Data Warehouse Analysis to Support UMKM Decisions using the Nine-step Kimball Method

I Gede Wira Darma[1*], Komang Sri Utami[2], and Ni Wayan Sri Aryani[3]

[1,2]Department of Electrical and Computer Engineering, Post Graduate Program, Udayana University,
[3]Department of Electrical and Computer Engineering, Udayana University
*Email: wiradarma@unud.ac.id

Abstract Micro, small and medium enterprises are one of the drivers of the economy. In the current technological era, companies have supported information systems to process business transactions. Naris Mart is one of the MSMEs engaged in retail with a business that has been running for 6 years. During the span of 6 years, of course, many buying and selling transactions were carried out. To help analyze how the business work, will be designed a data warehouse. Data warehouses need to be made so that they can receive information, reports, and can carry out multi-dimensional analysis that can ultimately help business owners. The design of the data warehouse uses the Nine-step Kimball methodology. Results obtained in the form of star schema and retail data warehouse analysis. The data warehouse can provide fast, accurate and continuous information that can help management in making policies for the future to come. In general, the benefit of this research is that additional references in building a data warehouse use the Nine-step Kimball methodology.

Keywords—Data warehouse, Nine-step Kimball, Star schema, web application, retail.

I. INTRODUCTION

Naris Mart is one of the MSMEs engaged in retail with a business that has been running for 6 years. During the span of 6 years, of course there have been many buying and selling data transactions.

During this time for the need for strategic information still using operational data that has not been optimized and the existing system has not been equipped with reporting features that help the process of analyzing existing transactions so that the process for obtaining the information needed requires a relatively long time and is not fully accurate. Based on the above problems, a data warehouse application is needed for extracting transaction data as material for preparing the required reports[1].

Data Warehouse is a database that is designed to work on the process of query, reporting and analysis. The data stored in a Data Warehouse is the history data of an organization/company to which the data is not stored in detail[2]

There are many methodologies in building a data warehouse. One of them is Kimball's method. Kimball & Ross (2010) says there are 9 steps in building a data warehouse, known as the nine-step methodology. The use of the Kimball approach is based on the ease of the user in analyzing the required reports[3]. The architecture that will be used in this design is a centralized data warehouse architecture. Centralized data architecture chosen because to facilitate the maintenance and supervision process[4]. The process of integrating and extracting data is done with the ETL concept (Extracts, Transformation, Loading).

II. PURPOSE OF PAPER

The main objective of this research is to design and implement a data warehouse application to help business owners analyze their business.

III. LITERATUR REVIEW

A. Data Warehouse

The data warehouse is "a subject-oriented, integrated data set, has a time variant, and does not change that can support the functions of the decision support system (DSS), where each unit of data is relevant for a moment at a time". According to Rainer and Turban "Data Warehouse is a place to store object-oriented historical data, which has been arranged so that it can be accessed and received for analysis process activities"[5].

B. Nine-steps Kimball

The data warehouse design method according to Kimball which is used includes 9 stages known as Nine-step Methodology[6], the nine stages are:

Step 1: Choose the Process.
Step 2: Choose the Grain.
Step 3: Identify And conform the dimensions.
Step 4: Choose the Fact.
Step 5: Store pre-calculations in the fact table.
Step 6: Rounding out the dimensions table.
Step 7: Decide the Duration of the database and periodicity of update.
Step 8: Track Slowly the changing dimensions.
Step 9: Decide the query priorities and the query mode.

C. Process Extract, Transform and Load (ETL)

ETL is a very important process, with ETL data can be incorporated into the data warehouse. ETL can also be used to integrate data with previously existing systems. ETL process includes the extraction of data from the source, transforming them into the new format in accordance with business needs and then put it in the target data structure. The purpose of ETL is collect, filter, manipulate and combine the relevant data from various sources to be stored into the data warehouse.

The result of the process of ETL is it generates data that meets the criteria data warehouse such as historical data, integrated, encapsulated, static and has a structure that is designed for the purposes of business processes[2].

IV. METODOLOGY

In the method section explained the steps taken in this study. In this section will explain 9 steps which used in this research.

1) Choose the Process

Choosing a process means determining the main subject. The main subject refers to a business activity of a company that can answer all important business questions and has certain characteristics. Analyzing data and information needs is an important process to improve business running and as an evaluation for business owners. The following is the process we chose to analyze:
1. Analysis of Sales Trends and Business Benefits
2. Item Procurement Trend Analysis.

2) Choose the Grain

Choosing grain means determining what will be represented or presented by a fact table. After determining the grain from the fact table, then it can be determined dimension tables related to the fact table. In this case grain can be obtained from invoices issued to buyers and invoices obtained from purchasing to suppliers.

3) Identify And conform the dimensions

At this stage the design of a matrix of comparison between the grain and dimension tables is determined based on the design requirements of the data warehouse to be built.

Table 1 Determine Grain of Dimension Table

Grain	Dimension			
	Item	Time	Supplier	Cashier
Trends and Business Benefits	v	v	x	x
Item Procurement Trend	v	v	v	x

Table 1 displays the selection process the dimension tables and the grain will be created in the design of data warehouse in which there are four dimensions to be formed that is dimensions, Item, time, supplier and cashier.

4) Choose the Facts

The next process is to determine which table to the fact table where the fact to be formed on the basis of selection process grain that has been done. In this study there are two fact tables will be set up, namely the fact of trends business benefit and the fact of the item procurement.

5) Store pre-calculations in the fact table

When the facts have been selected, each fact must be re-examined to determine whether there is an opportunity to use pre-calculation. The following is one example of the initial calculation stored in the fact table.

Figure 1 Fact Table have been selected

Fact Table fact_trx_jual including keytime, keybarang, keyuser total_item_trx,total_nominal_harga,total_nominal_trx,total _laba_trx;

Fact Table fact_trx_beli including keytime, keybarang, keysupplier, total_item_trx, total_nominal_harga, total_nominal_trx;

6) Round Out the Dimension Tables

At this stage return to the dimension table and add as many text descriptions as possible to the dimensions. Text description must be understood by the user. In this stage, it includes information in the form of a List of Dimension

Table Descriptions along with a table, and a schematic drawing that has been designed.
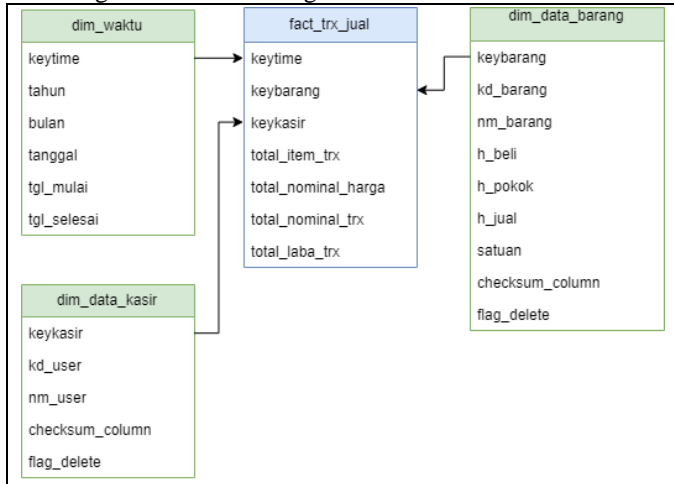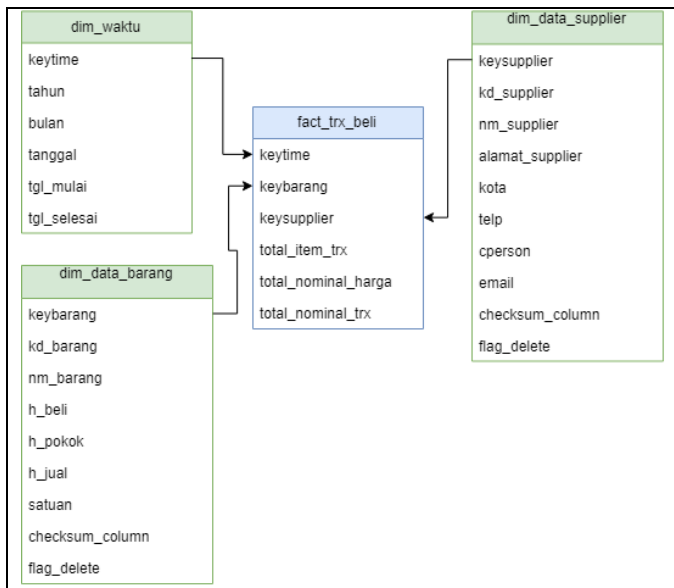

Figure 2 Fact Table fact_trx_jual with dimension

Figure 3 Fact Table fact_trx_beli with dimension

7) Decide the Duration of the database and periodicity of update

Determine the time limit of the age of data taken and will be moved into the fact table. The data used in this data warehouse design with a duration of 6 years starting from 2012 to 2018.

8) Track Slowly the changing dimensions

Attributes in the dimension table do not always have a fixed or static value. In overcoming and monitoring changes that occur in dimensions, the second type of three types of changes are described by Kimball and Ross (Kimball and Ross, 2002), namely if the dimensional attributes are changed it causes the creation of new dimension records with different surrogate keys.

Table changes dimensions are dim_data_supplier, dim_data_barang, dim_data_kasir.

9) Decide the query priorities and the query mode

In this stage ETL processing time (Extract, Transform, and Load) is carried out, periodic backups, and security of report use and analysis of storage media capacity and data growth, including the description of the perpetrator and the time period or process schedule.

V. RESULT AND DISCUSSION

In this section we will discuss the implementation of the design made. The implementation of the data warehouse design running on a web browser was built using PHP and using pivot table plugins. PHP functions as a service provider for data requests carried out by the client.
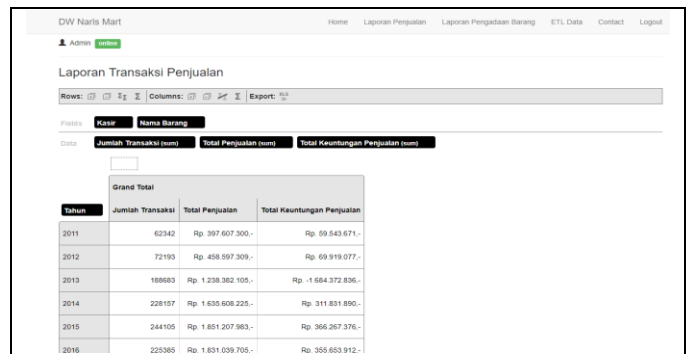

Figure 4 Report for Transaction and Benefit Income

Fig. 4 shows the report generated by the application. the report produced is a report that is used to analyze the business processes of sales and profit transactions.
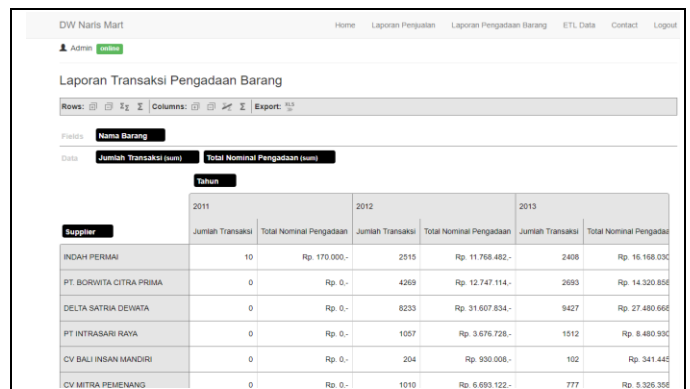

Figure 5 Reporting Item Procurement

Figure 5 shows the report generated by the application. the report produced is a report that is used to analyze the business process of procuring items against suppliers.

VI. CONCLUSION

From the implementation of this research conducted can be concluded as follows:

The data warehouse design that is made related to the business process of sales and procurement of goods transactions can be used as the basis for presenting integrated data as supporting data in the evaluation process and as the basis for the decision-making process for management.

The implementation of the data warehouse design that is

made can be implemented in the form of a report so as to facilitate the user in carrying out the analysis.

REFERENCES

[1] I. Purwanti, "Perancangan Data Warehouse pada PT. Olympindo Multi Finance Palembang Area Regional Sumatera II," p. 6.

[2] M. P. A. Ariawan, I. M. D. Ardiada, and Y. P. Sudarmojo, "Design of Library Data Warehouse Using OLTP Result of Services Analysis," vol. 3, no. 1, p. 4, 2018.

[3] H. Ramadhan and A. Soepriadi, "PENERAPAN MODEL KOMBINASI INMON DAN KIMBALL PADA PEMBANGUNAN ENTERPRISE DATA WAREHOUSE DAN BUSINESS INTELLIGENCE (EDW/BI)," p. 6, 2011.

[4] K. Khotimah, "PERANCANGAN DAN IMPLEMENTASI DATA WAREHOUSE UNTUK MENDUKUNG SISTEM AKADEMIK (STUDI KASUS PADA STKIP MUHAMMADIYAH KOTABUMI)," vol. 02, no. 01, p. 14, 2016.

[5] Y. S. Triana and A. Susilo, "Aplikasi Data Warehouse Untuk Menunjang Standar 3 Borang Akreditasi Prodi Informatika Universitas Mercu Buana," *J. Ilm. FIFO*, p. 8.

[6] J. Tunggono, M. R. Faisal, and D. T. Nugrahadi, "PEMANFAATAN DATA WAREHOUSE SEBAGAI SARANA PENUNJANG PENYUSUNAN BORANG AKREDITASI STANDAR 3 DAN STANDAR 4," vol. 02, p. 14, 2015.

[7] "System Usability Scale (SUS)," *usability.gov Improving the User Experience*. .

[8] A. Sergeev, "Efficiency," *User interfaces design and UX/usability evaluation*, 2010-1999. .

[9] J. U. Sari, "Perspektif Pengguna dalam Perancangan UserInterface pada Website Digital Library," vol. 13, p. 7, 2015.