PERANCANGAN DATA WAREHOUSE UNTUK PENJUALAN TIKET PESAWAT DI PT ABC

DESIGNING A DATA WAREHOUSE FOR AIRLINE TICKET SALES AT PT ABC

Sekar Aurannisa Ramdhani Qadriah[1], Jap Tji Beng[2,6*], Desi Arisandi[2], Fouad Nagm[3,6], Tasya Mulia Salsabila[4,6], Tiara Zahro[1,6], Vienchenzia Oeyta Dwitama Dinatha[5,6]
Information Systems Undergraduate Program, Universitas Tarumanagara, Jakarta, Indonesia[1]
Faculty of Information Technology, Universitas Tarumanagara, Jakarta, Indonesia[2]
BT Financial Group, Sydney, Australia[3]
Faculty of Computer Science, Universitas Indonesia, Jakarta, Indonesia[4]
Doctorate of Business Administration, INTI International University, Malaysia[5]
Science, Technology and Society Research Centre, Universitas Tarumanagara, Jakarta, Indonesia[6]
t.jap@untar.ac.id[2*]

ABSTRACT
The high demand for airline ticket purchases in Indonesia and ticket sales being one of the important assets for companies in the aviation industry. PT ABC requires a system that can effectively manage this data to support business analysis and strategic decision-making. This research aims to design a data warehouse that can integrate airline ticket sales data, enabling more efficient data storage, processing, and analysis. The methodology employs the Nine-Step Kimball approach for dimensional modeling, implemented through an ETL (Extract, Transform, Load) process using SQL Server Integration Services (SSIS). The approach used in the design of this data warehouse is the star schema, which consists of one main fact table and several dimension tables. The fact table contains ticket sales transaction data, while the dimension tables include information about time, flight routes, airlines, and customers. This research provides a primary contribution in the form of a data warehouse design that integrates airline ticket sales data, enabling faster and more accurate analysis. The resulting business impact includes increased operational efficiency, more accurate data-driven decision-making, and the ability to monitor sales trends in real-time through interactive dashboards. With this data warehouse, PT ABC can reduce data analysis time and improve the accuracy of sales trend predictions.

Keywords: *Data warehouse, Star Schema, PT ABC plane ticket sales, Business Intelligence.*

ABSTRAK
Permintaan yang tinggi terhadap pembelian tiket pesawat di Indonesia dan penjualan tiket menjadi salah satu aset penting bagi perusahaan di industri penerbangan. PT ABC memerlukan sistem yang dapat mengelola data ini secara efektif untuk mendukung analisis bisnis dan pengambilan keputusan strategis. Penelitian ini bertujuan untuk merancang data warehouse yang dapat mengintegrasikan data penjualan tiket pesawat, sehingga memungkinkan penyimpanan, pemrosesan, dan analisis data yang lebih efisien. Metodologi ini menggunakan pendekatan Nine-Step Kimball untuk pemodelan dimensional, yang diimplementasikan melalui proses ETL (Extract, Transform, Load) menggunakan SQL Server Integration Services (SSIS). Pendekatan yang digunakan dalam desain data warehouse ini adalah star schema, yang terdiri dari satu tabel fakta utama dan beberapa tabel dimensi. Tabel fakta berisi data transaksi penjualan tiket, sementara tabel dimensi mencakup informasi tentang waktu, rute penerbangan, maskapai penerbangan, dan pelanggan. Penelitian ini memberikan kontribusi utama berupa desain data warehouse yang mengintegrasikan data penjualan tiket pesawat, memungkinkan analisis yang lebih cepat dan akurat. Dampak bisnis yang dihasilkan meliputi peningkatan efisiensi operasional, pengambilan keputusan berbasis data yang lebih akurat, dan kemampuan memantau tren penjualan secara real-time melalui dashboard interaktif. Dengan data warehouse ini, PT ABC dapat mengurangi waktu analisis data dan meningkatkan akurasi prediksi tren penjualan.

Kata Kunci: Data warehouse, Star Schema, PT ABC Plane Ticket Sales, Business Intelligence.

INTRODUCTION

In the rapidly evolving digital era, significant changes have occurred in various aspects (Lukas et al., 2024), including data management, which has become a crucial factor in supporting the operational success of companies (Subroto & Endaryati, 2021). This is particularly evident in the aviation industry, which continues to experience rapid growth

alongside increasing mobility and global economic expansion, as reflected in the rising number of domestic private airlines being established (Utama & Rezki, 2021). PT ABC, as a provider of airline ticket sales services, generates a large volume of data daily from various transactions. This data includes critical information such as passenger details, flight schedules, ticket prices, and payments. However, PT ABC faces challenges in managing the growing volume of airline ticket transaction data, making effective analysis and strategic decision-making increasingly difficult.

PT ABC faces challenges in integrating ticket sales data from various sources, causing delays in analysis and strategic decision-making. The continuously growing data volume complicates manual data processing and analysis. The need for a centralized system like a data warehouse becomes urgent to ensure quick access to accurate sales information and to support data-driven decisions. While previous studies have implemented data warehouses in aviation contexts (Nambiar & Mundra, 2022; Dhaouadi et al., 2022), most focus on single system integration or operational reporting rather than comprehensive sales analytics across multiple source systems. This research addresses the gap by developing an integrated dimensional model that combines data and booking platforms to enable cross-system analytics and predictive demand forecasting.

A data warehouse enables the integration of data from various sources, helping companies unify and manage data efficiently to support business analysis and assist management in decision-making processes by providing quick and easy access to accurate and relevant information. In the context of airline ticket sales, a data warehouse can assist the company in consolidating data from various sales systems, thereby empowering management to make more effective decisions regarding airline ticket sales strategies (Nambiar & Mundra, 2022; Dhaouadi et al., 2022).

This research aims to design a data warehouse tailored to the needs of PT ABC for managing airline ticket sales data. The ETL (Extract, Transform, Load) process plays a crucial role in managing and integrating ticket sales transaction data from various platforms and systems. With an effective ETL process, data can be transformed into information ready for sales analysis and customer trend insights. Additionally, the developed ETL process is designed to handle large volumes of data efficiently (Dhaouadi et al., 2022; Peng et al., 2023). The data warehouse design in this research employs SQL Server Management Studio 2022. The outcome of this study is expected to provide strategic guidance for PT ABC in enhancing their operational efficiency and the effectiveness of their sales data analysis.

METHODS
Data Collection
Data collection for this study was conducted through interviews. This stage began with interviews involving PT ABC staff, conducted at PT ABC's office. These interviews focused on identifying the company's needs, leading to the conclusion that designing a star schema-based data warehouse is necessary to enhance. The entire data set covering the period from 2023 to 2024.

Nine-Step Kimball Methodology
The Nine-Step Kimball Methodology was utilized in designing the data warehouse because it facilitates the creation of a well-organized data warehouse, from selecting business processes to physical design (Sihombing, 2022). This methodology provides a clear framework for developing a data warehouse, encompassing steps such as selecting processes, identifying facts and dimensions, and adjusting dimensions as needed. This structured approach helps minimize the risk of errors during the system design process (Pratama & Raharja, 2023).

The Kimball Methodology emphasizes the importance of involving users in the data warehouse development process. By adopting a user-centered approach, the resulting solution is more responsive to their needs, thereby enhancing the accuracy of data-driven decision-making. The primary focus of the Nine-Step Kimball Methodology is to ensure that the developed data warehouse effectively supports faster and more accurate decision-making, ultimately improving the overall performance of the organization (Halim et al., 2024). This methodology enables the use of ETL (Extract, Transform, Load) techniques to ensure data is properly processed and prepared for business analysis. In the context of this research, the ETL method is employed to process data from various sources into a more structured form, which is then utilized for analysis (Pratama & Raharja, 2023).

ETL (Extract, Transform, Load)

The ETL (Extract, Transform, Load) process plays a crucial role in integrating data from various sources for the development of a data warehouse that supports analysis and decision-making in airline ticket sales at PT ABC (Barahama & Wardani, 2021). The ETL process is implemented using Microsoft SQL Server Integration Services (SSIS) by performing Extraction, Transformation, and Loading.

The ETL process allows for extracting data from multiple sources, transforming it according to business needs, and loading it into a unified, integrated database (Madhikerrni & Främling, 2019). The stages of the ETL process are as follows:

A. Extract

The first stage in the ETL process is data extraction from various sources. In this phase, the data used to design the star schema comes from CSV files. This data is read and processed to ensure that the information received is ready for further transformation and analysis.

B. Transform

The second stage is data transformation, which involves data cleaning. The extracted data typically comes in various formats, so it needs to be cleaned and transformed to meet the requirements of the data warehouse. The specific transformations performed include, standardization of dates or times across the entire system, matching and removal of duplicate customer IDs, and calculation of derived metrics such as booking time and revenue per passenger. This transformation process ensures that the data loaded into the data warehouse is consistent and easy to analyze.

C. Load

After the data is transformed, the next step is to load it into the PT ABC data warehouse. During this phase, the data is inserted into the appropriate fact tables and dimension tables within the data warehouse.

The choice of the star schema in this study is based on its simplicity in organizing large transaction data into a structure that is easier and more efficient for analytical queries. Compared to other methods like the snowflake schema, the star schema allows faster data access due to the direct relationship between fact tables and dimension tables. ETL was chosen for its ability to systematically integrate data from various sources through extraction, transformation, and loading processes, which is essential for ensuring data consistency and accuracy in the data warehouse.

RESULTS AND DISCUSSION

The design of the data warehouse for ticket sales at PT ABC demonstrates that the implementation of the star schema method and the ETL (Extract, Transform, Load) process is highly effective in integrating and managing ticket sales data. As shown in Figure 1, the star schema consists of the main fact table fact_sales connected to dimension tables for time,

customers, flight routes, airlines, and booking channels.

The data warehouse consists of a single fact table that stores transaction data related to ticket sales, and several dimension tables that include essential information such as time, customers, flight routes, and airlines.

This implementation allows PT ABC to access sales information quickly and in a structured manner, supporting more efficient business analysis processes. Additionally, the use of SQL Server Management Studio 2022 as the platform for managing the data warehouse simplifies query and data analysis tasks.
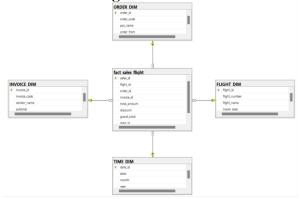
The ETL process applied in this study has proven to be effective in integrating data from various sources quickly and efficiently. The data extracted from CSV files and PT ABC's internal systems was transformed and loaded into the data warehouse without compromising integrity. Data cleansing was carried out to ensure that the data entered into the warehouse was free from duplicates and errors, thus enhancing the quality of the resulting analysis. This transformation stage ensured that the data format was consistent and ready for analysis.

With the implementation of this data warehouse system, PT ABC now has a more advanced tool for analyzing ticket sales trends, flight route performance, and customer behavior. The use of interactive dashboards also enables real-time data visualization, which facilitates decision-making to address issues promptly (Beng et al., 2023). The dashboard empowers management to make informed decisions based on accurate and relevant data, such as maximizing sales on specific routes or offering promotions to targeted customer segments.

The implementation results show that the ETL process can handle many transaction records with an average processing time of 30 minutes, demonstrating efficiency in large-scale data processing. Scalability testing was conducted by increasing the data volume, and the system maintained stable performance without significant degradation. Real-world testing at PT ABC showed a significant increase in sales analysis efficiency and enabled tracking of flight route performance through the dashboard.

Compared to the previous manual system, the data warehouse enables the consolidation of data from various sources into a centralized repository. Through the implementation of optimized indexing strategies, key implementation challenges such as poor data quality from legacy systems, timezone synchronization issues, and initial performance constraints were successfully resolved. While the old system took several hours to generate a sales report, the new system takes only minutes. Additionally, this solution is superior to traditional database approaches because it supports multidimensional analysis and allows data visualization through interactive dashboards, accelerating decision-making. The star schema design can be seen in **Figure 1.**

Figure 1. Star Schema Diagram
Source : Personal Documentation

CONCLUSION

This research successfully designed a data warehouse for airline ticket sales at PT ABC using the star schema method and the ETL (Extract, Transform, Load) process. The implementation of this data warehouse allows for the integration of data from various sources, data cleaning, and efficient processing, making ticket sales data easier and faster to access and analyze. However,

this study has several limitations, it uses only one case study, focuses only on structured data, and has not been implemented in a large-scale production environment. The use of SQL Server Management Studio 2022 also aids in data management and querying, supporting further visualization and analysis through interactive dashboards.

Although the designed data warehouse successfully improves data processing and analysis efficiency at PT ABC, there are some limitations. The system is not yet optimized for semi-structured or unstructured data, such as customer reviews. Future improvements include developing a more adaptive ETL process for various data formats and continuously evaluating performance to ensure scalability as data grows. Additionally, the application of machine learning techniques could enhance future sales trend predictions.

REFERENCES

Barahama, A. D., & Wardani, R. (2021, November). Utilization extract, transform, load for developing data warehouse in education using Pentaho Data Integration. In *Journal of Physics: Conference Series* (Vol. 2111, No. 1, p. 012030). IOP Publishing.

Beng, J. T., Amanto, A. F., Aurelia, A., Chandra, D., Mandey, K. Y. D., Ramadhani, L. A., ... & Tiatri, S. (2023, December). Designing mathematics, science, and reading competency dashboard using business intelligence algorithm. In *AIP Conference Proceedings* (Vol. 2680, No. 1, p. 020179). AIP Publishing LLC.

Dhaouadi, A., Bousselmi, K., Gammoudi, M. M., Monnet, S., & Hammoudi, S. (2022). Data warehousing process modeling from classical approaches to new trends: Main features and comparisons. *Data*, 7(8), 113.

Halim, N. A., Alhazmy, N. A., & Rasyad, M. A. (2024). Advanced Data Modeling dengan Metodologi Kimball: Systematic Literature Review. *Jurnal Sosial Teknologi*, 4(8), 639-650.

Lukas, K., Arisandi, D., & Jap, T. B. (2024). Perancangan Aplikasi Promosi dan Penjualan Produk Berbasis Website Pada Toko Baju Wearsarey. *Jurnal Ilmu Komputer dan Sistem Informasi*, 12(1).

Madhikerrni, M., & Främling, K. (2019, May). Data discovery method for extract-transform-load. In *IEEE International Conference on Mechanical and Intelligent Manufacturing Technologies* (pp. 174-181). IEEE.

Nambiar, A., & Mundra, D. (2022). An overview of data warehouse and data lake in modern enterprise data management. Big data and cognitive computing, 6(4), 132.

Peng, Y., Henke, E., Reinecke, I., Zoch, M., Sedlmayr, M., & Bathelt, F. (2023). An ETL-process design for data harmonization to participate in international research with German real-world data based on FHIR and OMOP CDM. *International Journal of Medical Informatics*, 169, 104925.

Pratama, I. P. A. E., & Raharja, I. M. S. (2023). Business Intelligence Based on Kimball Nine-Steps Methodology for Monitoring the Feasibility of Goods in Market. *International Journal of Advances in Data and Information Systems*, 4(2), 135-144.

Sihombing, D. J. C. (2022). Academic Data Warehouse Modeling In Higher Education Using Nine-Step Design Methodology. *Journal Of Information Systems And Informatics*, 4(4).

Subroto, V. K., & Endaryati, E. (2021). Business Intelligence Dan Kesuksesan Bisnis Di Era Digital. *Dinamika: Jurnal*

Manajemen Sosial Ekonomi, *1*(2), 41-47.

Utama, B. D., & Rezki, J. F. (2021). Perkembangan industri penerbangan dan pertumbuhan ekonomi di Indonesia. *Jurnal Ilmu Pemerintahan Suara Khatulistiwa*, *6*(2), 213-223.