Design of Library Data Warehouse Using OLTP Result of Services Analysis

Made Pasek Agus Ariawan[1*], I Made Dwi Ardiada[2], Yanu Prapto Sudarmojo[3]

[1,2]Department of Electrical and Computer Engineering, Post Graduate Program, Udayana University

[3]Department of Electrical and Computer Engineering, Udayana University

*Email: pasekagusariawan38@gmail.com

.

Abstract— **The library is the primary means used in development efforts as well as an increase in knowledge. Library holds a very large role in the spread of information because the library provides collections that can be used as a reference for the Civitas academic. Data Warehouse is a system that takes and consolidate data periodically from a source system into a dimensional or normalization of the data store. The purpose of this research was to analyze the information about the business process and create the initial means for taking quick decisions in related service quality evaluation in libraries and Understand the trends/interests and make forecasts better decisions in the evaluation of the quality of related services in the library.**

Index Terms --- data warehouse, OLTP, Pentaho.

I. INTRODUCTION

The library is the primary means used in development efforts as well as an increase in knowledge. The Library holds a very large role in the spread of information because the library provides collections that can be used as a reference for the Civitas academic [1].

Existence of the Library as a technical services unit in a College is extremely important to provide services against civitas Academica. With a growing number of books collection then the system databases are built should be able to improve services-oriented provision of data warehouse [2]. Especially at the managerial level require the full information, quickly and accurately to support processes and events planning, evaluation, and decision-making.

Database will store all the information in detail for any event or transaction performed. The data stored will be dynamic, always in accordance with the increased occurrence of transactions in the database. However, at the tingakatan of managerial needed in general is a result aggregation or summary of the transactional database. To be able to present the required data by managerial level, the existence of the data warehouse is very important [3].

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. Different Data Warehouse with data in OLTP (Online Transactional Processing) where data is stored in full [1].

The purpose of this research was to analyze the information about the business process and create the initial means for taking quick decisions in related service quality evaluation in libraries and Understand the trends/interests and make forecasts better decisions in related service quality evaluation in libraries.

II. OVERVIEW OF THE STUDY

A. Data Warehouse

Data Warehouse is a system that takes and consolidate data periodically from a source system into a dimensional or normalization of the data store. Typically contains history and query for business intelligence or other analysis activities. This is usually updated in batch, not every transaction that occurs in the source system [4]. Data Warehouse is a method in the design of the database, which support the DSS (Decission Support System) and EIS (Executive Information System). Physical data warehouse is a database, but the design of the data warehouse and database are very different. In the traditional database design using normalization, but on data warehouse normalization is not the best way [5].

Data warehouse is a system that takes and aggregate the data periodically from a data source into a dimension. Data warehouse can also be interpreted as a set of methods, techniques and tools used to support senior managers, directors, managers, and analysts to do data analysis that can assist in the process of making decisions and increase source of information[6].

Data warehouse is data that a very large that have a nature-oriented subject, integrated, time-variant, and is fixed on the stored data in support of decision-making process management, which can improve the accuracy and quality of information that is easily understandable and can be accesed easily by user containing desired information[7]

B. Process Extract, Transform and Load (ETL)

According to [8] 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[9].

C. Snowflake Schema

Snowflake schema is one method of data modeling for data warehouse which is a development of the star schema, namely the method of normalization to dimension table that exists in the star schema. According to [10] snowflake schema is a combination of several data modeling star schema normalized. In a snowflake schema, each dimension table can have sub-table dimensions again. This aims to minimize redundant data [5]. Dimension data this was the subject of the information to be material in decision-making, because on every dimension of the data it is possible to do more detailed breakdown again.

D. Pentaho

Pentaho Kettle is software (software) open source released by Pentaho. Applications developed by Pentaho Pentaho Corp. based in Orlando United States. Kettle Pentaho software can be used as tools for integrating data. Pentaho Kettle provides the facilities to conduct the process of ETL [1].

The main elements of the Pentaho Kettle is a Transformation and Job. Transformation is a set of instructions to transform the inputs into the desired output. While the Job is a set of instructions to run the transformation. There are three main components in Pentaho Kettle that is, Pan, Spoon and Kitchen. Spoon is the user interface to make Job and Tranformation. Pan is the tools that function read, modify and write data, and the Kitchen is a program that executes the job.

E. Design Data Warehouse

One of the data warehouse design methodology is a Four-Step Methodology by [7] [10] [11]which has four (4) steps.

1. Selection of Business Process

    The process is a natural business processes are performed within an organization that is usually supported by a system of data collection sources. The simplest way in choosing business processes is by interviewing users or stakeholders.

2. Choosing The Grain.

    Choose grain means deciding exactly what is stated by the record from the table facts. Only with the election of grain for the fact table so we can identify the dimensions. The decision of the grain for grain also determined the fact table for each dimension in the table facts.

3. Choose the dimensions

    Dimensions determine the context to provide questions on the facts – facts in the fact table. A collection of dimension that well constructed makes the data mart is understandable and easy to use. Dimensions identified in sufficient detail to describe the data as properties of the client and the right grain.

4. Identifying Facts Table

    Grain from the fact table to determine which facts can be used on data mart. All the facts should be stated based on the levels implied by the grain. Additional facts can be added to the fact table in each time with a note that fact is consistent with the grain of the table.

III. METODOLOGY

A. Analysis OLTP

The data used in this research is database OLTP library of the year 2009-2017 where the analysis is done by observing the parameters as follows.

   a) change the null values in the field.

   b) Select the data to be used in dimensional modeling

   c) Makes the new code

   d) Break down the data in a table

B. Choose Process

The process of the library that is used to design the data warehouse is the process of borrowing. The process of lending is meant is loaning books. As for the supporting data such as data books, data transaction, the data of visitors.

C. Select the Grain

Grain on a library that is used to design the data warehouse is loaning. Where the analysis is done on a loan is the number of books borrowed, types of books are often borrowed, the amount of visits. The analysis will be done per time period (month, year).

D. Determining Grain and table Dimensions

TABLE I. Determination of the grain and the table dimensions.

Grain / Dimension	The amount of the loan	Types of books often borrowed	The number of visits
Book		X	
Transaction	X		
Time	X	X	X
Visitors			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, transaction books, time and visitors.

E. Specify the table 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 borrowing and the fact of the visit

F. dimensional Modeling

This research uses modeling a snowflake consists of six tables of dimensions and fact tables for two. Table dimensions are composed of dim_visitors, dim_time, dim_transactions, dim_databooks, dim_ title, dim_typeofbook. Then to table the

fact is fact visit and fact lending. figure 1 dimensional modeling of the data warehouse is designed.
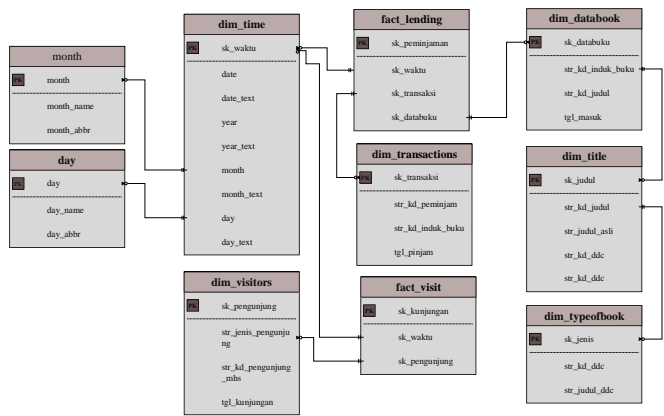

Fig 1. Dimensional modeling

G. ETL Process

In this study there are two processes ETL for the fact table. Figure 2 is the process of borrowing and ETL Figure 3 is the process of ETL from the fact of the visit.
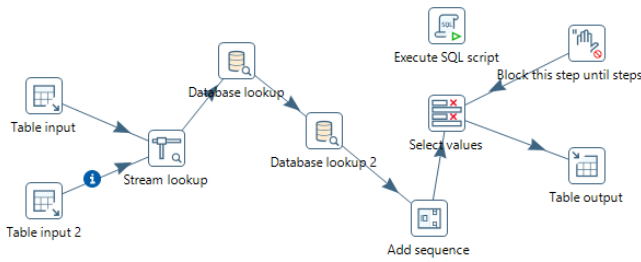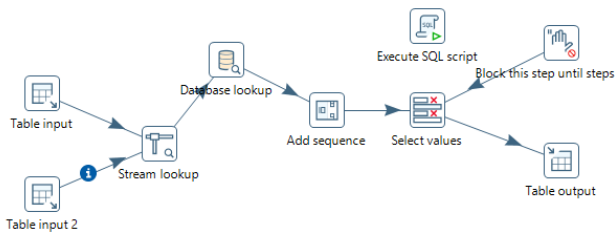

Fig 2. ETL process loan facts

Fig 3. ETL process fact visits

IV. RESULTS AND DISCUSSION

Following is the result from the design of the data warehouse has been done.

A. amount of borrowing books

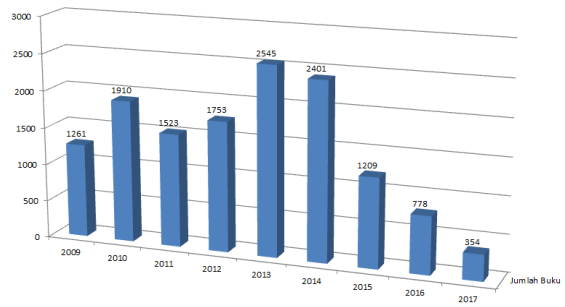Figure 4 is a graph of the number of borrowing books from the year 2009 to 2017.

Fig 4. The graph of the total loan book.

B. Total collection of books for every kind.

Figure 5 displays the total book belonging to perpuastakan from the year 2009 to 2017 are distinguished from each of its kind.
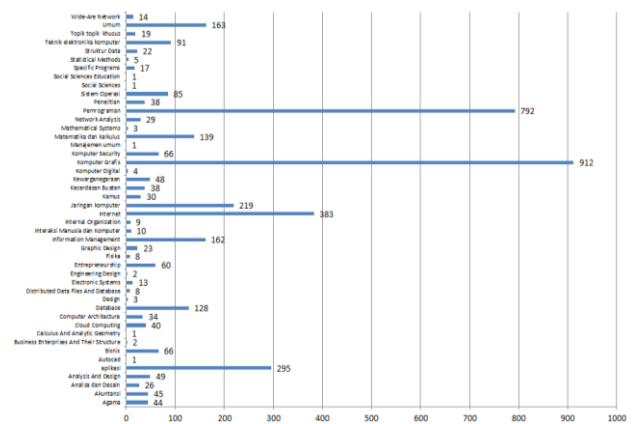

Fig 5. Graph of total library book collection

C. Types of books often borrowed

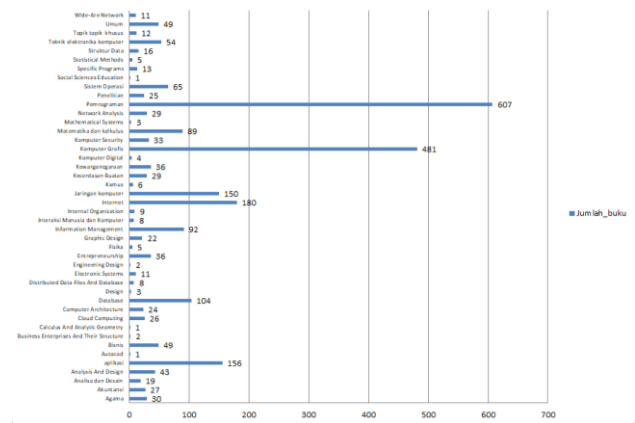Figure 6 displays the types of books are often borrowed from the year 2009 to 2017.

Fig 6. Graph types of books often borrowed.

D. Number of visitors

Figure 7 displays the graph of the number of visitors to the library either from the lecturers and students of the year 2009-2017
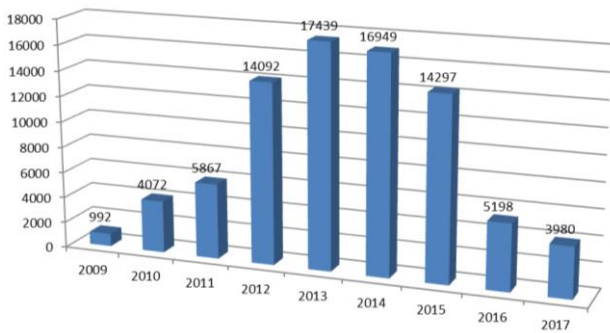

Fig 7. Graph of the number of visitors

E. Analysis of the number of visitors and the amount of borrowing books

Figure 8 displays the rate of the number of visitors and the amount of borrowing books.
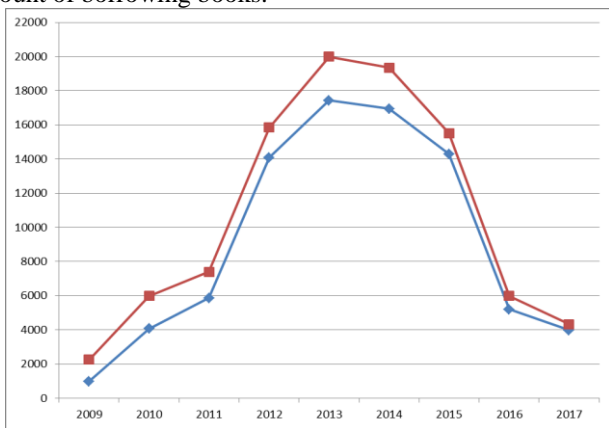

Fig 8. The graph of the number of visitors and the rate of borrowing books

in Figure 8 can be seen the rate declining graph starting from the year 2013 to 2017 for either borrowing or visit so that from these observations referable to steakholder to conduct an evaluation of the service quality in libraries.

V. CONCLUSION

Based on the results obtained in this research it can be concluded that the design of the data warehouse library can provide solutions related to the means used by steakholder to conduct evaluation related services diperpustakaan.

VI. BIBLIOGRAPHY

[1]   Andri dan B. Tujni, "Perancangan data warehouse perpustakaan," *J. Ilm. MATRIK*, vol. 19, no. 2, hal. 101–110, 2017.

[2]   M. A. Machmudi, "Rancang Bangun Data Warehouse Perpustakaan untuk Perguruan Tinggi," *Transform. J. Inf. Pengemb. Iptek*, vol. 11, no. 02, hal. 136–142, 2015.

[3]   A. Dahlan, E. Utami, dan E. T. Luthfi, "Perencanagan Data Warehouse Perpustakaan Perguruan Tinggi Xyz Menggunakan," hal. 1–14.

[4]   E. Turban, J. E. Aronson, dan T.-P. Liang, *Decision Support Systems and Business Intelligence*. 2007.

[5]   V. Rainardi, *Building s Data Warehouse with Examples in SQL Server*. 2008.

[6]   L. Izhariman, E. Darwiyanto, dan S. Y. Puspitasari, "Perancangan Executive Information System Rumah Sakit Untuk Evaluasi Kinerja Pelayanan Di Rsup Mohammad Hoesin Palembang Designing Hospital Executive Information System For Service Performace Evaluation At Rsup Mohammad Hoesin Palembang," in *e-Proceeding of Engineering*, 2017, vol. 4, no. 3, hal. 4957–4965.

[7]   P. Widiadnyana, M. A. Maricar, I. N. Arnawan, dan S. Ariyani, "Analysis of Data Warehouse for Transactional Analysis Methods Online Analytical Processing (OLAP) at Company XYZ Silver," *Int. J. Eng. Emerg. Technol.*, vol. 2, no. 1, hal. 72–75, 2017.

[8]   B. Thomas, Connolly, Carolyn, *Overview of SQL*. 2014.

[9]   A. B. Winnetou, S. A. Wicaksono, dan A. Pinandito, "Analisis Peningkatan Performa Proses ETL (Extract , Transform , Dan Loading) Pada Data Warehouse Dengan Menerapkan Delta Extraction Menggunakan Historical Table," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 4, hal. 1366–1371, 2018.

[10] W. H. Inmon, *Building the data warehouse*. 2002.

[11] K. P. B. Ary, G. B. Subiksa, dan M. Sudarma, "Design Data Warehouse For Centralized Medical Record," *Int. J. Eng. Emerg. Technol.*, vol. 2, no. 2, hal. 47–51, 2017.