Class Attendance and Academic Risk Condition in Higher Education: Evidence from University Students in Bangladesh

Vinicio Roblez-Torres

Universidad Técnica Particular de Loja, Ecuador

Abstract

This study aims to analyze the relationship between average class attendance and academic risk conditions among university students, through a statistical association approach in higher education. A quantitative, correlational, and cross-sectional design was conducted using a secondary open-access dataset comprising 328 observations. Class attendance was considered a behavioral academic indicator, while academic risk conditions were treated as the dependent variable. The analysis included descriptive statistics, contingency tables, and the application of Pearson's Chi-square test ($x^2 = 50.16$; $df = 35$; $p = 0.0466$), complemented by Cramer's V ($V = 0.391$). The results revealed a statistically significant association of moderate magnitude. From an applied perspective, the findings suggest that class attendance can be used as a useful indicator for early monitoring of academic risk in higher education, supporting evidence-based decision-making and institutional strategies focused on student performance.

Keywords: *Class attendance, Academic risk, Higher education, Data-driven monitoring, Student performance.*

1. Introduction

In recent years, higher education has experienced significant transformations in learning dynamics and in the factors that influence students' academic performance. These transformations have generated increasing interest in understanding not only academic outcomes but also the processes that explain them [1,2]. In this context, academic behavioural variables, such as class attendance, have gained relevance as potential determinants of student success. Attendance reflects the level of student commitment to their educational process and their interaction with the academic environment [3]. Likewise, it may influence content comprehension, active participation, and knowledge construction [4]. Therefore, the analysis of attendance is positioned as a key element for understanding academic performance in higher education.

One of the main challenges in higher education is the early identification of students in academic risk conditions. This issue is particularly relevant in contexts where dropout and low performance affect the quality and efficiency of educational systems [5]. Although multiple performance indicators exist, these often focus on final outcomes rather than on preceding behavioural factors. The lack of integration between academic and behavioural variables limits a comprehensive understanding of the phenomenon [6]. In this sense, class attendance emerges as a variable that could provide relevant information about academic risk. However, its joint analysis with indicators of academic vulnerability remains insufficient, which justifies the need for specific studies in this field [7].

Various studies by Ancheta and Kassarnig have shown that academic performance in higher education is influenced not only by cognitive factors but also by behavioural and contextual variables [1,8]. Among these, class attendance has been identified as a fundamental component of the learning process, as it facilitates interaction with the instructor and academic content [9]. Previous research has shown that students with lower attendance tend to present greater difficulties in their academic performance. This relationship suggests that attendance could be associated with academic risk conditions, such as sustained low performance [10]. Nevertheless, many of these studies have approached this relationship in a descriptive or limited manner. Therefore, it is necessary to apply statistical approaches that allow quantifying the magnitude of this association.

Despite the existing evidence, a gap remains in the literature regarding the formal statistical analysis between class attendance and academic risk condition. There is a limited number of studies that employ association tests to evaluate this relationship in university contexts. This limitation hinders the generation of robust empirical evidence to support institutional interventions. Furthermore, most studies have been conducted in specific contexts, restricting their generalizability. In this regard, analysing this relationship in different educational environments is essential to broaden the understanding of the phenomenon

[11,12]. Thus, the present study focuses on a scarcely explored international context, providing evidence from university students in Bangladesh.

The geographical context of the study corresponds to university students from Bangladesh, a developing country with characteristics in its educational system. This environment allows the analysis of the relationship between academic and behavioural variables in a context different from the Latin American one [13]. The dataset includes students admitted in 2021, enabling the evaluation of recent cohorts in higher education [14]. The sample consists of 328 students, providing an adequate size for statistical association analyses. Within this framework, two main variables are considered: average class attendance as the independent variable and academic risk condition as the dependent variable. This structure allows addressing the problem from a clear and coherent analytical perspective.

The objective of this study is to analyse the relationship between average class attendance and academic risk condition among university students admitted in 2021 in Bangladesh. To achieve this, a quantitative, associative, and cross-sectional approach was adopted. In the first stage, a descriptive analysis based on frequencies and contingency tables was conducted [15]. Subsequently, the Pearson Chi-square test was applied to evaluate the existence of association between the variables [16]. Likewise, Cramer's V was used to estimate the strength of this association; this methodological approach allowed quantifying the relationship between academic behaviour and student risk [17].

This study allows identifying the existence and magnitude of the relationship between class attendance and academic risk condition in higher education. Its results provide relevant empirical evidence to understand the role of attendance as a factor associated with student performance [11,18]. Likewise, it contributes to literature by incorporating formal statistical analysis in a scarcely explored international context. From an applied perspective, the findings can support the design of monitoring strategies and early warning systems in higher education institutions [19]. In addition, the study establishes a basis for future research focused on behavioural factors and their impact on academic performance. Overall, this research strengthens the analysis of academic variables from an integral and decision-oriented perspective [20].

2. Materials and Methods

2.1 Study Design

The present study aimed to analyse the relationship between average class attendance and academic risk condition among university students admitted in 2021 in Bangladesh. A quantitative, correlational approach with a non-experimental and cross-sectional design was adopted, allowing the examination of associations between categorical variables without direct manipulation. This design is appropriate for identifying relationships between variables in real-world contexts while preserving the integrity of observational data. The study was conducted within the higher education context of Bangladesh, using a previously collected dataset, which implies a retrospective analytical approach. This approach enabled the use of existing data to explore academic behavioural patterns in a specific student cohort.

2.2 Sample Description

The sample consisted of 328 university students corresponding to a cohort admitted in 2021. Inclusion criteria were based on the availability of complete information for the key study variables: average class attendance and academic risk condition. Consequently, records with missing values, inconsistencies, or ambiguous categories were excluded. Additionally, internal consistency checks were performed to identify potential duplication or recording errors. This data filtering process ensured internal validity and reliability of the analysis. The sample size was considered adequate for conducting statistical association tests involving categorical variables.

2.3 Data Collection Instrument

The data used in this study were obtained through a structured survey composed of closed-ended categorical items. The variable "average class attendance" was operationalized using ordinal categories representing different attendance levels (e.g., low, medium, and high attendance). The variable "academic risk condition" was measured through a dichotomous question identifying whether the student had experienced situations associated with low academic performance. The instrument was designed to capture relevant information about academic behaviors and performance-related conditions. Its structure enabled efficient coding and facilitated subsequent statistical processing.

2.4 Data Collection Procedure

The data used in this study were obtained from an open-access dataset available on the Mendeley Data platform (https://data.mendeley.com/datasets/dc3797vf3t/1), which contains academic and sociodemographic information from university students in a higher education context corresponding to Bangladesh. The survey had been previously administered in the university setting, and the data were stored in digital format for subsequent analysis. For the present study, the existing dataset was used without direct interaction with participants. The analysis was conducted retrospectively, always ensuring data anonymity and confidentiality. No additional data was collected, and no contact with students was established during the research process. This procedure allowed the use of secondary data under control conditions, minimizing potential biases associated with participant interaction. Ethical principles related to the use of pre-existing data were strictly observed, while the open-access nature of the dataset ensures transparency and reproducibility of the research.

2.5 Data Cleaning and Preparation

The data cleaning process was at a critical stage and was conducted in multiple phases. First, records with missing values in the variables of interest were identified and removed. Second, inconsistencies in responses were detected, including typographical errors, duplicated categories, and irregular categorical values. Third, category normalization was performed by standardizing equivalent labels. Additionally, variable levels were reorganized to ensure appropriate interpretation during statistical analysis. Finally, variables were converted into categorical (factor) format, and ordered levels were assigned to the attendance variable. This process ensured data quality, consistency, and reproducibility of the analysis.

2.6 Statistical Analysis in R

Statistical analysis was performed using R software (version 4.3.1) within the RStudio environment (version 2023.06.1). The following packages were used: readxl (1.4.3), dplyr (1.1.4), ggplot2 (3.4.4), writexl (1.4.2), DescTools (0.99.54), scales (1.3.0), stringr (1.5.1), and tibble (3.2.1). Initially, frequency tables were generated to describe variable distributions. Subsequently, contingency tables were constructed to examine the joint distribution of class attendance and academic risk condition. The Pearson Chi-square test was applied to assess the association between variables, after verifying expected frequency assumptions. Additionally, Cramer's V was calculated to determine the strength of the association. Finally, graphical representations, including bar charts, 100% stacked bar charts, heatmaps, and mosaic plots, were generated to support visual interpretation of the data.

2.7 Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki and applicable data protection regulations, ensuring participant anonymity and the exclusive use of aggregated data. Since the research was based on a previously anonymized dataset, ethical review and approval were waived, as the study was classified as "research without risk." Informed consent had been obtained from participants during the original data collection, ensuring voluntary and confidential participation. This approach ensured compliance with international ethical standards in academic research.

3. Results

3.1 Descriptive Analysis of Variables

To characterize the sample, a descriptive analysis of average class attendance and academic risk condition was conducted among university students admitted in 2021 in Bangladesh. This analysis allowed identifying the general distribution of the data and verifying the representativeness of the categories prior to the association analysis.

Table 1. Distribution of average class attendance (grouped)

Attendance level (%)	Frequency	Percentage (%)
Low (<70%)	20	6.10
Medium (70–89%)	83	25.30
High (≥90%)	225	68.60
Total	**328**	**100.00**

Table 1 presents the absolute and relative frequency distribution of the analyzed variables. Regarding average class attendance, a high concentration in upper values is observed, with the 100% attendance category accounting for 101 students (30.79%), representing the most frequent category. This is followed by 90% attendance with 42 students (12.8%), 80% with 27 students (8.23%), and 95% with 25 students (7.62%), indicating a general tendency toward high attendance levels within the sample.
In contrast, lower attendance categories show considerably smaller frequencies, many of them below 1%, such as values of 24%, 30%, 31%, and 34%, each with 1 case (0.3%), indicating minimal representation of students with low attendance levels. This distribution suggests a strong skew toward higher attendance values, which should be considered in subsequent analyses.

Table 2. Distribution of academic risk conditions

Academic risk condition	Frequency	Percentage (%)
Yes	84	25.61
No	244	74.39
Total	328	100.00

In Table 2, the distribution of academic risk conditions within the analyzed sample is presented. It is observed that 244 students (74.39%) were not at academic risk, while 84 students (25.61%) experienced academic risk conditions. This distribution indicates that although most students are not at academic risk, a considerable proportion has faced academic difficulties during their university trajectory. Overall, this evidence complements the general structure of the sample, characterized by a predominance of high attendance levels and a lower proportion of students in academic risk condition, supporting the feasibility of subsequent bivariate analysis.
To complement this description, a graphical representation of academic risk conditions is presented.

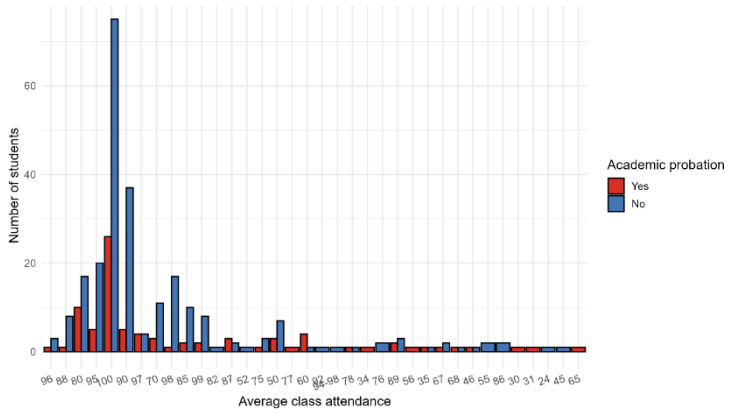

Fig. 1. Distribution of academic risk conditions.

Figure 1 presents the distribution of students with and without academic risk conditions across different levels of average class attendance. A clear concentration of cases is observed at higher attendance levels, particularly within the 95% to 100% range, where the highest absolute frequencies are recorded. Within these levels, students not in academic risk condition predominate, as evidenced by the greater height of the corresponding bars.
At intermediate attendance levels (approximately between 70% and 90%), a greater dispersion of cases is observed, including both students with and without academic risk condition. However, even within these ranges, the proportion of students not at academic risk remains higher.
At lower attendance levels (below 70%), the number of observations is considerably reduced, limiting the representativeness of these groups within the sample. Nevertheless, a relatively higher variability between categories can be observed in these ranges.

Overall, the figure reveals a skewed distribution toward higher attendance levels, as well as differences in the concentration of students according to their academic risk condition across attendance levels.

3.2 Joint Distribution Analysis

To examine the relationship between average class attendance and academic risk condition, a contingency table was constructed to display the joint distribution of both variables within the analysed sample.

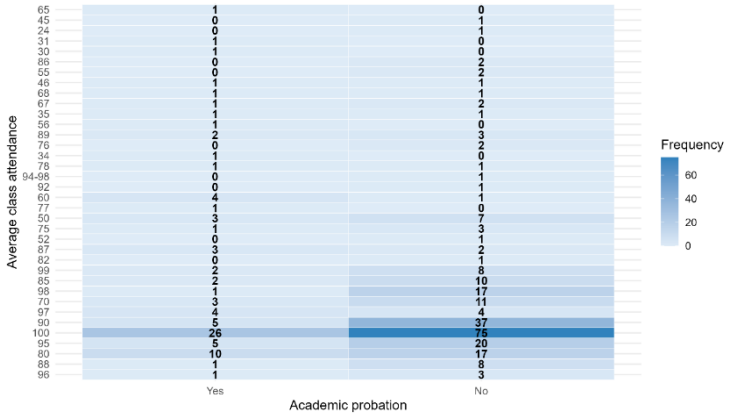

Fig. 2. Heatmap of the association between average class attendance and academic probation.

Figure 2 presents the distribution of academic risk condition across different levels of average class attendance, allowing the identification of specific patterns in the relationship between both variables. Overall, higher attendance levels concentrate a greater proportion of students without academic risk. The 100% attendance category records the highest frequency within the sample (n = 101), of which 75 students (74.26%) are not at academic risk, whereas 26 students (25.74%) are classified as academically at risk. Similarly, at the 90% attendance level (n = 42), 37 students (88.10%) are not at academic risk and only 5 students (11.90%) fall into the risk category, indicating a favourable distribution at high attendance levels.

Other high attendance categories show similar behaviour. At the 95% attendance level (n = 25), 80% of students are not at academic risk, compared to 20% who are classified as at risk. Likewise, at the 98% attendance category (n = 18), 94.44% of students are not at academic risk, representing one of the highest proportions observed in the figure. In the case of 96% attendance (n = 4), 75% correspond to students without academic risk and 25% to students at risk. These findings reflect a consistent tendency toward lower academic risk as attendance increases.

However, intermediate attendance categories exhibit more heterogeneous behaviour. At the 80% attendance level (n = 27), 17 students (62.96%) are not at academic risk, while 10 students (37.04%) are classified as at risk, representing one of the highest academic risk percentages among categories with substantial frequencies. Similarly, at the 87% attendance category (n = 5), 60% of students are at academic risk and only 40% are not. At the 70% attendance level (n = 14), 21.43% correspond to students at academic risk and 78.57% to students without risk, revealing important fluctuations within intermediate attendance ranges.

The figure also shows categories with perfectly balanced distributions. For example, attendance levels of 97% (n = 8), 78% (n = 2), 68% (n = 2), and 46% (n = 2) display an exact 50%–50% distribution between students with and without academic risk. These distributions indicate no dominance of either category, although they should be interpreted cautiously due to the limited sample size.

Finally, some attendance levels exhibit extreme distributions associated with minimal frequencies. Categories such as 77%, 34%, 30%, and 31% attendance show 100% of students classified as academically at risk, whereas categories including 82%, 52%, 92%, 94–98%, 76%, 55%, 86%, 24%, and 45% show 100% of students without academic risk. However, these results are conditioned by the presence of only one or two observations per category, limiting their statistical representativeness. Overall, the figure demonstrates a general tendency toward lower academic risk among students with high attendance levels, although important variations are observed across intermediate and lower attendance ranges.

3.3 Statistical Association Analysis

To formally assess the relationship between class attendance and academic risk condition, the Pearson Chi-square test was applied, complemented by Cramer's V as a measure of association strength.

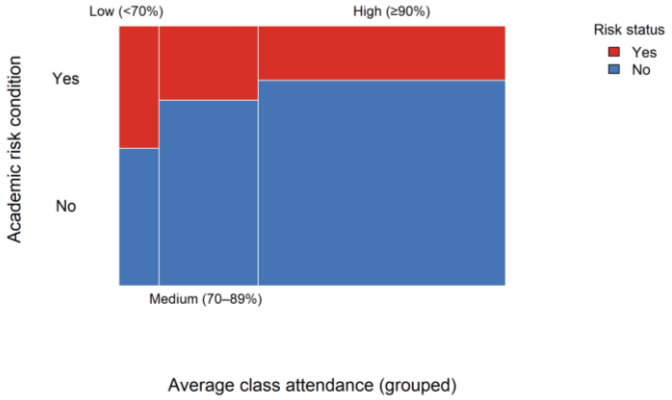

Fig. 3. Mosaic plot of grouped class attendance and academic risk.

Figure 3 presents a mosaic plot illustrating the relationship between grouped average class attendance and academic risk conditions among university students. The analysis was conducted using a total of 328 valid observations and revealed a statistically significant association between both variables, supported by the Pearson's Chi-square value ($\chi^2 = 50.16$; df = 35; p = 0.0466). This result indicates that the distribution of academic risk conditions varies across attendance levels and does not occur randomly. In addition, Cramer's V reached a value of 0.391, indicating a moderate association, suggesting that class attendance is related to academic risk, although it does not constitute a fully determining factor.

From a visual perspective, the high attendance category (≥90%) occupies the largest proportion of the plot, reflecting its predominance within the sample analyzed. This category accounts for approximately 68.60% of the total students (n = 225), indicating that most participants exhibit high attendance levels. Within this group, students without academic risk clearly predominate, represented by the blue area in the figure. This behavior is consistent with the previously observed descriptive results, where specific attendance categories such as 90%, 95%, 98%, and 100% showed high proportions of students without academic risk, reaching values of 88.10%, 80%, 94.44%, and 74.26%, respectively.

In contrast, the medium attendance category (70–89%) exhibits a more balanced distribution between students with and without academic risk. This category represents approximately 25.30% of the total sample (n = 83) and displays greater heterogeneity in the distribution of academic risk. Visually, the area corresponding to academic risk has a proportionally greater weight than in the high attendance category, indicating higher variability in students' academic behavior within this range. This pattern is consistent with specific attendance categories such as 80%, where 37.04% of students are classified as academically at risk, and 87%, where academic risk reaches 60% of cases.

On the other hand, the low attendance category (<70%) represents only 6.10% of the sample (n = 20); however, it shows a relatively higher proportion of students in academic risk conditions. In this category, the area corresponding to academic risk occupies a considerably larger visual proportion compared to the other groups, suggesting increased academic vulnerability associated with lower attendance levels. Although the size of this group is limited, the descriptive results revealed specific attendance categories with extreme distributions, where 100% of students were classified as academically at risk, particularly at attendance levels of 30%, 31%, 34%, and 65%.

Overall, Figure 3 reveals a progressive pattern in which higher attendance levels tend to be associated with a greater proportion of students without academic risk, whereas intermediate and lower attendance levels present less favorable and more variable distributions. However, the moderate magnitude of the association, reflected in Cramer's V, indicates that attendance only partially explains academic risk behavior, suggesting that additional academic, social, or contextual factors may influence this relationship.

4. Discussion

The results presented in Figure 3 (heatmap), reveal a statistically significant association between average class attendance and academic risk condition, supported by the Pearson's Chi-square value ($x^2 = 50.16$; df = 35; p = 0.0466). This indicates that the distribution of both variables is not independent, which is further reinforced by the Cramer's V value (V = 0.391), suggesting a moderate strength of association. From a visual perspective, the heatmap highlights differentiated frequency concentrations, while the mosaic plot shows a greater proportion of students without academic risk at higher attendance levels. This behavior is consistent with the findings of Ancheta, et al., as well as Kassarnig et al. who emphasize that attendance plays a significant role in academic performance by enhancing interaction with the learning environment. Therefore, the results confirm that attendance is associated not only with performance but also with academic risk conditions.

Based on the analysis of Figure 2 and its correspondence with Figure 3, structural patterns in the distribution of academic risk across attendance levels can be identified. Specifically, the high attendance category shows a clear concentration of students without academic risk, whereas the medium attendance category presents a more balanced distribution between conditions. This pattern reflects a non-linear relationship, consistent with the findings of Al Hazaa et al., and Karnik, et al. who argue that academic performance is influenced by multiple factors, with attendance being a relevant but not exclusive component. Similarly, studies by Ha et al. and Mia et al. indicate that attendance positively affects performance, although mediated by additional variables. In this regard, the moderate value of Cramer's V reinforces the interpretation that attendance contributes to academic risk but does not fully determine it.

Finally, although Figure 2 indicates that the low attendance category represents a smaller proportion of the sample, its combined analysis with Figure 3 reveals a higher relative concentration of students in academic risk within this group. This pattern is visually evident in the mosaic plot, where the segment representing academic risk occupies a larger proportion at lower attendance levels. This finding aligns with the work of Soares et al. and Nieuwoudt, who highlight that low attendance is associated with persistent academic difficulties. Moreover, recent studies by Rashid et al., Gray and Perkins, and Borges et al. propose the use of behavioral indicators such as attendance in early warning systems and academic monitoring frameworks. In this context, the findings not only confirm existing literature but also provide empirical evidence from a relatively underexplored international setting, strengthening the data-driven monitoring approach in higher education.

5. Conclusions

The present study aimed to analyse the relationship between average class attendance and academic risk condition among university students, demonstrating that a statistically significant association exists between both variables. The results obtained through Pearson's Chi-square test and Cramer's V indicate that class attendance constitutes a relevant factor in the analysis of academic performance, showing that higher levels of attendance are associated with a lower presence of academic risk. From an applied perspective, these findings highlight the importance of considering attendance as a key indicator within institutional student monitoring and tracking systems.

However, this study presents certain limitations that must be considered when interpreting the results. First, the cross-sectional design prevents the establishment of causal relationships between the analysed variables, restricting the findings to statistical association. Additionally, the use of a secondary dataset limits control over data quality and variable structure, including the presence of low-frequency categories, which may affect the robustness of the Chi-square analysis. Furthermore, complementary academic, socioeconomic, and behavioural variables that could influence academic risk were not included, suggesting the need for more comprehensive analytical approaches.

In this regard, future research should focus on developing multivariate models that incorporate additional variables such as study habits, technology use, prior academic performance, and contextual factors, to provide a more comprehensive explanation of academic risk. Likewise, longitudinal studies are recommended to analyse student performance over time and establish stronger causal relationships. Finally, the findings open opportunities for implementing early warning systems based on data analytics, contributing to the design of institutional strategies aimed at improving student retention and academic success in higher education.

Data Availability Statement

The data that support the findings of this study are openly available in the Mendeley Data repository at: https://data.mendeley.com/datasets/dc3797vf3t/1. This dataset corresponds to an open-access database containing information on university students. For the purposes of this study, specific variables were selected and processed, including average class attendance and academic risk condition. The data were used under open-access conditions, and no restrictions apply to their availability. Researchers may access the dataset directly through the provided link.

Acknowledgments

The author acknowledges the Mendeley Data repository for providing open access to the dataset used in this research, which enabled the development of the statistical analysis presented in this study.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this manuscript, artificial intelligence tools were used to support language editing and translation processes. Specifically, AI-assisted tools were employed to improve the clarity, coherence, and technical quality of the English version of the text. After using these tools, the author carefully reviewed and edited the content, ensuring accuracy and consistency. The author assumes full responsibility for the content of this publication.

REFERENCES

[1] Ancheta R.F., Daniel D., Ahmad R., “Effect of class attendance on academic performance”, European Journal of Education Studies, Vol. 8, 2021, doi:10.46827/ejes.v8i9.3887

[2] Ha W., Ma L., Cao Y., Feng Q., Bu S., “The effects of class attendance on academic performance: evidence from synchronous courses during COVID-19 at a Chinese research university”, International Journal of Educational Development, Vol. 104, 2024, doi:10.1016/j.ijedudev.2023.102952

[3] Nieuwoudt J.E., “Investigating synchronous and asynchronous class attendance as predictors of academic success in online education”, 2020

[4] Karnik A., Kishore P., Meraj M., “Examining the linkage between class attendance at university and academic performance in an international branch campus setting”, Research in Comparative and International Education, Vol. 15, 2020, pp. 371–390, doi:10.1177/1745499920958855

[5] Oyedotun S.A., Ejenarhome O.P., Oise G.P., “Learning analytics and predictive modeling: enhancing student success through data-driven insights”, Journal of Science Research and Reviews, Vol. 2, 2025, pp. 42–51, doi:10.70882/josrar.2025.v2i3.77

[6] Erdelyi V., Mizumoto T., Kitai Y., Ishimaru D., Adachi H., Higashino T., Ikeda M., “Detecting subtle signs of school attendance issues using smartphone-based sensing”, IEEE Access, Vol. 13, 2025, pp. 4652–4669, doi:10.1109/ACCESS.2024.3523108

[7] Rashid A., Yasmeen R., Khan R.A., “Early detection and intervention: a framework for preventing academic failure in medical students”, Pakistan Journal of Medical Sciences, Vol. 41, 2025, pp. 919–922, doi:10.12669/pjms.41.3.10883

[8] Kassarnig V., Bjerre-Nielsen A., Mones E., Lehmann S., Lassen D.D., “Class attendance, peer similarity, and academic performance in a large field study”, PLoS One, Vol. 12, 2017, doi:10.1371/journal.pone.0187078

[9] Al Hazaa K., Abdel-Salam A.S.G., Ismail R., Johnson C., Al-Tameemi R.A.N., Romanowski M.H., BenSaid A., Rhouma M.B.H., Elatawneh A., “The effects of attendance and high school GPA on student performance in first-year undergraduate courses”, Cogent Education, Vol. 8, 2021, doi:10.1080/2331186X.2021.1956857

[10] Soares D., Soares F.F.M., Baptista E., Carvalho N., Ximenes Z.A.C., Hastuti A.P., Hidayat S., “The effect of classroom attendance on academic achievement of management and leadership discipline of nursing students”, Journal of Quality in Health Care & Economics, Vol. 8, 2025, pp. 1–8, doi:10.23880/jqhe-16000438

[11] Hendrawan S.A., “Analysis of the digital system-based academic factors and the risk of student dropout using chi-square test for data-driven interventions”, Edu Cendikia: Jurnal Ilmiah Kependidikan, Vol. 5, 2025, pp. 602–616, doi:10.47709/educendikia.v5i02.6736

[12] Chan L.G., Ng Q.Y., “Using learning analytics in higher education: assessing students’ learning experience in an actuarial science course”, STEM Education, Vol. 4, 2024, pp. 151–164, doi:10.3934/steme.2024010

[13] Mia R., Selim S., Monir M., Hossain M.A., Ahmad B., Hasan M.R., “Effect of class attendance on academic performance in higher studies: a case study at Dhaka International University”, Khulna University Studies, 2025, doi:10.53808/kus.2025.22.02.1342-ss

[14] Suraiya S., Debnath C., Anam S., “International Journal of Engineering & Advanced Technology”, International Journal of Engineering & Advanced Technology, 2020, doi:10.35940/ijeat.C5334.029320
[15] Rost J.K., “Analyzing student success outcome variables in higher education utilizing the chi-square test of independence”, International Journal of Higher Education, Vol. 13, 2024, p. 100, doi:10.5430/ijhe.v13n2p100
[16] Chen D., Anderson C., “Categorical data analysis”, International Encyclopedia of Education, 2023, doi:10.1016/B978-0-12-818630-5.10070-3
[17] Nihan S.T., “Karl Pearson’s chi-square tests”, Educational Research and Reviews, Vol. 15, 2020, pp. 575–580, doi:10.5897/err2019.3817
[18] Gray C.C., Perkins D., “Utilizing early engagement and machine learning to predict student outcomes”, Computers & Education, Vol. 131, 2019, pp. 22–32, doi:10.1016/j.compedu.2018.12.006
[19] Borges G.A., Pedro C.F.S., Anjos J.C.S., Rodrigues A., Boavida F., Sá Silva J., “A platform for early class dropout prediction of university students”, IEEE Access, Vol. 13, 2025, pp. 109116–109133, doi:10.1109/ACCESS.2025.3581751
[20] Yang Y., “A design of academic diagnosis and early warning system based on big data mining technology”, in Proceedings of EDCS 2025, ACM, 2025, pp. 165–171