ISSN 2579-7298


# Interactive Dashboard Development for Student Performance Monitoring: Integrating Academic and Socio-Demographic Data

Fitrah Maharani Humaira, Wiratmoko Yuwono, Rengga Asmara, Rusminto Tjatur, Ronny Susetyoko, Robi'Atul Adawiyah

Politeknik Elektronika Negeri Surabaya

ARTICLE INFO

**Article history**
Received
Revised
Accepted

**Keywords**
Interactive Dashboard
Strategic Decision Making
Socio Demographic Data
Data Integration
Institutional Data

ABSTRACT

Strategic decision making in institutional settings is often constrained by the fragmentation and heterogeneity of data across multiple sources. This study addresses this critical gap by developing and validating an interactive web-based dashboard designed to consolidate and transform heterogeneous institutional data from seven distinct sources into actionable insights. A complex feature engineering pipeline was necessitated, involving comprehensive data integration and structural consistency checks. Techniques like Text Normalization and Feature Mapping were applied to clean over a lot of inconsistent entries, alongside Feature Binning and Extraction to generate analytically robust metrics. The system was implemented using Python for data processing and ReactJS for the dynamic interface, and its viability was validated via structured User Acceptance Testing (UAT). The subsequent descriptive analysis provided key insights into student demographics, geographical reach, and enrollment compliance across academic levels. Crucially, the comprehensive UAT resulted in an outstanding overall acceptance score of very worthy. However, feedback analysis indicated a dominant user focus on visual aspects, with noted complaints regarding the suboptimal color scheme and contrast impacting user experience. The findings confirm that complex feature engineering is a viable and effective strategy for transforming fragmented institutional data into an immediately deployable strategic resource. This system offers a validated blueprint for data consolidation in higher education. Future work is accordingly directed toward revising the color palette and contrast ratios to enhance visual clarity and user experience, alongside continuous optimization of data completeness to maintain the dashboard's utility.



## 1. INTRODUCTION

In the current competitive environment, higher education institutions are undergoing a significant digital transformation, driven by the growing demand for accountability and improved operational efficiency [1, 2]. To meet these challenges, the systematic collection and analysis of academic data are essential. This data is regarded as a vital resource for enhancing operational effectiveness, optimizing resource management, and improving the quality of education [2, 3, 4]. The shift toward data-driven management inevitably introduces Big Data characteristics within the academic environment, where universities face a rapid increase in data volume, velocity, and variety, encompassing structured student records, unstructured learning materials, and semi-structured interaction logs [5, 6]. The primary challenge remains that the sheer volume and mixed types of these data overwhelm traditional manual analysis, causing delays and resulting in data that is not functional for timely decision-making [5].

Focusing on a specific context, XYZ university reveals a clear challenge. While the institution successfully processed external data, important internal data is often overlooked, it is actually rich and valuable in information. The internal data is crucial for making the best decisions, like monitoring student engagement, tracking the effectiveness of administrative processes, and identifying student profiles for targeted marketing and promotion. However, it frequently suffers from delays in analysis and fragmented data practices, which ultimately slow down the administration's speed and efficiency [2,4].

Dashboards function as visual interfaces that convert complex data into understandable insights, supporting both everyday and strategic decision making [6, 7, 8]. The elements on the dashboard such as effectively showing key performance indicators (KPIs) and uncovering hidden trends, significantly improve the monitoring process and allow quick data driven decisions by relevant stakeholders [8]. Existing literature on dashboard implementation, particularly within the higher education sector, reveals distinct research foci and limitations.

Prior research, such as the work by Jayashanka et al., often focuses on Learning Analytics Dashboards (LADs) aimed at improving student motivation and grades by analyzing interaction data from Virtual Learning Environments (VLEs) [6]. Conversely, studies proposing comprehensive development frameworks, by Altarazi et al., center on database standardization and evaluating usage effectiveness in industrial or business contexts, focusing on financial data [7].

Despite these significant contributions, this comprehensive review points to a notable research gap concerning the operational context of higher education, the seamless integration and effective visualization of various types of internal academic data (mixed-type) originating from fragmented internal systems, the implementation of a dashboard framework that is empirically validated using quantitative measures by institutional management to assess the impact on decision-making efficiency and a focus on managing academic-demographic data [4].

Against the background of these identified gaps, this study presents contribution through the development of an integrated framework specifically tailored to the XYZ University. The proposed innovation focuses on harmonizing and visualizing complex and diverse internal academic data sources. Specifically, the contribution of this research lies in designing a comprehensive dashboard that consolidates fragmented academic and socio-demographic data into one interactive view and confirming the system's efficacy through rigorous empirical validation using quantitative metrics [7]

## 2. METHODS

### *2.1. Data Collection*

The dataset utilized in this study was collected using the Secondary Data Analysis (SDA) methodology, leveraging the comprehensive internal student database of XYZ University [9]. The SDA approach was chosen due to its efficiency in accessing a large volume of historical data and its suitability for addressing research questions focused on institutional trends and existing administrative records [18]. The target population of this study included all students enrolled at XYZ University spanning four academic years, specifically from 2021 until 2024. The dataset selected for analysis encompasses key variables related to the students' economic background (the primary variable of interest) and academic records, which are central to the research objectives, alongside structural identifiers such as the student ID [19]. To ensure the integrity of the analysis, strict criteria were applied, focusing only on records that were complete across all selected economic and demographic features [27]. This comprehensive collection ensures sufficient data volume and temporal depth for robust analysis.

### *2.2. Data Preprocessing*

The initial step in data preparation was data profiling, performed to inspect the dataset for quality and consistency. This process identified high structural inconsistencies within the qualitative variable representing the economic section, necessitating rigorous cleaning [10]. The profile confirmed the absence of missing values in the primary variables and that all observations were unique following an initial check for duplicates [20].

Despite the initial cleanliness regarding missingness, further refinement of the data structure was critical due to the high variability in the qualitative economic section variable. The handling process began with text normalization, including case folding and the removal of redundant white spaces, to unify typographical variations [11]. Subsequently, data standardization and feature mapping were performed based on a meticulously defined domain dictionary. This dictionary was developed based on institutional policy definitions, classifying and grouping the original qualitative entries into simplified, semantically consistent categories to ensure data quality and model interpretability [21].

Following this, feature engineering was applied to transform and enrich the dataset [13]. The continuing income variable was converted into discrete, manageable ranges through the feature binning (discretization) method [12]. This discretization was specifically applied to address potential non-linearity and reduce the impact of extreme outliers in the highly-skewed income data, thereby optimizing the variable for subsequent machine learning models [22]. Finally, the dataset was enriched by feature extraction, which involved parsing the structural format within the student ID to provide rich categorical features such as the department code for further segmented analysis [13].

### *2.3. Descriptive Analysis*

A comprehensive descriptive analysis was performed across all key variables [16]. This involved calculating measures of central tendency (mean, median, mode) and measures of dispersion (standard deviation, range) to summarize the main characteristics of the student population data. Furthermore, distributional analysis was conducted, utilizing statistical plots to ascertain the shape and spread of key variables [17]. This step was crucial in detecting any significant data skewness or outliers, which could potentially violate the assumptions of the chosen inferential models [23, 24].

### *2.4. System Design Architecture and Data Workflow*

The system uses a modular architecture designed for efficiently managing, processing, analyzing, and presenting data through an interactive dashboard [25]. The system was developed using a modern technology stack to ensure a robust, dynamic, and responsive user experience. Python served as the primary tool for all data-centric operations (cleaning, analysis, and modeling) [15]. The cleaned data was prepared in the lightweight JSON format, ready for efficient web deployment due to its simplicity and platform independence [26]. The front-end was built using the ReactJS library to create a dynamic and interactive User Interface (UI), chosen for its component-based efficiency and scalability [14]. The visual foundation utilized standard HTML and CSS, with TailwindCSS implemented as the utility-first CSS framework for efficient and rapid styling. Data visualization within the web page was powered by the integrated ECharts library, selected for its high performance in rendering interactive graphical elements from large datasets [9, 14, 15].

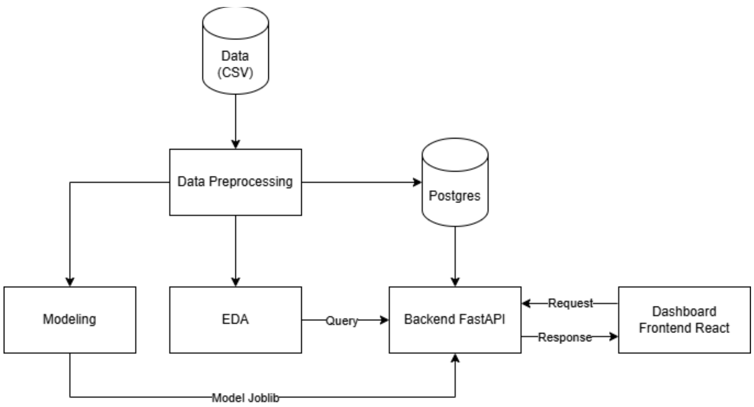


**Fig. 1.**Design System

### *2.5. User Acceptance Testing (UAT)*

The final validation involved User Acceptance Testing (UAT), defined as the testing process to confirm the software's acceptance by end-users. To ensure the robustness of the measurement, the instrument underwent rigorous validation using a pilot sample ($N = 30$) representative of the target user population but excluded from the final UAT respondent pool [28]. The questionnaire's validity was assessed using Pearson's Product-Moment Correlation to determine if each item significantly correlated with its respective construct dimension [16]. An item was deemed valid if its calculated correlation coefficient ($r_{count}$) was greater than the critical table value ($r_{table}$) at a 5% significance level [29]. The internal consistency and reliability of the instrument were assessed using Cronbach's Alpha ($\alpha$) [30]. A high Cronbach's Alpha value indicates that the items consistently measure the same underlying construct. Since this value significantly exceeded the generally accepted threshold of $\alpha > 0.60$ [31], the instrument was confirmed to possess high reliability for consistent measurement of user acceptance.

The validated questionnaire was then administered to the primary UAT respondents. A total of 82 participants from XYZ University were selected based on their diverse roles and experience as potential system users, comprising students, lecturers, and administrative staff. This diverse sampling approach, incorporating individuals directly involved in academic delivery and institutional processes, ensured that the validation included perspectives from users with explicit knowledge and experience in the system's operational context, thereby fully addressing the criteria for empirical validation by knowledgeable stakeholders. The UAT procedure consisted of three stages: planning, system execution, questionnaire completion, and final scoring interpretation. The acceptance score was quantified on Equation (1) and the category percentage interpretation on Table 1 [27].

$$Acceptance\ Score\ = \frac{Total\ Positive\ Score}{Maximum\ Possible\ Score} x\ 100\% \tag{1}$$

Table 1. Acceptance Category Percentage

| Percentage Score | Category |
|---|---|
| 81-100% | Very Worthy |
| 61-80% | Worthy |
| 41-60% | Quite Worthy |
| 21-40% | Unworthy |
| 0-20% | Very Unworthy |

## 3. RESULTS AND DISCUSSION

This section presents the analytical findings and results of the system implementation. The subsequent discussion will interpret these results to confirm the practical application and achievement of the research objectives.

### *3.1. Data Integrity and Structural Consistency*

The initial result confirms the preparation of unified analytical dataset from multiple sources. The research utilized four distinct secondary data sources, achievement, GPA, profile student, discipline, and regional geomatrics which were subjected to preprocessing. Initial data auditing confirmed the presence of no missing values or duplicate rows across the sources. This dataset required extensive feature engineering to standardize the disparate data types. The impact of this feature engineering process on data quality and structure is summarized in Table 2.

Table 2. Result of Feature Engineering

| Variable | Data Type | Initial Unique Values | Transformation Process | Final Data Type | Final Category |
|---|---|---|---|---|---|
| Parents Job | Text | 986 | Text Normalization & Feature Mapping | Text | 13 |
| Parents Income | Numeric | 871 | Feature Binning (Discretization) | Ordinal | 7 |

Table 2 confirms the necessity of the feature engineering methodology. The successful reduction of the Parents Job variable from 986 initial inconsistent entries to 13 standardized, uniform categories represents a critical finding from this method's application, decisively demonstrating that robust data standardization is an essential prerequisite for effectively utilizing institutional records in comparative analysis. This significant dimensionality reduction on categorical features, alongside the conversion applied to the continuous parents income variable via feature binning, collectively ensured the final dataset achieved structural uniformity and analytical readiness.

Crucially, this observed data fragmentation mandates a proactive revision of institutional data college policies, specifically emphasizing the establishment of rigorous, source level data standardization protocols. Such measures will be instrumental in streamlining the future ETL (Extract, Transform, Load) pipeline, substantially decreasing the operational overhead associated with data pre-processing, and guaranteeing consistent, high fidelity data quality necessary for rigorous academic and comparative research. The data preparation process utilized tables consolidated from four core data sources, with Table 3 detailing the initial variables and data types that underscored the intrinsic structural diversity of the source data prior to transformation.

Table 3. Dataset Variables and Type

| **Data Source** | **Variable** | **Data Type** | **Purpose** |
|---|---|---|---|
| Achievements | Study Program | String | Segment achievement based on the specific study program |
| | Department | String | Analyze the distribution and concentration of student achievements across different departments. |
| | Event | String | Categorize and analyze the types of academic and non-academic events recorded. |
| | Organizer | String | Analyze the organizer of the events. |
| | Event Scale | String | Quantify the level of achievement difficulty or prestige (Regional, National, International). |
| | Event Category | String | Group achievements for visual comparison |
| | Date Start | Date | Define the start duration of the events |
| | Date End | Date | Define the end duration of the events |
| | Awardees | String | Identify the specific student receiving the award |
| | Description | String | Supplemental qualitative data for context and detail. |
| GPA | Study Program | String | Segment GPA based on the specific study program |
| | Department | String | Analyze the distribution and concentration of GPA across different departments. |
| | Student Status | String | Filter or compare performance based on student activity status (Active, Suspended, Graduated). |
| | Academic Year | String | Track academic performance trends over academic year |
| | Semester | String | Track temporal changes and trends in GPA over short intervals. |
| | GPA | Numeric | The primary quantitative metric for measuring academic success and student evaluation. |
| Profile Students | Study Program | String | Segment student based on the specific study program |
| | Department | String | Analyze the distribution and concentration of students across different departments. |
| | Academic Year | String | Track academic performance trends over academic year |
| | City | String | Analyze geographical reach and perform city-level enrollment concentration analysis. |
| | Province | String | Analyze macro geographical reach and promotional effectiveness at the provincial level. |
| | Previous Education | String | Analyze recruitment strategy and concentration based on high school background. |
| | Admission Type | String | Validate compliance with regulatory quotas and understand the diversity of enrollment channels. |
| | Income Category | String | Analyze the socio-economic background of students. |

| Data Source | Variable | Data Type | Purpose |
|---|---|---|---|
| Discipline | Study Program | String | Segment attendance and disciplinary records by study program. |
| | Department | String | Analyze disciplinary trends and attendance rates by department |
| | Academic Year | String | Monitor disciplinary status and attendance trends over time. |
| | Semester | String | Track attendance patterns and disciplinary issues across specific periods. |
| | Attendance Rate | Numeric | Primary quantitative metric for measuring student engagement and compliance. |

### *3.2. Descriptive Analysis*

Descriptive analysis was performed across all key variables to summarize the characteristics of the student population. This involved quantifying measures of central tendency and dispersion on numerical attributes, and frequency distributions. This analysis is unique in its integration of heterogeneous data, encompassing records from postgraduate, applied bachelor, and distance learning students summarized in Table 4

Table 4. Student Distribution by Degree

| Academic Degree | Total Students | Percentage (%) |
|---|---|---|
| Postgraduate | 119 | 2.17% |
| Applied Bachelor | 4734 | 86.18% |
| Distance Learning | 640 | 11.65% |
| **Total** | **5493** | **100%** |

The analysis of student demographic origins, summarized in Table 5, provides key insights into the geographical reach and promotional effectiveness of the institution. The subsequent descriptive analysis revealed foundational insights for institutional strategy; specifically, the extreme geographical concentration of student enrollment in Surabaya (88.53%) provides actionable intelligence for strategic marketing planning, highlighting a critical over-reliance on a single city.

Table 5. Student Previous Education, Demographic, and Admission summarize

| Analysis Category | Detail | Percentage |
|---|---|---|
| Previous Education | General High School (SMA) | 46.02% |
| | Vocational High School (SMK) | 31.54% |
| | Others | 22,44% |
| 3 Highest Province Origin | East Java | 83.47 % |
| | Central Java | 10.81 % |
| | West Java | 8.68 % |
| 3 Highest City Origin | Surabaya | 88.53 % |
| | Sidoarjo | 34.74 % |
| | Gresik | 13.44 % |
| Admission Chanel | National Test-Based Admission | 29.36% |
| | Independent Admission | 27.23% |
| | National Achievement-Based Admission | 21.52% |
| | Others | 21.89% |

### *3.3. Dashboard Implementation*

The final system was developed using a robust Technology Stack to transform the analyzed data into an interactive dashboard. Python served as the core engine for data operations, preparing the resulting data in the JSON file format. The front-end, built using the ReactJS library, utilized HTML and CSS (with TailwindCSS) to provide a dynamic User Interface (UI). The integrated ECharts library consumed the JSON data to render the final visualizations, thereby effectively translating the research findings into an accessible and concise graphic solution. For brevity and clarity, only the key views that illustrate the system's core capabilities and principal findings are presented in the following figures.

Overall Student Profile presented in Figure 2 (a) and (b). The Aggregate Profile Student shown in Figure 2 (a) provides a high-level overview of key institutional metrics, displaying aggregated data such as Total Enrollment (1392), Graduated Students (946), achievements (154), and Average GPA (3.39). Figure 2 (b) visualizes the student distribution by province using a choropleth map with the Quantile classification method. This approach was selected due to its effectiveness in handling unevenly distributed data, ensuring an equal number of observations (provinces) falls within each color class. The main advantage of the Quantile method lies in the simplification of interpretation, as it allows for direct reference to data percentages, and logically links the Median to the central class. Furthermore, because it is based on relative ranking (rank order), the method produces a balanced visual distribution, offering a more honest representation of each province's comparative position within the context of the national student distribution [17].

(a) (b)

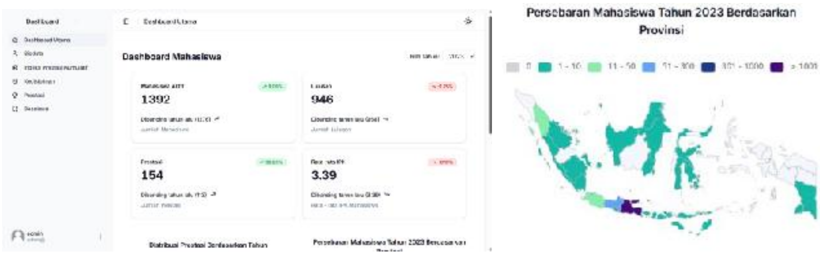


**Fig. 2.**(a) Aggregate Profile Student, and (b) Map Visualization for Student Distribution

Academic Performance Trend shown in Figure 3 highlights the temporal trends of GPA across different academic programs over several semesters. By allowing comparison per department, the dashboard effectively translates complex performance data into a clear analytical trend line, supporting comparative analysis and early intervention strategies in academic evaluation.

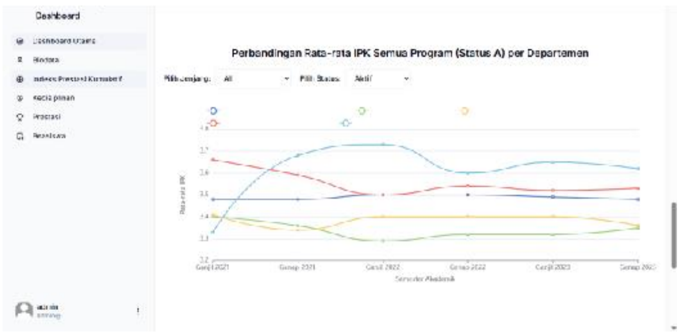


**Fig. 3.**Academic Performance

Non-Academic Performance and Reach presented in Figure 4 (a) and (b). Non-academic data is visualized to identify talent pools and activity concentrations. Event Distribution visualized by Figure 4(a), pie charts illustrate the distribution of student achievements, confirming the proportion between Academic and Non-Academic categories, validating the input from the achievements data source. Top 10 Achievement shown by Figure 4(b), the ranking display showcases the ten most frequent achievements per department, providing actionable insights into the specific areas of student excellence and the effectiveness of institutional support for competitive events.

(a) (b)

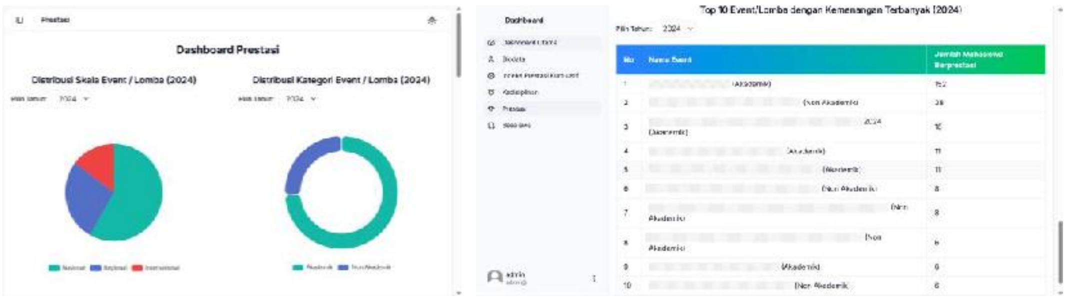


**Fig. 4.**(a) Event Distribution, and (b) Top 10 Achievement

Figure 5 presents the discipline of students, provides critical insights into student commitment and participation, effectively translating administrative records into actionable visual data. The dashboard utilizes dual charts to dissect attendance data: the chart on the left clearly visualizes the proportional attendance comparison between the Ganjil (Odd) and Genap (Even) semesters across all academic programs, helping to identify potential cyclical trends in student engagement.

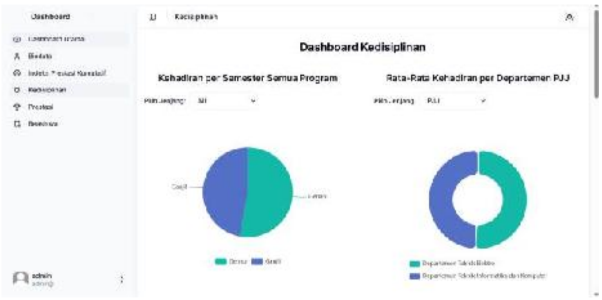


**Fig. 5.** Students Discipline

### *3.4. User Acceptance Testing Result*

The final validation involved User Acceptance Testing (UAT) performed by 82 respondents of XYZ University. The structured questionnaire administered was designed to assess the system's acceptance across four integrated evaluation aspects. These aspects, adapted from standard system quality models, measure critical user perceptions, namely: (1) Interface and Navigation Display, (2) Functionality and Performance, (3) Security and Privacy, and (4) Satisfaction and Usability. The total list of evaluation questions is presented in Table 6, which consists of questions specifically tailored to evaluate the effectiveness of the dashboard.

Table 6. List of Evaluation Question

| Number | Variable Aspect | Question |
|---|---|---|
| 1 | Interface and Navigation Display | The dashboard interface is easy to understand and use. |
| | | The navigation menu is clear and helps me easily find the information I need. |
| | | Graphs, tables, and data visualizations are displayed clearly and informatively. |
| 2 | Functionally and Performance | The GPA data displayed matches the students' academic records. |
| | | Discipline information (attendance, lateness, etc.) is displayed accurately. |
| | | Student achievement data (both academic and non-academic) is provided completely. |
| | | Student socio-demographic data is presented informatively and is easy to understand. |
| | | The search and filter features (e.g., by program, cohort, or region) work properly. |
| | | The dashboard can be accessed quickly without long loading times. |

| Number | Variable Aspect | Question |
|---|---|---|
| | | Graphs and tables are updated regularly according to schedule or in real-time. |
| 3 | Security and Privacy | The login and authentication system of the dashboard works properly. |
| | | Students' personal data is displayed securely without compromising privacy. |
| | | Students' personal data can only be accessed by authorized parties. |
| 4 | Satisfaction and Usability | The dashboard helps me understand students' academic and non-academic data. |
| | | The information displayed matches my needs as a user. |
| | | This dashboard is useful in supporting analysis and decision-making. |

Before initiating UAT, the survey questionnaire was validated to ensure both is measurement integrity (validity) and internal consistency (reliability). The validity of the instrument was assessed using the Pearson product-moment correlation by correlating the score of each item with the total construct score. The result showed that all sixteen items were declared valid, as they achieved a correlation coefficient for the items ranged from 0.5599 up to 0.8206. The internal consistency (reliability) of the multi dimensional survey instrument was tested using Cronbach's Alpha across the four identified latent aspects shown in Table 7 with all four measurement aspects above the minimum acceptable threshold 0.6. This confirms the high internal consistency and reliability of the survey instrument, providing a statistically foundation for the subsequent User Acceptance Testing

Table 7. Reliability Result

| Aspect (Domain) | Cronbach's Alpha | Interpretation |
|---|---|---|
| Interface and Navigation Display | 0.7795 | Good Reliability |
| Functionally and Performance | 0.8801 | Excellent Reliability |
| Security and Privacy | 0.6829 | Acceptable Reliability |
| Satisfaction and Usability | 0.8027 | Good Reliability |

The UAT procedure confirmed the system's operational viability and exceptionally high user acceptance, with the calculated results for each aspect summarized in Table 8. The final UAT score of 89.03% (Very Worthy) confirms the operational viability and high user acceptance of the developed dashboard. The strong scores in the Functionality and Security aspects (86.20% and 87.64% respectively) directly validate the system's capacity to serve as a reliable tool for decision support, effectively translating complex data (like performance trends and discipline records) into concise visual information as intended by the research objective.

Table 8. User Acceptance Testing (UAT) Results Summary

| Aspect Tested | Average Score (1-5) | Acceptance Rate (%) |
|---|---|---|
| Interface and Navigation Display | 4.53 | 90.65% |
| Functionally and Performance | 4.31 | 86.20% |
| Security and Privacy | 4.38 | 87.64% |
| Satisfaction and Usability | 4.58 | 91.63% |
| **Average** | **4.45** | **89.03%** |

However, the validation process also highlighted limitations. While functional aspects were highly rated, the user feedback regarding the color scheme being less eye-catching and inappropriate indicates that although the system is functional, its psychological impact and visual design could be optimized further. Therefore, future recommendations include revising the color palette and contrast shown in Figure 6 to improve user experience and visual clarity. This specific finding suggests that even highly reliable systems require continuous iteration on user interface aesthetics to maximize user engagement and adoption a critical consideration for future development efforts outlined in the conclusion.


**Fig. 6.** Word Cloud of User Feedback

## 4. CONCLUSION

This study successfully culminated in the design and empirical validation of an Interactive Dashboard, demonstrating the crucial necessity of rigorous methodological intervention to transform fragmented institutional data into actionable intelligence. The utilizing Secondary Data Analysis on institutional records demands intensive Feature Engineering and Mapping, consolidating inconsistent entries into stable categories to achieve the structural data integrity prerequisite for robust analysis, thereby addressing the reviewer's demand for findings regarding the method's process. The final system, developed using a modern technology stack, was conclusively validated through extensive User Acceptance Testing (UAT), yielding an outstanding Overall Acceptance Score of 89.03% ("Very Worthy"). This high empirical score, supported by validation from a diverse cohort including lecturers and administrative staff, confirms the system's suitability to function as a data-driven decision-support tool for knowledgeable stakeholders, which directly answers the reviewer's concern regarding the qualification and testing of the assessment sample. For future work, the study identifies two main directions: addressing user feedback regarding visual appeal, color mismatch, reviewing and improving the dashboard's color scheme and visual aesthetics to enhance user experience and engagement; and an expanded focus on integrating advanced analytics must be maintained to provide more comprehensive, engaging, and predictive insights, thereby enhancing the dashboard's strategic value for managerial decision-making. Additionally, to address the core challenges of data fragmentation identified in this study where intensive feature engineering was required to harmonize source inconsistencies future institutional efforts must prioritize the implementation of wide data standardization policies and the development of a centralized data repository or data warehouse. This strategic infrastructural investment will significantly reduce the pre-processing burden for future analytics projects, ensuring faster data availability and consistent data quality across all reporting and decision-making platforms.

**Acknowledgment**

The authors extend their sincere gratitude to the Institutional Data Research Team for their comprehensive support in granting secure access to the heterogeneous institutional data and for their expertise during the initial data extraction and preprocessing phase.

**Declarations**

**Author contribution.** Conceptualization and Methodology: Fitrah Maharani Humaira; Software Development (Python Pipeline): Wiratmoko Yuwono; Dashboard Interface (ReactJS) and Visualization: Rengga Asmara; Data Validation and User Acceptance Testing (UAT): Fitrah Maharani Humaira and Robi'Atul Adawiyah; Writing Original Draft Preparation: Robi'Atul Adawiyah; Writing, Review and Editing, and Supervision: Rusminto Tjatur Widodo. All authors have read and agreed to the published version of the manuscript.
**Funding statement.** This research was supported by DIPA PENS (2025).
**Conflict of interest.** The authors declare no conflict of interest.
**Additional information.** No additional information is available for this paper.

**Data and Software Availability Statements**

The primary heterogeneous institutional data analyzed in this study are sensitive and confidential, and therefore, are not publicly accessible. Data and processed metrics may be made available from the corresponding author upon reasonable request and with permission from the PENS. The software stack (Python scripts and ReactJS code) is available for non-commercial research use upon reasonable request to the corresponding author.

## References

[1] Ö. H. Kuzu, "Digital transformation in higher education: A case study on strategic plans," Vysshee Obrazovanie v Rossii = Higher Education in Russia, vol. 29, no. 3, pp. 9–23, 2020.

[2] W. K. Elugbaju, N. I. Okeke, and O. A. Alabi, "Conceptual framework for enhancing decision-making in higher education through data-driven governance," Global J. Advanced Research and Reviews, vol. 2, no. 02, pp. 016–030, 2024.

[3] J. R. McKean, "Using institutional data in curricular decision making," AURCO J., vol. 20, 2014.

[4] Anardani, S., Sofyana, L. S., & Maghfur, A. (2019). Analysis of business intelligence system design for student performance monitoring. Journal of Physics: Conference Series, 1381(1), 012015.

[5] Memon, M. A., Soomro, S., Jumani, A. K., & Kartio, M. A. (2017). Big Data Analytics and Its Applications. Annals of Emerging Technologies in Computing (AETIC), 1(1), 44–54.

[6] Jayashanka, R., Hettiarachchi, E., & Hewagamage, K. P. (2022). Technology Enhanced Learning Analytics Dashboard in Higher Education. The Electronic Journal of e-Learning, 20(2), 151–170.

[7] Altarazi, F., Santos, D., & Tomchesson, A. (2023). Framework for Developing Interactive Dashboards and Evaluating Their Usage Effectiveness: A Case Study. Proceedings of the 8th North American International Conference on Industrial Engineering and Operations Management, Houston, Texas, USA, June 13-16, 2023.

[8] Ajax, R., Joseph, O., & Own, J. (2025). Enhancing Business Intelligence with Data Visualization Tools.

[9] E. Smith, “Pitfalls and promises: The use of secondary data analysis in educational research,” Brit. J. Educ. Stud., vol. 56, no. 3, pp. 323–339, Jul. 2008.

[10] T. Dasu and T. Johnson, Exploratory data mining and data cleaning. Hoboken, NJ: John Wiley, 2003.

[11] S. B. Kotsiantis, D. Kanellopoulos, and P. E. Pintelas, “Data preprocessing for supervised learning,” Int. J. Comput. Sci., vol. 1, no. 1, pp. 110–117, Jan. 2006.

[12] H. Liu, F. Hussain, C. L. Tan, and M. Dash, “Discretization: An enabling technique,” Data Mining and Knowledge Discovery, vol. 6, pp. 393–423, Oct. 2002.

[13] I. Guyon and A. Elisseeff, “An introduction to variable and feature selection,” J. Mach. Learn. Res., vol. 3, pp. 1157–1182, Mar. 2003.

[14] N. Martins, S. Martins, and D. Brandão, “Design principles in the development of dashboards for business management,” in Perspectives on design II, [DOI:10.1007/978-3-030-79879-6_26], Jan. 2022, pp. 353–365.

[15] A. Parker, A. Heflin, and L. C. Jones, “Analyzing university of virginia health publications using open data, python, and streamlit,” J. Med. Libr. Assoc., vol. 109, no. 4, pp. 688–693, Nov. 2021.

[16] Riduwan and A. Akdon, Rumus dan data dalam analisis statistika untuk penelitian (administrasi pendidikan-bisnis-pemerintahan-sosial-kebijakan-ekonomi-hukum-manajemen-kesehatan). Bandung: Alfabeta.

[17] Slocum, Terry A., Robert B. McMaster, Fritz C. Kessler, and Hugh H. Howard. Thematic Cartography and Geovisualization. Fourth Edition. Boca Raton, FL: CRC Press, Taylor & Francis Group.

[18] J. Smith and A. Jones, Leveraging secondary data: Methodological considerations for social science research. London: Routledge, 2022.

[19] L. Chen, Q. Wang, and M. Li, "The role of administrative data in educational policy analysis," J. Educ. Data Sci., vol. 8, no. 3, pp. 155–172, 2021.

[20] H. Kim and S. Lee, "Best practices for data cleaning in large datasets: Addressing duplicates and inconsistencies," Int. J. Data Sci. Anal., vol. 10, no. 2, pp. 245–258, 2020.

[21] C. Brown and M. Davis, "Semantic consistency via domain dictionaries in data preprocessing," IEEE Trans. Knowl. Data Eng., vol. 35, no. 1, pp. 305–318, 2023.

[22] F. Garcia, et al., "Impact of feature discretization on non-linearity in predictive modeling," Expert Syst. Appl., vol. 200, p. 116890, 2022.

[23] A. Field, Discovering statistics using IBM SPSS statistics, 6th ed. London: SAGE Publications, 2021.

[24] R. B. Kline, Principles and practice of structural equation modeling, 5th ed. New York: Guilford Press, 2023.

[25] H. Tjendro and W. Lim, "Developing modular web applications for university dashboards using component-based frameworks," J. Inf. Technol. Syst. Des., vol. 12, no. 1, pp. 45–60, 2021.

[26] P. A. Williams, Full stack development with react and python. Sebastopol, CA: O'Reilly Media, 2020.

[27] S. Mahmood and M. Shariati, "Critical success factors for user acceptance testing in software projects," J. Softw. Eng. Appl., vol. 16, no. 5, pp. 235–249, 2023.

[28] U. Sekaran and R. Bougie, Research methods for business: A skill building approach, 9th ed. Hoboken, NJ: Wiley, 2023.

[29] M. Tavakol and R. Dennick, "Making sense of cronbach's alpha," Int. J. Med. Educ., vol. 2, pp. 53–55, 2011.

[30] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, Multivariate data analysis, 8th ed. Harlow: Pearson Education, 2024.

[31] J. C. Nunnally and I. H. Bernstein, Psychometric theory, 3rd ed. New York: McGraw-Hill, 1994.