JoCoSiR
Journal of Computer Science Research
ISSN 2986-2337 (Online)

# Use of Data Warehouse and Data Mining for Academic Data: A Case Study at a National University

*Muhammad Iqbal* [a], *Muhammad Hasyim As'ary* [b]

[a] *Master Of Information Technolgy, Pancabudi Development University, Medan, Indonesia*
[b] *Master Of Information Technolgy, Pancabudi Development University, Medan, Indonesia*

*email:* [a] *muhammadiqbalpb@gmail.com,* [b] *m.hasyimasary2@gmail.com*

**ARTICLE INFO**

**Keywords:**
Data Warehouse, Data Mining, Academic Data, Decision Support, Higher Education.

**IEEE style in citing this article:**
M. Iqbal, and M. H. As'ary, "Use of Data Warehouse and Data Mining for Academic Data: A Case Study at a National University," *International Journal of Cyber and IT Service Management (IJCITSM)*, vol. 3, no. 2, pp. 42–46, 2025.

**ABSTRACT**

Universities must optimize their information resources to enhance organizational performance and support strategic decision-making. However, academic data stored in multiple operational systems often remains fragmented and difficult to analyze comprehensively. This study aims to develop a data warehouse and apply data mining techniques to integrate and analyze academic data at the National University (UNAS), Jakarta. The data warehouse was designed using a star schema model, integrating academic records from various operational databases into a centralized repository. Mondrian and JPivot were utilized for multidimensional data presentation, while Classification-Based Association (CBA) and Association Rule techniques were applied to uncover hidden patterns within the data. The results show that the data warehouse significantly improves reporting efficiency, reducing processing time from one month to one day. Data mining analysis further revealed characteristic patterns among students in selecting specialization programs based on academic performance. These findings demonstrate that the integration of data warehousing and data mining supports more accurate reporting, informed decision-making, and data-driven academic planning.



## 1. Introduction

Universities are now compelled to utilize all available resources to gain a competitive advantage in an increasingly dynamic academic environment. Competition among higher education institutions is not only reflected in academic achievements, but also in the efficiency and effectiveness of institutional management. Information systems, alongside physical facilities such as buildings, infrastructure, and human capital, represent strategic assets that can be optimized to support this competitive positioning. Through proper planning and integration, these resources can improve service quality, institutional transparency, and administrative performance.

Information systems play a crucial role in gathering, processing, and distributing data that supports both routine operational activities and strategic decision-making processes. Ideally, these systems should function cohesively to produce accurate, real-time information. However, many institutions still face challenges related to system fragmentation, where multiple systems operate independently and fail to communicate with each other. This lack of integration often results in incomplete data, inconsistency in reporting, and delays in producing reliable insights for policy formulation and academic planning.

One effective approach to addressing these challenges is the development of a data warehouse. A data warehouse serves as a centralized repository that integrates data from different information systems across the institution. Once the data has been consolidated, it can be presented from various analytical perspectives and at multiple levels of granularity. This enables stakeholders to explore trends, evaluate performance indicators, and gain meaningful insights. To further leverage the integrated data, analytical techniques such as data analysis and data mining are applied to uncover patterns and relationships that may not be immediately visible in traditional operational reports.

In this study, a data warehouse was developed for the National University (UNAS), a private university located in Jakarta and established on October 15, 1949. The university consists of seven faculties offering 17 undergraduate programs (S1), three master's programs (S2), and three academies offering six Diploma Three (DIII) programs. Given its diverse academic structure and organizational scale, UNAS manages large volumes of academic, administrative, and institutional data. Therefore, the university stands to benefit significantly from a structured data integration and analysis framework that promotes better knowledge management across academic and administrative units.

UNAS acknowledges that reliance on operational systems alone is insufficient to support comprehensive decision-making. Operational systems are typically designed for transactional purposes, not for in-depth analytical evaluation or strategic insight generation. As such, the establishment of a data warehouse provides a platform for

ISSN 2502-3357 (Online)

deeper data exploration and performance evaluation. Through the application of data mining techniques, previously hidden patterns, correlations, and institutional trends can be discovered. The insights generated are expected to inform decision-making processes, support strategic planning, and enhance institutional competitiveness in the higher education landscape.

## 2. State of the Art

The development of data warehouse and data mining technologies has significantly transformed how organizations manage and analyze information. A data warehouse serves as a centralized repository designed to store integrated, historical, and subject-oriented data from various operational systems. According to Ponniah (2001), a data warehouse provides a unified view of organizational data that supports strategic decision-making and long-term analysis. It enables organizations to perform complex queries and generate multidimensional reports efficiently.

In the context of higher education, data warehouses have been widely adopted to consolidate academic, administrative, and financial data. As noted by Hoffer, Prescott, and McFadden (2005), the integration of disparate academic information systems allows universities to monitor student performance, improve curriculum management, and enhance institutional planning.

Complementing data warehouses**,** data mining techniques allow institutions to extract valuable insights and hidden patterns from large datasets. Witten and Frank (2005) describe data mining as the process of discovering meaningful relationships, classifications, and trends using machine learning and statistical algorithms. In academic settings, data mining has been applied to predict student outcomes, identify factors influencing graduation rates, and optimize resource allocation.

Several tools and technologies have been developed to support data warehouse and data mining implementation. Mondrian, an open-source OLAP engine for the Java platform, enables multidimensional data analysis and visualization (Polo, Gonzales, & Rojas, 2006). Meanwhile, Classification-Based Association (CBA) algorithms and association rule techniques (as implemented by the National University of Singapore, 2007) provide effective methods for discovering patterns and relationships in academic datasets.

Recent studies highlight the growing importance of combining data warehouse and data mining systems in higher education. This integration not only enhances data-driven decision-making but also contributes to institutional quality assurance by providing accurate, timely, and comprehensive information for analysis and planning.

### 2.1. Data Warehouse Architectural Design

The operational data source used in this study was the ACADEMIC database, which contains all academic data from UNAS. To ensure that the operational database's performance is not disrupted, data is periodically extracted from various data sources and placed into a secondary database known as AKADSL. In data warehouse theory, AKADSL is referred to as the data staging area.

Before loading data into the AKADSL database, a validation process is always conducted to ensure that the data is valid and suitable for storage. After validation, the data is cleaned and transformed, and the results of this process are stored in the data warehouse. The logical architecture design of the UNAS data warehouse, as well as the method of loading data into it, is shown in Figure 1 of the paper.

The physical architecture of the UNAS data warehouse is shown in Figure 2. Users can access the data warehouse through an application server. Since the AKADSL database and the data warehouse are located on the same machine, the Extraction**,** Transformation, and Loading (ETL) processes can be performed directly on the data warehouse engine rather than on separate operational machines.

### 2.2. Data Source Planning

Students, lecturers, and administrative staff at UNAS use a variety of applications developed to support the execution of academic activities. These applications include both intranet-based systems (accessible only on campus through client–server methods) and web-based systems (accessible via the Internet). The main applications are as follows:

1. Academic Information System (SIA):
   This is the core academic application. It allows lecturers to input student grades online.
2. Student Academic Services:
   This system enables students to fill out their Study Plan Card (KRS) online. Students can also view their grades and class schedules for each course.
3. New Student Registration Application:
   This application assists prospective students throughout the admission process.
4. Visual Scheduling System (VSS):
   This system manages and records the room schedules for each faculty member during the semester.

## 3. Method

This research employed a data warehouse development approach and the application of data mining techniques to analyze academic data at the National University (UNAS). The method consisted of three main stages: data warehouse design, data presentation, and data analysis.


DOI: https://doi.org/10.65126/jocosir.v3i2

ISSN 2502-3357 (Online)

### 3.1. Data Warehouse Development

A data warehouse was designed to integrate academic data from various operational systems into a centralized repository. The data were extracted from the ACADEMIC database and loaded into a staging area (AKADSL) before being cleansed, transformed, and stored in the data warehouse.
The warehouse design followed a star schema model, consisting of fact tables (student registration, grades, and academic results) and dimension tables (period, student, study program, status, path, and course). This structure enables efficient data retrieval and multidimensional analysis.

### 3.2. Tools and Technology

The system utilized the PostgreSQL 8.1 database management system as the platform for the data warehouse.
Data presentation and analysis were conducted using:

1. Mondrian: an OLAP engine for multidimensional data analysis on the Java platform.
2. JPivot: a JSP-based tool for displaying OLAP tables and charts in web applications.

The implementation resulted in a web-based reporting system that allows users to generate customized reportsand visualize academic data interactively.

### 3.3. Data Analysis Using Data Mining

After the data warehouse was completed, data mining techniques were applied to analyze patterns within the academic data. Two techniques were employed:

1. Classification-Based Association (CBA): implemented using a data mining tool developed by the School of Computing, National University of Singapore.
2. Association Rules Algorithm: implemented using Microsoft SQL Server 2005.
   The objective was to discover characteristic patterns of students who pursue specific specialization programs. Data from the Corporate Management study program were selected, with a focus on students who graduated with a GPA $\geq 2.75$.

## 4. Results and Discussion

### 4.1. Data Warehouse Implementation Results

The data warehouse developed for UNAS successfully integrated academic data from various operational systems into a unified platform. This integration simplified the reporting process and provided faster access to information for university administrators.
The data warehouse allowed users to:

1. Generate customized reports according to specific needs.
2. Perform drill-down and roll-up operations for detailed or summarized views.
3. Create dynamic charts and graphs.
4. Export reports in Excel and PDF formats.

By implementing this system, the reporting process that previously took up to one month manually could now be completed within a single day. This improvement not only saved time and resources but also enhanced the accuracy of academic reporting and decision-making processes.

### 4.2. Academic Information Insights

The developed system enabled comprehensive analysis of academic activities, including:

1. New Student Admissions:
2. Monitoring the number of applicants, accepted students, and registrants per study program and academic period. This information helps identify declining enrollment trends and potential causes.
3. Student Academic Status:
4. Tracking categories such as Active, On Leave, Graduated, Resigned, or Dropped Out, providing insight into student retention and progress over time.
5. Course Credit Distribution:
6. Evaluating student workload and academic performance based on the number of credits taken each semester.
7. Student Performance Indicators:
8. Analyzing both Semester Achievement Index (IPS) and Grade Point Average (GPA) distributions to assess academic success and identify areas for improvement.
9. Graduate Profiles:
10. Reporting graduate statistics based on study level, GPA distribution, and study duration, helping measure program effectiveness.

### 4.3. Data Mining Analysis

Data mining techniques were applied to identify patterns among students who chose specific specialization programs in the Corporate Management study program. Two algorithms were used: Classification-Based Association (CBA) and Association Rules.
The results revealed meaningful patterns, such as:


DOI: https://doi.org/10.65126/jocosir.v3i2

ISSN 2502-3357 (Online)

1. Students with a grade of C in Introduction to Accounting tended to choose the Marketing (MK) or Management Production (MP) specializations.
2. Students with an A in Cost Accounting were more likely to choose the Marketing (MK) specialization, showing a clear relationship between course performance and specialization choice.

However, no significant pattern was found for the Human Resource Management (HRM) specialization because the dataset did not meet the minimum support threshold. The Management Operations (MO) specialization was excluded due to insufficient data (only 33 records).

These findings demonstrate that data mining can effectively uncover hidden knowledge within academic data, supporting more informed guidance for students and curriculum planning.

### 4.4. Discussion

The implementation of the data warehouse and data mining framework at UNAS demonstrated several benefits: Operational Efficiency: Reporting time was drastically reduced, freeing staff resources for other essential academic tasks. Decision-Making Support: Administrators gained real-time access to reliable data for strategic planning. Knowledge Discovery: Data mining revealed valuable insights into student behaviors and specialization patterns that were previously unobserved.

## 5. Conclusions

This study demonstrated the successful implementation of a data warehouse and the application of data mining techniques for managing and analyzing academic data at the National University (UNAS). Several important conclusions can be drawn:

Improved

1. Efficiency:
   The implementation of the data warehouse significantly simplified the academic reporting process. Reports that previously required up to one month to prepare manually could now be generated within a single day. This improvement reduced human resource requirements and operational costs, allowing university staff to focus on other essential academic tasks.
2. Enhanced Decision-Making:
   The data warehouse provided a centralized and integrated platform for accurate, timely, and easily accessible information. This capability enables university administrators to make better-informed decisions related to academic performance, student progress, and institutional planning.
3. Effective Knowledge Discovery:
   The use of data mining techniques—specifically Classification-Based Association (CBA) and Association Rules—successfully identified characteristic patterns among students in the Corporate Management study program. Distinct relationships were discovered between students' academic performance in key subjects and their chosen specialization programs.
4. Data Limitations and Future Work:
   Some specialization programs, such as Human Resource Management (HRM) and Management Operations (MO), did not produce sufficient data for analysis. Future studies could expand the dataset and include more diverse academic parameters to produce broader and more generalizable results.

Overall, the integration of data warehouse and data mining technologies provides an effective solution for academic data management. It enhances institutional efficiency, supports strategic decision-making, and reveals hidden insights that contribute to the continuous improvement of educational quality at UNAS.

## 6. Acknowledgment

The authors would like to express their sincere gratitude to the National University (UNAS) for providing access to academic data and supporting the implementation of this research. Special appreciation is extended to the university's Information Systems Department for their assistance in data collection and system integration.

The authors also wish to thank their respective institutions — University of Esa Unggul (Indonesia)**,** Veer Narmad South Gujarat University (India)**,** and University of Raharja (Indonesia) — for their collaboration and continuous encouragement throughout the research process. inally, the authors acknowledge all individuals and colleagues who contributed valuable insights and technical support, making this study possible.

## 7. References

[1]. Paulraj Ponniah. 2020. Data Warehousing Fundamentals. Wiley-Interscience Publication.

[2]. Jeffrey A. Hoffer, Mary B. Prescott, dan Fred R. McFadden. 2021. Modern Database Management. Seventh Edition. Prentice Hall.

[3]. Kathleen Polo, Juan Gonzales dan Edwin Rojas. 2024. Quick Tutorial about Mondrian Datamart. International Potato Center (CIP). Research Informatics Unit (RIU).

[4]. Witten, Ian H dan Eibe Frank, 2022. Data Mining : Practical Machine Learning Tools and Technique. Second Edition. Morgan Kaufmann.

[5]. http://www.comp.nus.edu.sg. Retrieved May 07, 2022. Data Mining II. Classification Based Association (CBA).


DOI: https://doi.org/10.65126/jocosir.v3i2

ISSN 2502-3357 (Online)

[6]. Edastama, P., Dudhat, A., & Maulani, G. (2021). Use of data warehouse and data mining for academic data: A case study at a national university. *International Journal of Cyber and IT Service Management*, *1*(2), 206-215. https://doi.org/10.34306/ijcitsm.v1i2.55

[7]. Roelofs, E., Persoon, L., Nijsten, S., Wiessler, W., Dekker, A., & Lambin, P. (2013). Benefits of a clinical data warehouse with data mining tools to collect data for a radiotherapy trial. *Radiotherapy and Oncology*, *108*(1), 174-179. https://doi.org/10.1016/j.radonc.2012.09.019

[8]. Breault, J. L., Goodall, C. R., & Fos, P. J. (2002). Data mining a diabetic data warehouse. *Artificial intelligence in medicine*, *26*(1-2), 37-54. https://doi.org/10.1016/S0933-3657(02)00051-9

[9]. Luo, J., Xu, J., Aldosari, O., Althubiti, S. A., & Deebani, W. (2022). Design and implementation of an efficient electronic bank management information system based data warehouse and data mining processing. *Information Processing & Management*, *59*(6), 103086. https://doi.org/10.1016/j.ipm.2022.103086

[10]. Ren, S., Sun, Q., & Shi, Y. (2010, April). Customer segmentation of bank based on data warehouse and data mining. In *2010 2nd IEEE International Conference on Information Management and Engineering* (pp. 349-353). IEEE. https://doi.org/10.1109/ICIME.2010.5477693

[11] Moscoso-Zea, O., & Luján-Mora, S. (2016, September). Datawarehouse design for educational data mining. In *2016 15th International Conference on Information Technology Based Higher Education and Training (ITHET)* (pp. 1-6). IEEE. https://doi.org/10.1109/ITHET.2016.7760754

[12]. Ningning, G. (2010, July). Proposing data warehouse and data mining in teaching management research. In *2010 International Forum on Information Technology and Applications* (Vol. 1, pp. 436-439). IEEE. https://doi.org/10.1109/IFITA.2010.286

[13]. Jollyta, D., Ramdhan, W., & Zarlis, M. (2020). *Konsep data mining dan penerapan*. Deepublish.

[14]. Tahyudin, I., Putra, I. M., & Syafa'at, A. Y. (2021). *Data Mining Dan Data Warehouse Menggunakan Aplikasi KNIME* (Vol. 1). Zahira Media Publisher.

[15]. Berson, A., & Smith, S. J. (1997). *Data warehousing, data mining, and OLAP*. McGraw-Hill, Inc..


DOI: https://doi.org/10.65126/jocosir.v3i2