East African Journal of Education and Social Sciences
EAJESS January-February 2025, Vol. 6, No. 1, pp. 88-101.
ISSN: 2714-2132 (Online), 2714-2183 (Print). Published by G-Card
DOI: https://doi.org/10.46606/eajess2025v06i01.0427.

Determinants of Students' Performance in Electrical and Electronics Engineering at Mbeya University of Science and Technology, Tanzania

Zacharia Katambara
ORCiD: https://orcid.org/0000-0001-9983-1207
Department of Civil Engineering, Mbeya University of Science and Technology, Tanzania
Email: zacharia.katambara@must.ac.tz

Abstract: The Electrical and Electronics Engineering program requires a balance between theoretical knowledge and practical application, making students' performance optimization essential in meeting industry demands. This study utilized descriptive statistics, Pearson Correlation Analysis, and Principal Component Analysis (PCA) to evaluate academic performance in the EEE program at Mbeya University of Science and Technology (MUST). By examining 16 core courses, the study identified key determinants of students' success, course interdependencies and areas for curriculum enhancement. Descriptive statistics revealed significant variability in performance, with EE 8401 (Industrial Practical Training 3) recording the highest mean (79.98) and EE 8402 (Phase AC Synchronous Machines) the lowest (48.11), highlighting disparities in instructional effectiveness. Pearson Correlation Analysis shows strong correlations among theoretically aligned courses, moderate correlations among related subjects, and weak or negative correlations in distinct learning domains, emphasizing the need for targeted interventions and curriculum adjustments. PCA findings confirmed that three Principal Components explained 58.85% of the variance, representing theoretical foundations, applied project-based learning and specialized hands-on training. Scree plot and eigenvalue analysis validated dimensionality reduction, enhancing data interpretation. Principal Component loadings highlight academic constructs, with PC1 reflecting analytical competencies, PC2 capturing project-based courses and PC3 representing specialized training. This study recommends aligning theoretical courses with standardized assessments, integrating industry collaborations in project-based learning and refining assessment models for specialized training. Future research should explore longitudinal trends in Principal Components, external influences on high-uniqueness courses and students' feedback integration. By implementing data-driven strategies, institutions can refine engineering curricula, bridge performance gaps and enhance student success outcomes.

Keywords: Principal component analysis; engineering education; academic Performance.

How to cite: Katambara, Z. (2025). Determinants of Students' Performance in Electrical and Electronics Engineering at Mbeya University of Science and Technology, Tanzania. East African Journal of Education and Social Sciences 6(1), 88-101. **DOI:** https://doi.org/10.46606/eajess2025v06i01.0427.

Introduction

The Electrical and Electronics Engineering (EEE) program is among the most technically rigorous and impactful disciplines, requiring students to integrate complex theoretical concepts with practical applications. This field plays a crucial role in addressing global challenges, such as renewable energy development, industrial automation and adoption of emerging technologies like the Internet of Things (IoT) and Artificial Intelligence (AI) (Alawin et al., 2016; Nurhidayat, et al., 2024). With the increasing demand for highly skilled professionals in these domains, assessing and enhancing students' performance remains a critical priority in engineering education (Kabakchieva, 2012). Traditional assessment methods often fail to capture the multifaceted nature of academic success, necessitating data-driven analytical approaches to identify key determinants of student performance.

Principal Component Analysis (PCA) has emerged as a powerful multivariate statistical technique for evaluating academic performance by simplifying complex datasets and extracting essential performance determinants. PCA enables educators to analyze interdependencies among courses, identify key predictors of success and offer actionable insights for curriculum enhancement (Borges et al., 2018; Alam & Khatun, 2021).

Prior research suggests that foundational knowledge in mathematics and physics, teaching methodologies and technological proficiency are critical factors influencing students' performance (Mills & Treagust, 2003; Felder & Brent, 2024). PCA has also demonstrated strong correlations between active learning approaches, such as problem-based and project-based learning and improved academic outcomes (Mills & Treagust, 2003).

Moreover, technological integration, including simulation tools and virtual laboratories like MATLAB and Simulink, significantly enhances students' ability to tackle real-world engineering challenges (Nurhidayat et al., 2024; Li and Liang et al., 2024). Active students' engagement in academic and extracurricular activities further strengthens learning outcomes, as evidenced by PCA analysis (Felder & Brent, 2024; Kuh et al., 2008). Additionally, timely and constructive feedback is essential in mastering complex engineering concepts and preparing students for the industry's evolving demands (Nicol & Macfarlane-Dick, 2006; Gibbs & Simpson, 2005). By leveraging PCA, this study sought to establish the most influential factors affecting student performance in the Electrical and Electronics Engineering programme at Mbeya University of Science and Technology (MUST), informing targeted interventions and optimising teaching strategies.

Literature Review

Academic performance analysis is crucial in understanding students' success factors and developing effective educational strategies. Traditionally, descriptive statistics, regression analysis and correlation studies have been employed to assess academic performance (Field, 2024). For instance, Martínez-Cervantes et al. (2013) examined the effect of basic infrastructure on academic achievement in Mexican technological high schools. Using regression analysis and structural modelling, they found that an increase in students per classroom negatively impacted reading and mathematical abilities, lowering the overall academic performance. In addition, Hyytinen et al. (2014) explored the relationship between students' critical thinking and epistemological beliefs in problem-solving contexts. The study applied correlation analysis and found that students with deeper processing approaches to learning demonstrated stronger essential abilities of thinking, which in turn influenced their academic performance. In another study, Serebwa et al. (2017) investigated performance target setting and service delivery at Kirinyaga University, Kenya, using descriptive statistics and multiple linear regression analysis. Their study revealed that performance contracting parameters significantly affected service delivery and, by extension, academic effectiveness (Serebwa et al., 2017). These conventional approaches, however, often fail to capture the multidimensional relationships among academic variables, resulting in an incomplete understanding of students' success determinants (Jolliffe & Cadima, 2016). Given the complexity of students' performance data, more sophisticated methods, such as Principal Component Analysis (PCA), have been introduced to extract meaningful insights.

PCA was first introduced by Karl Pearson (1901) as a statistical technique to transform correlated variables into a set of linearly uncorrelated components, primarily to reduce dimensionality while preserving essential information. Later, Harold Hotelling (1933) extended PCA to multivariate analysis, making it widely applicable across various domains, including education. Over the past two decades, PCA has increasingly been used in student performance analysis to extract meaningful patterns from complex educational datasets. This allows researchers to identify key factors influencing academic success. By reducing the dimensionality of large student performance datasets, PCA enables more effective analysis of factors, such as socioeconomic background, learning behaviours and institutional resources, which traditional methods like regression analysis and correlation studies might overlook (Pearson, 1901; Hotelling, 1933). Studies leveraging PCA in education have demonstrated its ability to enhance the predictive modelling of students' outcomes and optimise teaching strategies based on data-driven insights. This shift towards more sophisticated analytical methods underscores the growing need to harness advanced statistical tools to improve educational assessment and intervention strategies.

Previous studies used descriptive and inferential statistical methods, including mean, standard deviation, t-tests and ANOVA to analyse students' performance trends (Cohen, 2013). For example, t-tests have been used to compare students' performance between two groups, such as gender-based differences or different teaching methods (Field, 2024). ANOVA has been widely applied to assess performance differences across multiple categories, such as the impact of different instructional approaches or institutional factors on academic success (Heiman, 1992). Descriptive statistics, including mean and standard deviation, provide a foundational understanding of students' performance trends, helping educators and policymakers to make data-driven decisions to enhance learning outcomes (Tabachnick et al., 2019). While these methods provide essential insights, they do not effectively account for the interdependencies between multiple academic subjects. Pearson correlation analysis, commonly used to explore relationships between course performance variables, is particularly limited when dealing with large datasets containing numerous variables (James et al., 2013). Regression models and machine learning techniques, such as decision trees and support vector machines have also been used to predict academic performance (Borges et al., 2018). While these methods provide prediction accuracy, they often lack interpretability and require extensive feature engineering, which can undesirably introduce bias (Alam & Khatun, 2021).

PCA has emerged as a robust tool for addressing these limitations by reducing dimensionality and identifying key academic success factors (Tabachnick & Fidell, 2019). By transforming highly correlated variables into independent principal components, PCA allows educators and researchers to analyse academic performance more effectively (Jolliffe & Cadima, 2016).

Methodology

This study examined factors influencing students' performance in the Bachelor of Electrical and Electronics Engineering (EEE) program, using a retrospective correlational research design and Principal Component Analysis (PCA). The study identified key performance determinants, course interdependencies and curricular implications through descriptive statistics, Pearson correlation analysis and PCA by analysing students' records.

Research Design

This study employed a retrospective correlational research design, utilising PCA to assess the key factors influencing students' performance in the Bachelor of Electrical and Electronics Engineering (EEE) program at Mbeya University of Science and Technology. A correlational research design examines relationships between variables without manipulating them, making it appropriate for this study as it seeks to identify underlying academic performance patterns based on existing students' records (Tabachnick & Fidell, 2019). The retrospective nature of the design enables the analysis of historical data to uncover trends and associations. At the same time, PCA was selected for its effectiveness in dimensionality reduction, allowing for the identification of essential factors that significantly impact academic success (Jolliffe & Cadima, 2016).

Population and Sampling

The population comprised 169 final-year students in the Bachelor of Electrical and Electronics Engineering (EEE) programme at Mbeya University of Science and Technology, who had completed all 16 core courses. A purposive sampling technique was employed to ensure that only students with complete academic records were included in the study. Creswell and Creswell (2017) emphasised that purposive sampling is beneficial when researchers aim to select participants with specific characteristics relevant to the study's objectives, ensuring the collection of rich and meaningful data. They argued that this technique is commonly used in quantitative and qualitative research to enhance the validity of findings by focusing on cases that best represent the population under investigation. They recommended that researchers clearly define their inclusion criteria to minimise bias and ensure data reliability. In this study, purposive sampling was appropriate as it allowed for selecting students with comprehensive academic records, ensuring accurate Principal Component Analysis and meaningful insights into the key determinants of academic performance.

A sample of 169 students was deemed adequate and adhered to the widely accepted rule of 5 or more cases per variable (Hair et al., 2010). For this study, there were 16 courses, which are the variables, and for the application of PCA, it is adequate since the sample is more than 5 times the number of variables.

Instruments

The study used the Students' Information Management System (SIMS) to extract academic records, course grades and GPA scores. The author developed a structured data extraction sheet to ensure consistency and eliminate redundant data (Field, 2024).

Validity and Reliability

Content validity was established by confirming that the selected academic variables, the courses and the respective students' marks were directly relevant to students' performance analysis since they are the inputs to the statistical analysis (Bolarinwa, 2015). Construct validity was tested through factor loadings, and eigenvalues and cumulative variance were explained, ensuring that the extracted components meaningfully represented students' performance factors (Tabachnick & Fidell, 2019). Tabachnick and Fidell (2019) discussed using factor loadings, eigenvalues and cumulative variance explained as key indicators for assessing construct validity in exploratory factor analysis (EFA). They emphasised that factor loadings should typically exceed 0.40 to indicate meaningful relationships between variables and factors. Eigenvalues are more significant than one and are commonly used to determine the number of factors to retain based on Kaiser's criterion. The cumulative variance explained should ideally exceed 50-60% to ensure that the extracted factors adequately represent the underlying construct.

Internal consistency was assessed using the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy to evaluate the reliability of the dataset. A KMO value above 0.70 was considered acceptable, indicating that the sample was adequate for Factor Analysis (Kaiser, 1974). According to Kaiser (1974), KMO values between 0.70 and 0.79 are considered "middling," between 0.80 and 0.89 are considered "meritorious" and values above 0.90 are "marvellous," demonstrating strong factorability.

Statistical Treatment of Data

The collected data underwent several statistical procedures, ensuring a robust analysis: The Jamovi software and the MS Excel spreadsheet were used to process and analyse data, as well as statistical analysis, visualisation and PCA implementation. Descriptive statistics were calculated to summarise the data distribution and assess normality, including mean, standard deviation, skewness and kurtosis (Field, 2024). According to Field (2024), skewness values between -1 and 1 indicate approximately normal distribution, while kurtosis values should fall within the range of -2 to 2 to avoid significant deviations from normality.

Pearson Correlation Analysis examined relationships between course performances, determining the strength and direction of linear associations (Cohen, 2013). Cohen (2013) suggests that correlation coefficients of 0.10–0.29 represent a small effect, 0.30–0.49 a moderate effect and 0.50 or above a strong effect, guiding the interpretation of course relationships. Principal Component Analysis (PCA) was conducted following the Kaiser criterion, where components with eigenvalues greater than 1 were retained to ensure meaningful data reduction (Jolliffe & Cadima, 2016). Jolliffe and Cadima (2016) emphasised that retaining components with eigenvalues above 1 preserves significant variance while eliminating redundant factors. Additionally, a scree plot was used to determine the number of meaningful components visually.

A varimax rotation was applied to improve interpretability, maximising variance among factor loadings and aiding in the distinct separation of components (Hair et al., 2010). Hair et al. (2010) highlight that varimax rotation enhances factor clarity by reducing complex cross-loadings, making the extracted components more interpretable. Cluster analysis was performed using Principal Component Scores to classify students based on performance patterns (Parmar & Bhavsar, 2020). According to Parmar and Bhavsar (2020), clustering based on principal components reduces dimensionality while preserving essential performance-based distinctions among students. The Shapiro-Wilk Test was applied to assess the normality of data distribution, ensuring that PCA assumptions were met (Ghasemi & Zahediasl, 2012; Shapiro et al., 1965). Ghasemi and Zahediasl (2012) report that Shapiro-Wilk is particularly effective for small to medium-sized samples, providing a reliable test for normality.

All statistical analyses were conducted using Jamovi Software (James et al., 2013), an open-source statistical platform that integrates PCA, correlation analysis and clustering methodologies. James et al. (2013) described Jamovi as a user-friendly alternative to traditional statistical software, offering robust analytical tools for educational research.

Ethical Considerations

The study adhered to ethical principles to ensure the confidentiality and protection of student records. Research clearance was obtained from Mbeya

University of Science and Technology, ensuring the study adhered to institutional ethical guidelines. In addition to university approval, authorisation from relevant government authorities was sought to comply with national research regulations and ensure lawful data collection. To protect participants' rights, anonymity and confidentiality were strictly maintained. No personally identifiable information was collected or disclosed and all data was anonymised to prevent the identification of individual respondents. Confidentiality was further upheld by securely storing research data and limiting access to authorised personnel only.

Results and Discussion

This section presents the findings and discussion. The results are systematically analysed and interpreted to reflect the study's objectives, ensuring clarity and coherence. Understanding students' performance across engineering courses requires an in-depth analysis of descriptive statistical measures, including central tendency, variability, skewness and normality. These statistical indicators provide critical insights into academic consistency, performance disparities and areas requiring intervention.

Evaluating Sampling Adequacy: Kaiser-Meyer-Olkin (KMO) Measure and Its Implications for Factor Structure

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was used to assess the suitability of the dataset for Factor Analysis. The overall KMO value obtained was 0.80, which falls within the "meritorious" range (Kaiser, 1974), indicating that the data is adequately suited for Factor Analysis (Table 1). The individual Measure of Sampling Adequacy (MSA) values for each variable provide insights into their factorability.

Most variables demonstrated acceptable KMO values above 0.70, suggesting they contribute effectively to the factor structure. Notably, EE 8405 (0.94), EE 8406 (0.94), EE 8410 (0.92), EE 8411 (0.91) and EE 8415 (0.95) showed "marvellous" sampling adequacy, further reinforcing the dataset's robustness. However, EE 8401 (0.28), EE 8408 (0.43) and EE 8416 (0.49) exhibited low KMO values, indicating weak factorability and potential issues with multicollinearity or insufficient correlation with other variables. GPA (0.68) also falls slightly below the recommended 0.70 threshold, suggesting borderline adequacy. These variables may require further examination, such as removal or transformation, to enhance the overall factor structure.

Table 1: KMO Measure of Sampling Adequacy

Description	MSA	Description	MSA	Description	MSA	Description	MSA
Overall	0.8						
EE 8413	0.75	EE 8404	0.8	EE 8408	0.43	EE 8412	0.79
GPA	0.68	EE 8405	0.94	EE 8409	0.88	EE 8414	0.89
EE 8401	0.28	EE 8406	0.94	EE 8410	0.92	EE 8415	0.95
EE 8402	0.83	EE 8407	0.9	EE 8411	0.91	EE 8416	0.49
EE 8403	0.74						

Student Performance across Subjects

By examining the descriptive statistics of students' performance across 16 engineering courses (Table 2), it is possible to identify trends that inform curriculum development, teaching strategies and academic support mechanisms (Field, 2024; Tabachnick et al., 2019).

The measures of central tendency, specifically mean and median, offer a foundational understanding of students' performance across courses. The findings indicate that EE 8401 (Industrial Practical Training 3) recorded the highest mean score (79.98), suggesting a course structure that supports students' success. In contrast, EE 8402 (Phase AC Synchronous Machines) had the lowest mean score (48.11), highlighting significant performance challenges. According to Biggs et al., 2022), mean scores often reflect the effectiveness of instructional methodologies, assessment designs and students' engagement levels. The discrepancy between mean and median values in some courses suggests the presence of outliers or non-uniform performance distributions, which may require further pedagogical adjustments (Heiman, 1992).

Students' performance variability is assessed using Standard Deviation (SD), which measures score dispersion. Courses such as EE 8402 (SD = 10.52) and EE 8409 (SD = 11.04) exhibit high variability, suggesting diverse students' competencies or inconsistencies in teaching approaches. High variability in students' performance, as indicated by standard deviations of 10.52 and 11.04, suggests significant differences in achievement levels, meaning that some students excel while others struggle. Whether high variability is considered good or bad depends on contexts. On the one hand, it highlights diverse learning needs, necessitating tailored instructional strategies, such as differentiated teaching and adaptive learning (Biggs et al., 2022). On the other hand, excessive variability might indicate inconsistencies in instruction, assessment difficulty or gaps in prior knowledge, potentially disadvantaging some students (Cohen, 2013; Allen & Yen, 2001). Conversely, EE 8405, with a low standard deviation (SD = 3.94), reflects uniform performance, which could indicate effective instructional delivery and a well-structured assessment framework (Hattie, 2008). However, too little variability may also suggest a lack of differentiation, where assessments fail to capture the full range of student abilities.

Table 2: Descriptive statistics on the students' performance

Description	EE 8401	EE 8402	EE 8403	EE 8404	EE 8405	EE 8406	EE 8407	EE 8408	EE 8409	EE 8410	EE 8411	EE 8412	EE 8413	EE 8414	EE 8415	EE 8416	GPA
Mean	79.98	48.11	53.11	63.54	62.27	56.24	61.31	65.08	66.38	59.07	55.12	59.84	58.15	69.74	70.07	57.62	3.61
Median	80	46	52	64	63	56	61	65	67	59	55	61	60	70	71	58	3.6
Standard deviation	5.23	10.52	9.26	6.64	3.94	10.43	8.83	7.51	11.04	8.02	10.74	9.13	12.85	9.36	7.27	9.01	0.51
Minimum	54	25	31	45	52	32	33	40	30	41	30	30	25	44	52	27	2.5
Maximum	92	82	76	78	77	82	83	85	89	81	88	80	84	95	84	81	4.8
Skewness	-0.68	0.56	0.14	-0.13	0.18	0.09	-0.11	-0.15	-0.6	0.39	0.19	-0.42	-0.41	-0.17	-0.38	-0.28	-0.06
Std. error skewness	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19
Kurtosis	2.88	0.21	-0.1	-0.48	0.73	-0.4	0.02	0.23	0.46	-0.24	0.24	-0.11	-0.35	-0.07	-0.4	0.38	-0.32
Std. error kurtosis	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37
Shapiro-Wilk W	0.96	0.98	0.99	0.99	0.98	0.99	0.99	1	0.97	0.98	0.99	0.98	0.98	0.99	0.98	0.99	0.99
Shapiro-Wilk p	<.001	0.005	0.486	0.315	0.049	0.689	0.245	0.85	0.003	0.034	0.545	0.033	0.009	0.665	0.009	0.462	0.13

Where: EE 8401 = Industrial Practical Training 3, EE 8402 3 = Phase AC Synchronous Machines, EE 8403 = Digital Control System Engineering, EE 8404 = Programmable Logic Controller, EE 8405 = Electrical Power Systems Dynamics, EE 8406 = Power Electronics Converters, EE 8407 = Engineering Operation Management, EE 8408 = Project I, EE 8409 = Law for Engineers, EE 8410 = Electrical Drives, EE 8411 = Digital System Engineering, EE 8412 = High Voltage Engineering, EE 8413 = Engineering Economic, EE 8414 = Switchgear and Protection Engineering, EE 8415 = Renewable Energy Technologies and EE 8416 = Project II.

A standard deviation interpretation scale should be applied to classify variability levels objectively. One common approach categorises SD values as: Low variability: SD < 5 (indicating uniform performance). Moderate variability: SD between 5 and 10 (indicating some differences but manageable diversity). High variability: SD > 10 (indicating significant disparities in performance)

Using this scale, 10.52 and 11.04 fall within the high variability range, requiring instructional adjustments. Meanwhile, 3.94 represents low variability, suggesting consistency in student performance. Thus, high variability is neither inherently good nor bad—it signals the need for pedagogical reflection and targeted interventions to support all learners effectively.

Skewness provides further insights into the distribution of students' scores. A negative skew in EE 8401 (-0.68) suggests that most students performed well, with a few lower outliers. This pattern may indicate effective teaching strategies and comprehensive students' understanding of course materials. Conversely, EE 8402 (0.56) exhibits a positive skew, meaning more students scored lower, which may indicate course difficulty or inadequate students' preparation. According to Tabachnick et al., (2019), positively skewed distributions often signal the need for additional academic support, such as remedial programs, peer mentoring or adjustments to course content. Tinto (1993) emphasizes that courses with positive skewness may require prerequisite reinforcement and targeted academic interventions to help students grasp fundamental concepts.

The kurtosis values further explain the shape of the distribution and the presence of extreme scores. EE

8401 (Kurtosis = 2.88) demonstrates a highly peaked distribution, indicating that most students scored close to the mean, suggesting a structured grading system or standardized learning outcomes (Field, 2024). On the other hand, EE 8404 (-0.48) exhibits a flatter distribution, meaning that students' performance varies significantly. Research by DeCarlo (1997) indicates that flatter distributions can be linked to inconsistent assessment methods or a wide range of students' competencies, necessitating instructional modifications to ensure a more balanced learning experience.

The Shapiro-Wilk test is a widely used statistical test to assess whether a dataset follows a normal distribution. The interpretation of the test results typically follows this scale: $p > 0.05$: Data does not significantly deviate from normality (assumption of normality is met); $p < 0.05$: Data significantly deviates from normality (non-normal distribution).

In this study, several courses had p-values below 0.05, indicating a significant departure from normality. However, only EE 8401 ($p < .001$) and EE 8409 ($p = .003$) were highlighted, which raises the question of why other courses with similar p-values were excluded from the discussion. A more comprehensive analysis should either include all courses with $p < 0.05$ or clarify the criteria for selecting only these two courses. For example, this should be explicitly stated if EE 8401 and EE 8409 exhibited the most extreme deviations from normality (e.g., lowest p-values or highest skewness/kurtosis). The deviation from normality suggests that students' performance is unevenly distributed, which may be attributed to factors, such as assessment methods, grading scales, instructional quality or differences in students' preparedness (Razali & Wah, 2011). In practical terms, this non-normality might indicate the presence of outliers, skewed distributions or multiple student subgroups' performing at different levels. Depending on the severity of the deviation, appropriate data transformation (e.g., logarithmic or square root transformation) or non-parametric statistical techniques could be considered for further analysis. To strengthen the discussion, including a table of Shapiro-Wilk test results for all courses or a histogram of selected distributions would provide more precise insights into the extent of non-normality across different courses.

Meanwhile, EE 8403 ($p = .486$) and EE 8416 ($p=0.462$) approximate normality, suggesting that the course assessments align well with expected students' performance trends (Cohen, 2013). Non-normality in students' performance can affect the reliability of inferential statistical analyses, and in such cases, non-parametric statistical methods may be more appropriate for further evaluation (Tabachnick et al., 2019).

These statistical findings have significant implications for targeted academic interventions. Courses with low mean scores and high variability, such as EE 8402 and EE 8409, require curricular adjustments, additional instructional support and structured assessment modifications to enhance students' comprehension (Tinto, 1912; Biggs et al., 2022). Positively skewed courses, where most students scored lower while a few excelled, indicate performance gaps that may stem from learning difficulties, instructional challenges or assessment complexity. Early diagnostic assessments help identify struggling students, enabling targeted interventions before gaps widen. They allow educators to differentiate instruction, tailor teaching strategies and provide immediate feedback to correct misconceptions early. Individualised learning plans—including remedial sessions, tutoring or adaptive learning tools—offer personalised support to enhance student performance in such courses (Hattie, 2008).

Patterns and Implications of Pearson Correlation Coefficients among Course Performances: Insights into Academic Factors Influencing Student Success

Analysing the relationships among students' performance across engineering courses using Pearson correlation coefficients provides insights into common academic competencies, curricular alignment, and areas requiring targeted interventions. In this analysis, students' scores in various engineering courses serve as independent and dependent variables, as the goal is to assess how performance in one course correlates with performance in another. The correlation matrix in Table 4 demonstrates the associations of variables. Correlation analyses offer evidence-based pathways for refining curriculum design, instructional methodologies and students' support mechanisms (Cohen, 2013; Tabachnick et al., 2019).

The Pearson correlation coefficients (r) measure the strength and direction of the relationship between two variables (Table 3). The values range from -1 to +1 (Cohen, 2013; where: +1 indicates a perfect positive correlation (both variables increase together); -1 indicates a perfect negative correlation (one variable increases while the other decreases); 0 suggests no correlation (no relationship between variables).

Table 3: Description of correlation values

Pearson r Value	Strength of Correlation	Interpretation
± 0.00 – 0.10	Very Weak	Negligible or no meaningful relationship
± 0.11 – 0.30	Weak	Low correlation, minimal shared variance
± 0.31 – 0.50	Moderate	Noticeable association, some shared variance
± 0.51 – 0.70	Strong	Substantial relationship, significant shared variance
± 0.71 – 0.90	Very Strong	High correlation, large shared variance
± 0.91 – 1.00	Near Perfect	Extremely high association variables are highly related

Table 4: Pearson Correlation Coefficients

Courses	EE 8413	EE 8401	EE 8402	EE 8403	EE 8404	EE 8405	EE 8406	EE 8407	EE 8408	EE 8409	EE 8410	EE 8411	EE 8412	EE 8414	EE 8415	EE 8416
EE 8413	1															
EE 8401	-0.11	1														
EE 8402	0.54	-0.09	1													
EE 8403	0.44	-0.08	0.56	1												
EE 8404	0.34	-0.01	0.38	0.36	1											
EE 8405	0.38	0.06	0.37	0.33	0.28	1										
EE 8406	0.54	-0.05	0.59	0.48	0.46	0.37	1									
EE 8407	0.47	-0.04	0.53	0.54	0.36	0.33	0.49	1								
EE 8408	0.15	0.05	0.26	0.27	0.23	0.18	0.21	0.2	1							
EE 8409	0.47	-0.04	0.54	0.46	0.35	0.3	0.47	0.54	0.11	1						
EE 8410	0.55	-0.14	0.52	0.48	0.41	0.28	0.56	0.43	0.25	0.44	1					
EE 8411	0.6	-0.03	0.46	0.51	0.4	0.29	0.51	0.45	0.23	0.5	0.52	1				
EE 8412	0.6	-0.02	0.53	0.39	0.35	0.27	0.59	0.42	0.2	0.43	0.5	0.56	1			
EE 8414	0.49	-0.08	0.58	0.49	0.36	0.3	0.52	0.46	0.16	0.54	0.48	0.48	0.56	1		
EE 8415	0.54	0.08	0.49	0.44	0.24	0.24	0.47	0.41	0.15	0.46	0.42	0.45	0.44	0.53	1	
EE 8416	0.23	-0.01	0.22	0.34	0.32	0.18	0.26	0.22	0.32	0.18	0.21	0.33	0.29	0.31	0.17	1

Strong and positive correlations, such as those between EE 8413 (Engineering Economics) and EE 8411 (Digital System Engineering) (r = 0.60) and between EE 8412 (High Voltage Engineering) and EE 8411 (r = 0.56), suggest a positive interrelationship between these courses. The associations indicate that the competencies developed in one course are relevant in enhancing the performance in the other. This indicates the coherent curriculum structures at the university are under investigation (Heiman, 1992). This aligns with literature, asserting that courses sharing theoretical underpinnings or similar problem-solving approaches exhibit strong correlations due to the reinforcement of shared cognitive skills (Biggs & Tang, 2022).

Moderate correlations observed between courses, such as EE 8410 (Electrical Drives) and EE 8411 (r = 0.52), as well as EE 8414 (Switchgear and Protection Engineering) and EE 8406 (Power Electronics Converters) (r = 0.52), indicate partial but not complete transferability of skills. Moderate interdependence suggests that students benefit from instructional strategies explicitly bridging content between the courses. These findings support educational theories that advocate integrative teaching approaches, reinforcing interconnected knowledge across related subjects (Hattie, 2008; Ramsden, 2003). For instance, structured reinforcement activities, such as integrated problem-solving exercises, could enhance skill transfer and reinforce underlying concepts, improving academic outcomes (Biggs aet al., 2022).

Weak correlations were also identified in pairs like EE 8416 (Project II) with EE 8412 (r = 0.29) and EE 8416 (Project II) with EE 8415 (Renewable Energy Technologies) (r = 0.17). Weak associations indicate low transferability of skills, knowledge and attitudes. This suggests that these courses may be pedagogically or cognitively distinct. The weak correlations may result from differences in course objectives, instructional methods or assessment procedures (Field, 2024; Ramsden, 2003). For example, project-based courses, such as EE 8416 emphasise practical application and higher-order cognitive skills, which differ from theoretical or concept-focused courses (Hattie, 2008). These findings underscore the importance of tailoring teaching methods and assessments to meet each course's specific cognitive demands and learning outcomes.

Negative correlations were found between EE 8413 (Engineering Economics) and EE 8401 (Industrial Practical Training 3) (r = -0.11), as well as between EE 8410 (Electrical Drives) and EE 8401 (r = -0.14). Negative correlations indicate an inverse relationship between course performances, meaning that as a student's score in one-course increases, their score in another course tends to decrease. This suggests that strong performance in specific subjects may be associated with weaker performance in others. This observation aligns with theories suggesting that variations in students' learning styles, cognitive strengths and course-specific assessment strategies contribute to inverse relationships in performance. Negative correlations indicate that as performance in one course improves, performance in another tends to decline. This could be due to differences in learning preferences, where students who excel in practical, hands-on courses may struggle with theoretical, abstract subjects or vice versa (Tabachnick et al.,, 2019). The correlation matrix also reveals distinct clusters of strong correlations among EE 8412, EE 8413 and EE 8411, indicating common foundational skills across engineering fundamentals, digital systems, and economic principles. These clusters suggest that integrated instructional approaches, such as interdisciplinary projects or cross-course assignments, could reinforce shared competencies and enhance overall learning outcomes (Biggs et al., 2022). Conversely, courses with weak or negligible correlations indicate areas where independent instructional strategies, specialised support, and tailored assessments may be necessary to address specific course objectives effectively.

Scree Plot and Eigenvalues in Principal Component Selection: Implications for Dimensionality Reduction and Variance Interpretation in Engineering Course Performance

The scree plot (Figure 1) and initial eigenvalues table (Table 5) provide critical insights into selecting principal components for dimensionality reduction in engineering course performance data.

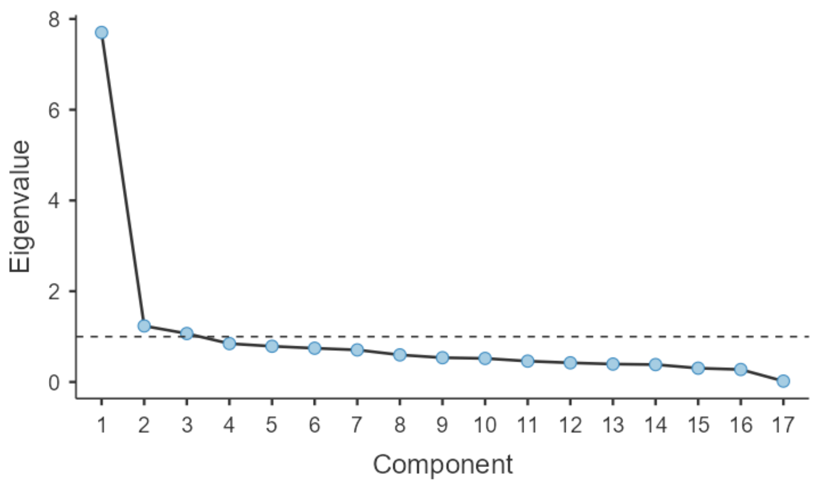

Figure 1: Scree Plot of the Courses

Table 5: Principal Components, eigenvalues, % of variance and the cumulative % of the variances

Principal Component	Eigenvalue	% of Variance	Cumulative %
1	7.7	45.31	45.31
2	1.23	7.26	52.57
3	1.07	6.28	58.85
4	0.85	4.98	63.83
5	0.79	4.63	68.46
6	0.74	4.38	72.84
7	0.71	4.16	76.99
8	0.6	3.51	80.51
9	0.54	3.15	83.66
10	0.52	3.05	86.71
11	0.46	2.71	89.42
12	0.42	2.49	91.9
13	0.39	2.32	94.22
14	0.38	2.26	96.48
15	0.3	1.79	98.27
16	0.28	1.62	99.89
17	0.02	0.11	100

The Principal Component Analysis (PCA) is a widely used technique that can transform a dataset into a set of uncorrelated principal components, prioritizing those that explain the most variance (Jolliffe & Cadima, 2016).

The scree plot (Figure 1) visualizes each component's eigenvalues, highlighting each one's contribution to the total variance. The steep decline in eigenvalues after the first principal component and the subsequent elbow at the second or third component indicate the optimal number of components to retain. This pattern suggests that additional components contribute minimal variance, aligning with standard PCA practices (Tabachnick & Fidell, 2019).

The initial eigenvalues table (Table 5) confirms this observation in Figure 1. The Principal Component 1 has an eigenvalue of 7.70, accounting for 45.31% of the total variance, making it the most dominant factor influencing course performance. Principal Component 2 explains an additional 7.26% (eigenvalue = 1.23), while Principal Component 3 contributes 6.28% (eigenvalue = 1.07). The cumulative variance explained by these three components reaches 58.85%, supporting their retention as key explanatory variables.

In contrast, components beyond the third component exhibit eigenvalues below 1 and contribute progressively smaller portions of variance, with the 17th component explaining only 0.11% (eigenvalue = 0.02). According to the Kaiser criterion, which suggests retaining components with eigenvalues greater than 1, only the first three components (Principal Component 1, Principal Component 2 and Principal Component 3) are meaningful for summarising the dataset (Kaiser, 1960).

These findings have significant implications for interpreting variance in engineering course performance. The dominance of PC1 suggests that a single underlying factor—potentially a core academic competency or a common instructional framework—accounts for a substantial portion of students' performance variations (Jolliffe, 2002). PC2 and PC3 are the next significant predictors, which explain an additional 13.54% of the variance, indicating that course performance is influenced by multiple dimensions, possibly reflecting disciplinary specializations, assessment methodologies, or learning strategies (Wold et al., 1987). The sharp decline in variance contribution beyond the third component supports dimensionality reduction, allowing for a more concise yet meaningful representation of students' performance patterns (Tabachnick et al., 2019).

It is fair to note that the dimensionality reduction suggested by PCA has direct implications for academic planning and curriculum optimization. Retaining only three principal components simplifies the complexity of analyzing students' performance while preserving the majority of meaningful variance. This approach enables educators and administrators to identify key performance drivers, optimize course design and reduce redundancy in academic evaluations.

Interpreting Principal Component Loadings and Uniqueness Values: Insights for Identifying Underlying Factors in Engineering Course Performance and Curriculum Design

In Table 6, the component loadings represent the correlation between each course and its Principal Component, helping to identify the primary academic constructs that drive students' performance. Jolliffe (2002) stated that loadings above 0.70 are considered strong indicators of a variable's contribution to a Principal Component, and variables with high loadings should be interpreted based on conceptual coherence rather than just numerical thresholds. He emphasised that Principal Components should be understood in terms of their underlying structure, meaning that courses grouped under a single component should reflect a meaningful academic construct rather than an arbitrary selection of high-loading variables. For Principal Component 1 (PC1), strong loadings were observed for EE 8413 (Engineering Economics) (0.77), EE 8402 (Phase AC Synchronous Machines) (0.76), and EE 8409 (Law for Engineers) (0.74). These courses were selected because they form a coherent academic category, likely emphasising theoretical and analytical competencies. The decision to exclude other courses with loadings above 0.70 is based on cross-loadings. If a course loads strongly on multiple components, it may not contribute uniquely to a single construct, making its interpretation less clear. Additionally, courses with strong loadings but distinct content areas may align better with another principal component, which will be analysed separately. The high correlations (i.e., strong positive associations between a course and a principal component, typically above 0.70) suggest that PC1 represents fundamental engineering competencies, analytical reasoning, and shared assessment methodologies. This aligns with Wold et al. (1987), who found that courses emphasising mathematical modelling and theoretical analysis exhibit strong interdependencies in PCA, relying on similar cognitive skills, problem-solving techniques, and grading criteria.

Principal Component 2 (PC2) shows strong loadings for EE 8408 (Project I) (0.76) and EE 8416 (Project II) (0.76), suggesting that this component reflects applied, project-based learning experiences. These courses require practical implementation of engineering concepts, teamwork and capstone project execution, distinguishing them from theoretical subjects. Principal Component 3 (PC3) is uniquely dominated by EE 8401 (Industrial Practical Training 3) (0.95), indicating that this course's performance is driven by specialized hands-on training distinct from broader engineering competencies (Kaiser, 1958).

Courses with moderate loadings across multiple components suggest interdisciplinary learning influences. For example, EE 8403 (Digital Control System Engineering) loads onto both PC1 (0.63) and PC2 (0.36), indicating a hybrid structure integrating both theoretical concepts and practical applications. This supports findings that engineering curricula benefit from aligning theoretical instruction with hands-on application to improve competency across domains (Biggs & Tang, 2022).

Influence of Moderate and Shared Loadings: Interdisciplinary Dependencies

Uniqueness Values and their Implications

The uniqueness values indicate the proportion of variance not captured by the extracted principal components. Low uniqueness values suggest that the identified components well represent a course, whereas high uniqueness values signal additional unexplained variance, suggesting that other instructional or assessment-related factors may influence students' performance (Tabachnick et al., 2019). For instance, EE 8401 (Industrial Practical Training 3) has a very low uniqueness value (0.09), meaning its variance is almost entirely explained by PC3, confirming its specialized nature. Conversely, EE 8405 (Electrical Power Systems Dynamics) (uniqueness = 0.69) and EE 8404 (Programmable Logic Controller) (uniqueness = 0.58) exhibit higher uniqueness values, indicating additional external influences on performance. These could stem from variations in instructional delivery, assessment structures, or student engagement levels (Field, 2024).

The Role of Varimax Rotation in Component Interpretation

In Table 6, the varimax rotation applied in this PCA analysis enhances component clarity by maximizing variance explained by each factor while reducing cross-loadings (Kaiser, 1958). This rotation enables distinct clustering of courses under relevant principal components, confirming that PC1 captures analytical and theoretical courses, emphasizing mathematical, economic and legal principles. Secondly, PC2 corresponds to project-based and applied learning courses, requiring practical engagement and implementation. Finally, PC3 uniquely represents specialized hands-on training, primarily associated with Industrial Practical Training 3. This approach validates PCA's effectiveness in distinguishing independent performance dimensions and supporting targeted curriculum adjustments (Jolliffe & Cadima, 2016).

Table 6: Principal Components Loadings

Courses	Component			Uniqueness
	1	2	3	
EE 8413	0.77			0.39
EE 8401			0.95	0.09
EE 8402	0.76			0.39
EE 8403	0.63	0.36		0.47
EE 8404	0.43	0.49		0.58
EE 8405	0.44			0.69
EE 8406	0.73			0.4
EE 8407	0.68			0.5
EE 8408		0.76		0.4
EE 8409	0.74			0.45
EE 8410	0.68			0.44
EE 8411	0.69			0.45
EE 8412	0.71			0.46
EE 8414	0.74			0.43
EE 8415	0.72			0.43
EE 8416		0.76		0.4

Implications for Curriculum Design and Academic Assessment

The findings present several strategic insights for improving engineering education and students' learning outcomes. They include optimizing course structuring based on key academic competencies, enhancing practical and capstone learning support, addressing unexplained variance in high-uniqueness courses and differentiating assessment strategies for course-specific learning goals.

Conclusions and Recommendations

Conclusions

The findings from this study provide valuable insights into students' performance trends, course interdependencies and academic structures in engineering education. The KMO measure confirmed data suitability for factor analysis while descriptive statistics highlighted performance disparities across courses, necessitating targeted instructional adjustments. Pearson correlation coefficients revealed strong, moderate and weak relationships between subjects, indicating areas for curriculum reinforcement or differentiation. Principal Component Analysis identified three key components: theoretical knowledge, applied learning and specialised practical training, emphasising the importance of a balanced curriculum. These insights collectively inform curriculum optimisation, assessment strategies and academic interventions, ensuring improved student learning outcomes and institutional effectiveness.

Recommendations

The study recommends that core theoretical courses should be aligned with standardised assessments that measure conceptual understanding and application to enhance curriculum effectiveness and reinforce analytical and problem-solving competencies. Project-based courses, such as Project I (EE 8408) and Project II (EE 8416), should integrate competency-based learning models and industry collaborations to strengthen applied engineering skills and bridge the gap between academia and professional practice. Specialised training courses, including Industrial Practical Training (EE 8401), require mentorship programs and skill-based evaluations to ensure that students develop the technical expertise and industry-relevant competencies necessary for career readiness.

Additionally, instructors should incorporate active learning techniques like problem-based learning, collaborative exercises and real-world case studies to improve students' engagement and conceptual understanding. The findings further suggest bridging theoretical knowledge with practical application through experiential learning opportunities, integrative projects and applied assessments. Future research should examine the long-term stability of Principal Components, identifying external factors influencing high-uniqueness courses and incorporating students' feedback for a more holistic understanding of academic performance trends. By adopting these data-driven recommendations, institutions can enhance curriculum effectiveness, close performance gaps and ensure that engineering education remains adaptive, equitable, and responsive to industry demands.

Acknowledgement

The Author acknowledges the support from the Mbeya University of Science and Technology, particularly the Department of Electrical and Power Engineering and the Examination Office.

References

Alam, M. J. and Khatun, F. (2021). Factors affecting academic performance of undergraduate students: evidence from the public university of Bangladesh. Journal of Education and Practice, 12(5), 50-57.

Alawin, A. A., Rahmeh, T. A., Jaber, J. O., Loubani, S., Dalu, S. A., Awad, W. and Dalabih, A. (2016). Renewable energy education in engineering schools in Jordan: Existing courses and level of awareness of senior students. Renewable and Sustainable Energy Reviews, 65, 308-318.

Allen, M. J. and Yen, W. M. (2001). Introduction to measurement theory. Waveland Press.

Biggs, J., Tang, C. and Kennedy, G. (2022). Teaching for quality learning at University 5e. McGraw-hill Education (UK).

Bolarinwa, O. A. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science research. Nigerian Postgraduate Medical Journal, 22(4), 195–201.

Borges, V. R. P., Esteves, S., de Nardi Araújo, P., de Oliveira, L. C., & Holanda, M. (2018, October). Using principal component analysis to support students' performance prediction and data analysis. In Brazilian symposium on computers in education (simpósio brasileiro de informática na educação-SBIE) 29(1).

Cohen, J. (2013). Statistical power analysis for the behavioural sciences. Routledge.

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

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. Psychological methods, 2(3), 292.

Felder, R. M. and Brent, R. (2024). Teaching and learning STEM: A practical guide. John Wiley & Sons.

Field, A. (2024). Discovering statistics using IBM SPSS statistics. Sage publications limited.

Ghasemi, A. and Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. International Journal of Endocrinology and Metabolism, 10(2), 486–489. https://doi.org/10.5812/ijem.3505.

Gibbs, G. and Simpson, C. (2005). Conditions under which assessment supports students' learning. Learning and Teaching in Higher Education, 1, 3–31.

Hair Jr, J. F., Black, W. C., Babin, B. J. and Anderson, R. E. (2010). Multivariate data analysis. In Multivariate data analysis (pp. 785-785).

Hattie, J. (2008). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. routledge.

Heiman, G. W. (1992). Basic statistics for the behavioural sciences. Boston, MA: Houghton Mifflin.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24(6), 417-441.

Hyytinen, H., Holma, K., Toom, A., Shavelson, R. J. and Lindblom-Ylänne, S. (2014). The complex relationship between students' critical thinking and epistemological beliefs in the context of problem-solving. Frontline Learning Research, 2(4), 1-24.

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, No. 1). New York: Springer.

Jolliffe, I. T. (2002). Principal component analysis for special types of data (pp. 338-372). Springer New York.

Jolliffe, I. T. and Cadima, J. (2016). Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202.

Kabakchieva, D. (2012). Student performance prediction by using data mining classification algorithms. International journal of computer science and management research, 1(4), 686-690.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. Psychometrika, 23(3), 187–200. https://doi.org/10.1007/BF02289233.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20(1), 141–151. https://doi.org/10.1177/001316446002000116.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36. https://doi.org/10.1007/BF02291575.

Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J. and Gonyea, R. M. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. The Journal of Higher Education, 79(5), 540-563.

Li, J. and Liang, W. (2024). Effectiveness of virtual laboratory in engineering education: A meta-analysis. *PloS one*, *19*(12), e0316269.

Martínez-Cervantes, T. J., Soto-Mendivil, E. A., Silva-Salazar, P. A. and Velasco-Arellanes, F. J. (2013). Effects of Basic Infrastructure on ENLACE Test of High School Technology Mexican. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, 11(4), 93-107.

Mills, J. E. and Treagust, D. F. (2003). Engineering education—Is problem-based or project-based learning the answer? Australasian Journal of Engineering Education, 3(2), 2–16.

Nicol, D. J. and Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education, 31(2), 199–218. https://doi.org/10.1080/03075070600572090.

Nurhidayat, E., Mujiyanto, J., Yuliasri, I., & Hartono, R. (2024). Technology integration and teachers' competency in the development of 21st-century learning in EFL classroom. Journal of Education and Learning (EduLearn), 18(2), 342-349.

Parmar, P. and Bhavsar, M. (2020). Achieving trust using rot in IAAS cloud. Procedia Computer Science, 167, 487-495.

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. Philosophical Magazine, 2(11), 559–572.

Ramsden, P. (2003). *Learning to Teach in Higher Education* (2nd ed.). Routledge.

Razali, N. M. and Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, and Anderson-Darling tests. Journal of Statistical Modelling and Analytics, 2(1), 21-33.

Serebwa, P., Njeru, W., & Maina, K. (2017). Performance Target Setting and Employee Service Delivery at Kirinyaga University, Kenya.

Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52(3/4), 591–611. https://doi.org/10.2307/2333709.

Tabachnick, B. G., Fidell, L. S. and Ullman, J. B. (2019). Using multivariate statistics (Vol. 6, pp. 497-516).

Tinto, V. (2012). Leaving college: Rethinking the causes and cures of student attrition. University of Chicago Press.

Wold, S., Esbensen, K. and Geladi, P. (1987). Principal component analysis. Chemometrics and intelligent laboratory systems, 2(1-3), 37-52.