EJMR
ERA'S JOURNAL OF MEDICAL RESEARCH
p-ISSN: 2348-9839, e-ISSN: 2394-5222 | www.ejmr.org

Volum 12 Issue 1
ORIGINAL ARTICLE

Evaluating the Impact of Data Analytics Training in Health Professions Education: A Pretest-Posttest Analysis

Shahid Hassan[1], Sonia Jaiswal[2], Gauhar Hassan[3]

Received on: 05-02-2025 Published on: 30-04-2025

ABSTRACT

Background: In health professions education, data analytics is increasingly recognized as a critical skill for educators to enhance decision-making and improve student outcomes. A session on data analytics was conducted as part of a Postgraduate Certificate in Health Professions Education (PGCHPE) program to equip educators with foundational skills in data analysis and interpretation. Immediate effectiveness of educational interventions is critical for justifying delivery of a course or workshop to readjust the learning outcomes. Pretest-posttest data analytics offers an evidence-based approach to evaluate short-term learning gains.

Aims and Objectives: This evaluation aimed to determine whether the training effectively improved participants' understanding of core data analytics concepts and their ability to apply these skills in educational settings. The objective of this study was to determine the immediate impact of the session using a pretest-posttest design to assess participants' knowledge and confidence before and after the session.

Methods: A one-group pretest-posttest design was employed. Participants (n = 22) enrolled in the Postgraduate Certificate Programme of Health Profession Education (PGCHPE) completed a knowledge-based assessment immediately before and after the intervention. Descriptive statistics, paired sample t-tests, and effect size (Cohen's d) were used to analyse the data.

Results: Mean pretest score was 70.90 (11.72), which increased to 86.34 (11.58) in the posttest. The difference was statistically significant, $t = 3.18$ (9), $p < 0.05$). The effect size (Cohen's $d = 1.009$) indicated a large immediate impact of the intervention.

Conclusions: The pretest-posttest data analytics confirmed a statistically significant immediate improvement in learners' knowledge. This supports the intervention's effectiveness and provides justification for its continued or expanded use.

Keywords: *Data analytics, Educational intervention, Effect size, Immediate impact, Pretest-posttest.*

Era's Journal of Medical Research. 12(1);2025 [doi: 10.24041/ejmr.2025.2]

INTRODUCTION

In contemporary medical education, the utilization of data analytics has emerged as a cornerstone for enhancing the quality and effectiveness of teaching and learning. Online assessment platforms generate a wealth of data that can be harnessed to provide actionable insights into student performance and curricular alignment. By analysing assessment data through targeted parameters, educators can make informed decisions to optimize both instructional strategies and learner outcomes. With regular practice of learning analytics in curriculum and assessment evaluation we can predict students' performance. Leveraging learning analytics to support the medical education continuum can provide faculty of indigenous data that can be used for evidence-based decisions to predict future trends to be accommodated in periodic and major curriculum review.[1]

Assessment data encompasses diverse metrics, including item-level statistics, student responses, and overall test performance. Categorizing this data based on parameters such as learning outcomes, subject-wise distribution of items, Bloom's Taxonomy and the Accreditation Council for Graduate Medical Education (ACGME) six core competencies[2] provides a structured framework for interpretation. For instance, analysing items according to Bloom's Taxonomy enables the evaluation of cognitive levels assessed, ranging from knowledge recall to application and synthesis, thus ensuring alignment with desired educational objectives[3] Similarly, mapping items to the ACGME six core competencies allows for a focused assessment of domains such as medical

[1]Department of Anatomy, American University of Barbados School of Medicine, School of Medicine, Barbados.
[2]Department of Medical Education, American University of Barbados School of Medicine, School of Medicine, Barbados.
[3]Department of Anatomy, Era's Lucknow Medical College and Hospital, Era University, Lucknow, UP, India
Corresponding Author: Shahid Hasan
Email: shahid21225@gmail.com
How to cite: Hassan S, Jaiswal S, Hassan G. Evaluating the Impact of Data Analytics Training in Health Professions Education: A Pretest-Posttest Analysis. Era J Med Res. 2025:12(1);6-12.

knowledge, patient care, professionalism, interpersonal and communication skills, and systems-based practice and practiced-based learning and improvement, fostering comprehensive competency-based medical education. Professionalism can best be evaluated through students attitude towards continuing and life -long learning behaviour, which is a difficult area of students' performance evaluation using numeric data. The qualitative inferences drawn about their learning behaviour observed for attitude and professionalism[4] towards punctuality (option of digital attendance), visit to post instructional uploaded self-assessment, online formative assessment and enrolling in online provided training such as continuous professional development (CPD).

Faculty and institutional preferences further enrich the analytical process, permitting customization of parameters to align with specific educational goals.[5] These analytics not only identify areas for student improvement but also highlight trends in question design and curriculum delivery, paving the way for evidence-based curricular reforms. By leveraging data analytics, medical educators can transition from traditional assessment review methods to a more dynamic, data-driven approach, thereby advancing the standards of medical education.

Training faculty in data analytics is paramount in effectively implementing CBME practices. Faculty must be equipped to interpret complex assessment data, monitor student progression, and identify gaps in both learner and curriculum performance.[6,7] Data analytics empowers educators to track competency milestones, ensure timely feedback, and tailor instructional strategies to individual student needs. Continuous peer feedback and competency analytics in the workplace can help employees develop transversal skills and focus on performance objectives.[8,9] Moreover, it facilitates a culture of continuous quality improvement by highlighting trends and inefficiencies in curriculum delivery, assessment design, and teaching practices. Faculty development in data analytics thus becomes a critical enabler for the successful adoption of CBME, ensuring that competency frameworks translate into meaningful educational outcomes.

OBJECTIVE

The main objective was to determine the impact of data analytic session for immediate transfer of knowledge.

The specific objective was to measure the effect size to authenticate it's clinical and practical usability.

Hypothesis

1. *Null Hypothesis (H0):* There is no significant difference between the pretest and posttest scores.
2. *Alternative Hypothesis (Ha):* There is a significant difference between the pre/posttest scores.

Research Methodology

A pretest-posttest quasi-experimental design was employed to evaluate the effectiveness of the data analytics session. The participants, enrolled in the PGCHPE program, were given a structured pretest prior to the session to assess their baseline knowledge and confidence in data analytics. Following the session, a posttest with a similar structure was administered to evaluate any changes in their performance after the intervention.

The session content included key concepts of data analytics, practical applications, and the interpretation of results in an educational context. The tests comprised multiple-choice questions items, measuring confidence and perceived competence in data analytics.

Data Analysis

The data collected from the pretest and posttest were analysed using a paired t-test to compare mean scores and assess the significance of the changes. The paired t-test was chosen because it accounts for the repeated measures nature of the data, where the same participants were tested before and after the intervention. Following are the steps of analysis:

Data Preparation: The pretest and posttest scores were downloaded as Microsoft Excel sheet to export as SPSS data sheet from the online survey using google form.

Descriptive Statistics: Mean and standard deviation were calculated for both pretest and posttest scores to summarize the data.

Paired t-test statistic was calculated using the Statistical Package for the Social Sciences (SPSS 29).

In addition to t-test, Cohen's d was calculated to determine the magnitude of the effect.

Significance Level: A *p-value* of <0.05 was considered statistically significant.

RESULT

The % correct pretest and posttest data opened in Microsoft Excel was downloaded from the online survey prior to session and after the session, respectively (Table 1). Data was analysed using paired t-test and found significant at $p = .011$ (Table 2). The result of pair t-statistics indicates a mean difference of 15.44 between the pretest score = 70.90 (11.58)) and post-test score = 86.34 (11.72), respectively (Table 2) significant at, $t = 3.189$, $p < 0.05$ (Table 2) associated with a poor correlation of 0.081. The Cohen's d as the indicator of effect size = 1.009 is significant for its practical or clinical importance (table 3). For individual items analysis please see the appendix.

DISCUSSION

Pretest-posttest analysis evaluates the effectiveness of an instructional session by comparing participants' knowledge or skills before and after the intervention has been established.[10] The analysis of the pretest and posttest data (Table 1) in current study revealed a statistically significant

Table 1: %correct pretest and postest score result significant at <0.05.

Pre/Postest Item	Pretest Result	Postest Result
	% Correct Score	% Correct Score
Question 1	68.2	90.5
Question 2	68.2	59.1
Question 3	50	90.9
Question 4	72.7	90.9
Question 5	90.9	95.5
Question 6	63.6	90.9
Question 7	72.7	77.3
Question 8	63.6	95.5
Question 9	86.4	95.5
Question 10	72.7	77.3
Average	70.9	86.3
Correlation	.137	

improvement in participants' knowledge and confidence following the data analytics session (p = 0.014, <0.05) (Table 2). The effect size, as measured by Cohen's d = 0.959, was notably high, indicating a strong impact of the session on the participants' learning outcomes. Cohen's d is a standardized effect size measure used to indicate the magnitude of the difference between two means, commonly interpreted as small, medium, or large.[11] The findings from this analysis provide insights into the session's effectiveness noted with high impact, and highlight areas for future improvement in designing targeted educational interventions in data analytics.

Analysing the individual questions (see appendix) all the items except item 2 (see question 2 in appendix) have shown a positive impact on knowledge after the intervention. Item 2 has shown a reverse impact and deterioration of knowledge from what it existed prior to the session and after the intervention of 2-hours long interactive session in selecting the correct answer. This item apparently simple has asked about the example of descriptive analysis, which refers to summary of the past data but, many chose the option of predicting students dropout. Surprisingly word "predicting" deceived them instead of giving the cueing effect of predicting being referred to forecasting the future trend cannot be a descriptive analytic. In this question the other two options were about recommend tailored study resources and interventions for improvement selected by 1 (4.5%) each and seems more deceiving and plausible than the easy to reject predictive option selected predominantly by respondents in postest. However, a big number of 49.9% remained incorrect shows misconception not cleared and perhaps participants need more interaction to define and develop right concepts and differentiation among the perceptive, predictive and perspective learning analytics.

The situation observed in item 2 can be further explored studying item 5 on the same topic of predictive analytics, where majority 90.9% in pretest and 95.5% in postest (with a minor difference of 4.6) have rightly answered by selecting the option, "forecasting future outcomes based on patterns." Here words forecast and prediction provided them the clue, which was ignored by the respondents in item 2. Item 10 (see appendix question 10) has same score distribution and the difference between the pretest and postest as item 7 but does not need elaboration because it is about software among the option and power BI is maintained in both the tests may be attributed to limitation of knowledge about a software less commonly practices by the participants may have been answered by guessing.

A similarly close response in item 7 (see appendix question 7) has been noted. Here question about ethical consideration also straight forward with a correct option of ensuring, "data privacy and confidentiality" with the very words used leading to obvious cueing effect was misconceived for, sharing student data with, "external agencies" though was

Table 2: Paired t-statistics to determine the significant difference of means between pretest and postest knowledge about the learning analytics.

Pair Pretest and Postest Score	Pretest Mean (SD)	Postest Mean (SD)	Mean (SD) Score Difference (95%CI)	t-statistics (df)	*p*-value	Cohen's d
N =22	70.90 (11.72)	86.34 (11.58)	15.44 (15.30) (4.48, 26.39)	3.18 (9)	.011	1.009

Table 3: Cohen's d range for its effect size, interpretation and for their practical importance.

Cohen's d range	Interpretation	Practical/Clinical importance
d<0.2	Very Small/ Negligible	Effect is likely trivial in most contexts but might matter in highly sensitive or specific settings (e.g., rare outcomes).
d>0.2 — ≥0.5	Small Effect	Detectable but modest; practical importance depends on the context (e.g., population-level changes or low-cost interventions).
d>0.5—<0.80	Medium Effect	Moderately noticeable; often practical or clinically relevant depending on benefits, risks, and costs.
d≥0.8	Large Effect	Substantial and obvious; likely to have strong practical or clinical significance.

very plausible and close to the correct response, identified in this survey. Besides, the misconceptions developed in items 2 and 7 due to their being highly plausible to correct answer are suggestive of participants deep-dived approach reading each options with utmost attention and overindulgence with a newly acquired knowledge in a difficult subject area that often a reason for respondent to hit the wrong option, nevertheless makes the item high on the discrimination index.

However, the small number of participants (22 out of 23 registered) is a limitation that may affect the generalizability of the findings. This reduced sample size underscores the importance of encouraging higher participation in future studies to strengthen the reliability of results.

CONCLUSION

A significant increase in posttest scores, along with a large effect size, indicated that the data analytics session had a meaningful impact on participants' learning outcomes. This supports the intervention's effectiveness and provides justification for its continued or expanded use.

Limitation

Despite the limitation of number of participant, the significant *p-value* and large effect size suggest that the session was effective in meeting its learning objectives and equipping participants with essential data analytics skills for educational settings.

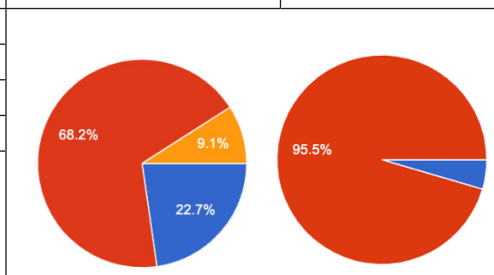
Recommendations

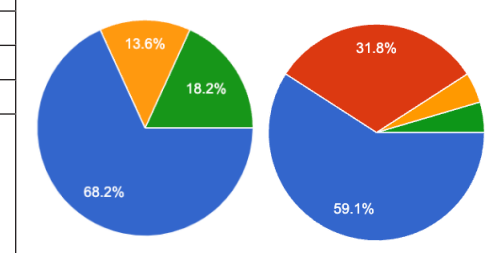
Further studies with larger sample sizes are recommended to validate these findings and refine the session's content and delivery.

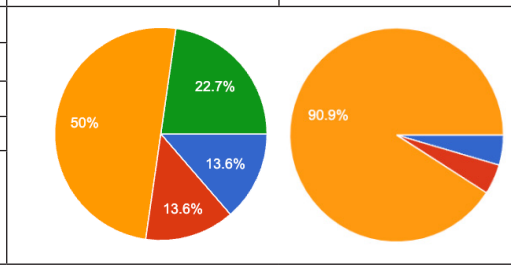
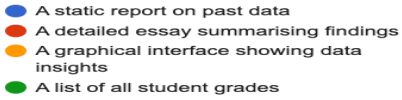
Acknowledgement

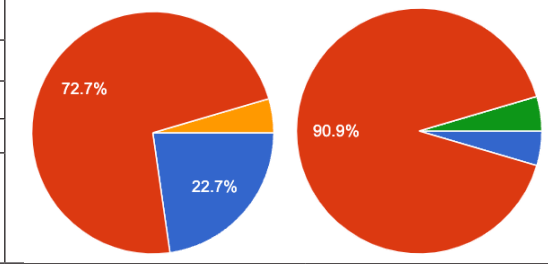
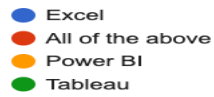
My appreciation for all those attending the session and particularly those responding to complete the pretest and postest survey online.

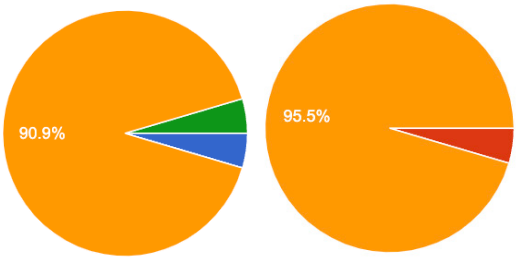
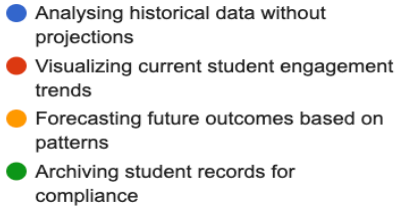
Appendix: Individual items analysis for correct response in pretest and postest presented with pie chart together with the question and the option list.

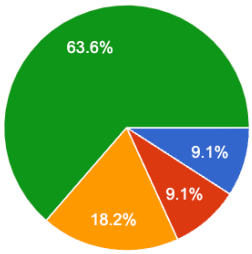
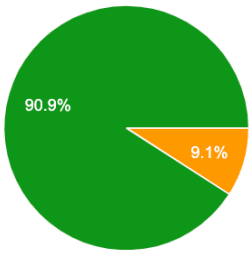
Question No 1	What is the primary purpose of data analytics in medical education?	Pretest % Correct 68.2%	Postest % Correct 95.5%
Options		Impact (High +)	+22.3
A	To automate grading processes		
B	To identify trends for informed decision.		
C	To reduce faculty workload.		
D	To replace human judgment		
Correct **B**	To automate grading processes To identify trends for inform decision To reduce faculty workload To replace human judgment		
Interpret.	A positive impact of session indicated by 22.3% increment in knowledge however, a small number of 4.5% though, are still incorrect.		

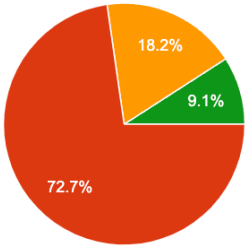
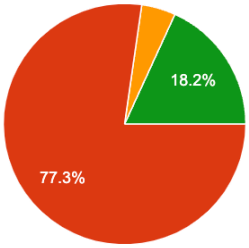
Question No 2	Which of the following is an example of descriptive analytics?	Pretest % Correct 68.2%	Postest % Correct 59.1%
Options		Impact (Minor -)	-9.1
A	Identifying the top-performing students		
B	Predicting student dropout rates.		
C	Recommending tailored study resources.		
D	Creating early intervention plans.		
Correct **A**	Identifying the top-performing students Predicting student dropout rates Recommending tailored study resources Creating early intervention plans		
Interpret.	A misconception shown by reverse impact of intervention indicated by a deterioration of knowledge by -9.1% from the existing concept recorded in data analytics and a big number of 49.9% remained incorrect shows misconception not cleared (see the interpretation).		

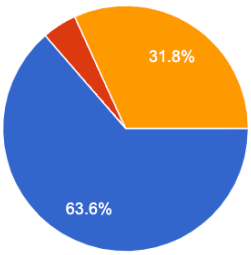
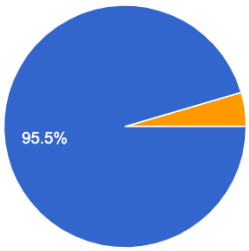
Question No 3 Options	In data visualization, what does a dashboard typically represent?	Pretest % Correct 50%	Postest % Correct 90%
		Impact (Major +ive)	+40
A	A static report on past data.		
B	A detailed essay summarizing findings.		
C	A graphical interface showing data insights.		
D	A list of all student grades.		
Correct C			
Interpret.	A positive impact of session has been noted with a high improvement of knowledge by 40% however, a small number of 9.1% still incorrect.		

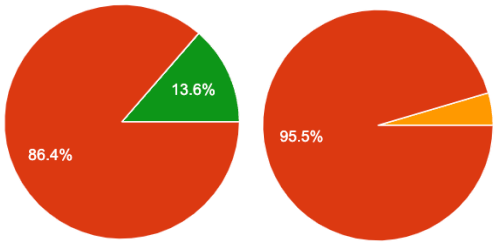
Question No 4 Options	Which tool is commonly used for data visualisation in education?	Pretest % Correct 72.7%	Postest % Correct 90.9%
		Impact(Moderate +)	+13.1
A	Excel.		
B	All of the above		
C	Tableau.		
D	Power BI		
Correct B			
Interpret.	A positive impact of session indicated by 13.1% increment in knowledge however, a small number of 9.1% though, are still incorrect.		

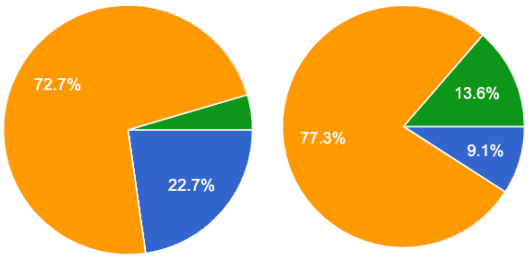
Question No 5 Options	What does "predictive analytics" in education focus on?	Pretest % Correct 90.9%	Postest % Correct 95.5%
		Impact (Minor +ive)	+4.6
A	Analysing historical data without projections		
B	Visualizing current student engagement		
C	Forecasting future outcomes based patterns		
D	Archiving student records for compliance		
Correct C			
Interpret.	A positive impact of session minor though, indicated by 4.6% increment in knowledge however, a small number of 4.5% though, are still incorrect.		

Question No 6 Options	In Moodle and Canvas, what feature helps identify students at risk of failing?	Pretest % Correct 63.6%	Postest % Correct 90.9%
		Impact (Minor -ive)	+27.3
A	Learning Mastery Gradebook		
B	Behavior Tracking Module.		
C	Automatic Grading System		
D	Predictive Analytics Dashboard		
Correct D	Learning Mastery Gradebook Behaviour Tracking Module Automatic Grading System Predictive Analytics Dashboard		
Interpret.	A positive impact of session indicated by 27.3% increment in knowledge however, a small number of 9.1% though, are still incorrect.		

Question No 7 Options	Which of the following is a critical ethical consideration in data analytics?	Pretest % Correct 72.7%	Postest % Correct 77.3%
		Impact (Minor -ive)	+4.6
A	Maximizing faculty workload.		
B	Ensuring data privacy & confidentiality		
C	Collecting as much data as possible		
D	Sharing student data with external agencies		
Correct B	Maximizing faculty workload Ensuring data privacy and confidentiality Collecting as much data as possible without limits Sharing student data with external agencies		
Interpret.	A positive impact of session minor though, indicated by 4.6% increment in knowledge however, a moderate number of 22.7% though, are still incorrect.		

Question No 8 Options	What is the primary advantage of using analytics to track lifelong learning?	Pretest % Correct 63.6%	Postest % Correct 95.5%
		Impact (Minor +ive)	+31.9
A	Promoting engagement in CPD.		
B	Monitoring faculty workload		
C	Automatically assigning course to students.		
D	Ensuring 100% attendance in lectures		
Correct A	Promoting engagement in CPD Monitoring faculty workload Automatically assigning courses to students Ensuring 100% attendance in lectures		
Interpret.	A positive impact of session indicated by 31.9% increment in knowledge however, a small number of 4.5% though, are still incorrect.		

Question No 9 Options	How can faculty use behavioural data in analytics?	Pretest % Correct 86.4%	Postest % Correct 95.5%
		Impact (Minor -ive)	+11.1
A	To reduce class sizes.		
B	To track identify disengaged students		
C	To calculate final grades		
D	To eliminate the need for feedback		
Correct B	To reduce class sizes To track identify disengaged students To calculate final grades To eliminate the need for feedback sessions		
Interpret.	A positive impact of session has been noted with an improvement of knowledge in data analytics of 5.4% however, 10% still incorrect		

Question No 10 Options	Which software can integrate with both Moodle and Canvas for advanced analytics?	Pretest % Correct 72.7%	Postest % Correct 77.3%
		Impact (Minor +ive)	+4.6
A	Zoom		
B	Photoshop		
C	Power BI		
D	H5P		
Correct C	Zoom Photoshop Power BI D. H5P Option 5		
Interpret.	A positive impact of session has been noted with a high improvement of knowledge by +10.8% in data analytics however, 20% still incorrect		

REFERENCES

1. Artino AR, Holmboe ES, Durning SJ. Can we predict performance? Leveraging learning analytics to support the medical education continuum. Academic Medicine. (2020); 95(5), 678-685.
2. Frank JR, Snell LS, Sherbino J. CanMEDS Physician Competency Framework. Royal College of Physicians and Surgeons of Canada. Online. (2015); Available: https://www.royalcollege.ca
3. Bloom BS. Taxonomy of Educational Objectives: The Classification of Educational Goals. New York: Longmans, Green. 1956; ISBN-10 0679302093
4. Mak-van der Vossen M, Peerdeman S, Kusurkar RA. The role of assessment in promoting medical professionalism. Perspectives on Medical Education. 2020; 9(4), 209-213. https://doi.org/10.1007/s40037-020-00600-y
5. Harden RM, Laidlaw JM. Essentials of Medical Education. Edinburgh: Elsevier. 2020.
6. Thoma B, Ellaway RH, Chan TM. From utopia through dystopia: charting a course for learning analytics in competency-based medical education. Academic Medicine. 2021 Jul 1;96(7S):S89-95.
7. Chan Teresa, Sebok-Syer Stefanie, Thoma Brent, Wise Alyssa, Sherbino Jonathan, Pusic Martin. Learning Analytics in Medical Education Assessment: The Past, the Present, and the Future. AEM Education Training. 2018; 2(2).
8. Walsh E, Neelen M, Kapros E. Competency Analytics in the Workplace through Continuous Peer Feedback. In2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT) 2017 Jul 3 (pp. 224-226). IEEE.
9. Kumar V, Boulanger D, Seanosky J, Kinshuk, Panneerselvam K, Somasundaram TS. Competence analytics. Journal of Computers in Education. 2014 Dec;1:251-70.
10. Hassan S, Venkateswaran SP, Nadarajah VD. Evaluation of immediate impact of Faculty Development Programme using a pretest–post-test study design format. Innovation and Education. 2021 Apr 21;3(1):1.
11. Cohen J. Statistical power analysis for the behavioral sciences. routledge; 2013 May 13.

Orcid ID:
Shahid Hassan - https://orcid.org/ 000000032982660X
Sonia Jaiswal - https://orcid.org/ 0000-0003-2291-6178
Gauhar Hassan - https://orcid.org/ 0009-0000-0275-0436