis may significantly influence students' data literacy competencies.

Recommendations

Based on the study's findings, the following recommendations are proposed:

1. Training Workshops for Faculty:

Universities should organize workshops and training sessions for teachers, facilitated by the National Academy of Higher Education (NAHE). Enhancing faculty proficiency in statistical tools will directly contribute to improving students' understanding and use of data literacy software such as SPSS, STATA, and R.

2. Curriculum Enhancement:

The National Curriculum Review Committee (NCRC) of the Higher Education Commission (HEC) is encouraged to integrate a dedicated unit on statistical data analysis software—including SPSS, STATA, and R—into relevant undergraduate programs. This would ensure students are formally introduced to these tools during their academic journey.

References

1. Adams, M 1993 'Beginning to read: an overview' in Beard, R Teaching Literacy Balancing Perspectives Hodder & Stoughton: London 'Introduction' pp. 204-215
2. Bhargava, Rahul, et al. (2015). Beyond Data Literacy: Reinventing Community Engagement and Empowerment in the Age of Data." Data-Pop Alliance, Data-Pop Alliance https://datapopalliance.org/item/beyonddata-literacy-reinventing-community-engagement-and-empowerment-in-the-age-of-data.
3. Coles, P. (2013). Global media and information literacy assessment framework: Country readiness and competencies.
4. Creswell, J. W. (2012). Educational research: Planning, conducting, and evaluating quantitative. 4th Edition Upper Saddle River, NJ: Prentice Hall. Pp. 146-166.
5. Fransman, J. (2005). Understanding literacy: a concept paper. The EFA Global Monitoring Report, Literacy for Life, 31.
6. Goodman, K. (1996) On Reading Scholastic Canada. Ch 8 'Learning and Teaching Reading and Writing' Pp. 117-125.
7. Hassan A, (2019), Investigating the Relationship between University Students‟ Statistical Literacy and Attitudes Towards Research. Phd Thesis, Department of Education Faculty 0f Social Sciences Abdul Wali Khan University Mardan.
8. Keefe, E. B., & Copeland, S. R. (2011). What is literacy? The power of a definition. Research and Practice for Persons with Severe Disabilities, 36(3-4), 92-99.
9. Koltay, T. (2017). Data literacy for researchers and data librarians. Journal of Librarianship and Information Science, 49(1), 3-14.
10. National Defence. The Department of National Defence and Canadian Armed Forces Data Strategy Canada, (2019), https://www.canada.ca/en/department-national-defence/corporate/reports-publications/data-strategy.html.
11. National Science Board (2005) Long-Lived Digital Data Collections: Enabling Research and Education in the 21
12. National Education Policy (2017) Government of Pakistan Ministry of Federal Education and Professional Training Islamabad.P.32.

13. Soares, M. B. (1992). Literacy Assessment and Its Implications for Statistical Measurement. Current Surveys and Research in Statistics.
14. Schield, M. (2002a). Three kinds of statistical literacy: what should we teach? ICOTS6. Durban, South Africa.
15. Street, B. 2004 Understanding and defining literacy**.** Background Paper for EFA Global Monitoring Report 2006.
16. Ridsdale, Chantel, et al. (2015), Strategies and Best Practices for Data Literacy Education. Dalhousie University.
17. Sperry, J. (2018). Data Literacy: Exploring Economic Data. U.S. Census Bureau, St. Mary's. St. Mary's Practice Test.
18. Wolff, Annika, et al.(2016,), Creating an Understanding of Data Literacy for a Data-Driven Society. Journal of Community Informatics, vol. 12, no. 3.