



DATAWAREHOUSE SALES AND SUPPLY OF GOODS MODEL BASED ON HTML5

MICHAEL YOSEPH RICKY

School of Computer Science,
Bina Nusantara University,
Jakarta-Indonesia
Email: mricky@binus.edu

ABSTRACT

The purpose of this study was to conduct an analysis of sales and inventory systems that exist in the company, then designing a data warehouse for PT. Delta Djakarta, Tbk. to support the decision-making process in estimating the number of sales in the future, then found the existing problems and provide solutions to the company so that the company can overcome the existing problems and provide the best service to its customers. The method used in this research is a nine step methodology that is a model of systematic data warehouse in building data warehouse with through the stages of choosing the process, choosing the grain, identifying and confirming the dimension, choosing the facts, storing pre-calculation in the fact table, rounding out the dimension table, choosing the duration of the database, tracking slowly changing, deciding the query priorities and the query modes. The results of this study are the form of websites using HTML5 and data a warehouse information extraction model that provide sales and inventory companies and an application that can be used for all trading companies that sell products finished material. The conclusions obtained from this research is a data warehouse models and applications that can help executives (stakeholders) in obtaining the multidimensional report, where the report can be directly obtained by performing customized according to the wishes of stakeholders as a decision maker.

Keywords: *Data Warehouse, Nine Step Methodology, HTML5*

1. INTRODUCTION

Today's intense competition led to the need for the use of information technology in the enterprise is increasing. Today many companies have invested funds for information technology for the betterment of their business. The use of technology-based applications in enterprise information can vary in scope. With the information technology can produce a database. If the data has been integrated in the database, the data search time can be shortened, thus increasing the performance of the company.

To improve competitiveness in a globalized world, companies must be able to make decisions quickly and accurately to determine the proper steps in running the business. Companies that have implemented technology -based information database to store its data in the long run will certainly have the data with a very large number. The data that has been stored in the database can be processed in such a way that can be generated DTA is useful for further analysis. The result of such processing is called a data warehouse.

PT. Delta Djakarta Tbk is a company engaged in the production of carbonated beverages. The company doing millions of transactions per month in area of inventory and sales of products based on the transaction data. These data are entered into a computer to be analyzed and processed by the part of management to be useful information for the company. The problem faced by the company PT. Delta Djakarta Tbk is making the final report of sales and inventory is still using manual systems and un-structured data. With the amount of information generated, it takes a quick access to information for executives to make decisions. Obtaining and display information residing in enterprise data quickly and accurately then be made to a data warehouse.

Data warehouse is the heart of the architect and the environment as the basis of all the decision-making system [6], the data warehouse is more easily understood by the customer visits their needs in assessing the financial situation, products and services, as well as time to market and new services. Most importantly the data warehouse can be seen as a technology, with a unique ability to



process information, aiming to generate high revenue and increase profits.

Development of a data warehouse includes development of systems to extract data from the system and installation of warehouse database system that allows managers to access data flexibly. In general, the term refers to the data warehouse database different collected from all parts of the company.

2. DATAWAREHOUSE

Data warehouse is the combined technologies aimed at effective integration into the relational database environment that enables the strategic use of data [6].

The data warehouse is a collection of data that have the nature of the subject-oriented, integrated, time span and data collection has not changed in supporting management decision-making process [1].

Based on the above notions, it can be concluded that the data warehouse is a collection of data or records in a computer system that logically relate to each other with each other for decision making at the managerial level.

2.1 Purposes of Designing Data Warehouse

The design of data warehouse aims to allow organizations to use the data set so that the organization can help organizations to gain business advantage [1]. Data warehouse attributes include [6]:

2.1.1 Subject Oriented

Oriented data warehouse on the major subject areas of the company that has been defined in the data model of a high-level enterprise.

Data warehouses are arranged by subjects in a company, in which each subject area is physically implemented as a collection of related tables in the data warehouse and not oriented to a specific process or application functions. In accessing user data must be oriented in a particular subject.

2.1.2 Integrated

The data in the data warehouse from various sources of data collected from the whole system of different enterprise applications. Data sources are often inconsistent. Therefore, an

integrated data source must be made consistent to ensure the accuracy of the information generated.

2.1.3 Time Variant

Each unit of data in the data warehouse is only valid and accurate at a certain time or a certain time interval. To view the time interval used to measure the accuracy of a data warehouse.

2.1.4 Non-volatile

The data in the data warehouse is not updated in real time, but is refreshed from operational systems on a regular basis. The new data is added to the database as a complement rather than a substitute, so that historical data remains stored in the database.

2.2 Data warehouse Structure

There are several levels of detail in the data warehouse environment. Is categorized into four levels, namely Older Detail Level, Current Level Detail, Lightly and Highly Summarized Data Level Data Level. Initially the data flow occurs from the operational environment to the data warehouse. In this data flow transforms process occurs. The flow of data in the data warehouse has been in the level of detail. Over time, the data from the Current Level Details flows toward Older Detail Level. Summarize the event, and then the data will be switching from toward Lightly Current Level Detail Level Summarized data which will then be heading Highly Summarized Data Level [2].

2.3 Data warehouse Architecture

Identify the components contained in the data warehouse data warehouse architecture [1]:

2.3.1 Operational Data

Operational data focuses on transactional functions. This data is part of the enterprise infrastructure, detail, there is no redundancy (repetitive data), can be updated (be changed), and these data reflect the current value.

2.3.2 Operational Data Store (ODS)

ODS is a temporary storage of operational data currently used for the integrated analysis. Building the ODS can be a useful step in building a data warehouse for an ODS can supply the data that has been extracted from the source system and cleaned. This means integrating work and structuring of data to the data warehouse becomes simpler.

2.3.3 Load Manager

Load Manager displays all the operations associated with the extraction and loading of data into the data warehouse. Data can be extracted directly from the data source or in general of ODS.



2.3.4 Warehouse Manager

Warehouse Manager displays all the operations related to the management of data from the data warehouse. Operation shown by the warehouse manager includes analysis of data to ensure consistency, changes and the incorporation of data sources from a temporary storage area for data warehouse tables, index creation and appearance of the table base, making de-normalization (if required), making aggregation (if required), creating a data backup and archive data.

2.3.5 Query Manager

Query Manager displays all the operations related to the management of user queries. Operation displayed by this component includes directing queries to the appropriate table and schedule the execution of the query.

2.3.6 Detailed Data

This component stores all the details of the data in the database schema. Details of data are divided into two are:

2.3.6.1 Current details of the data

This data comes directly from the operational databases, and always refers to the current company data. Current details of the data set along the sides of the subject such as customer profile data, customer activity data, sales data, demographic data, and others.

2.3.6.2 Old details of Data

These data show detail data that was current or history of the subject area. These data were used to analyze the trends that will be generated.

2.3.7 Lightly and Highly Summarized the Data

Area data warehouse stores all the data lightly and highly summarized that have been defined previously made by the warehouse manager. The purpose of this information is summarized improve query performance.

2.3.8 Archive / Backup Data

The warehouse area store detailed data and summarized data archiving purposes and do backup data.

2.3.9 Metadata

Metadata is data about data that describes the data warehouse. Metadata are used to build, maintain, organize and use the data warehouse. Metadata contains the location and description of the components of the data warehouse; names, definitions, structure, and content of the data warehouse and the end user view; identification of the manufacturer's data sources (record system); integration rules and transformations that are used to populate history of updating the data warehouse

and warehouse refresh data; matrix patterns are used to analyze the performance of the data warehouse, and so on.

2.3.10 End-User Access Tool

This tool includes reporting and query tool, application development tool, executive information system (EIS) tool, online analytical processing (OLAP) tools, and data mining tools.

3. DIMENSIONALITY MODELING

Dimensionality Modeling is a logical design technique that aims to present the data in a standard form and intuitive which allows accessing the database with high performance. [7]

3.1 Fact Table

Fact table is a table where the center of a star schema data that often appears to be placed on the table. Fact table is called the primary table (major table), is at the core of a star schema and contains the actual data to be analyzed (quantitative data and transactions). Fact table is a table that generally contains numbers and historical data where the key generated very unique because it is a collection of primary keys and foreign keys that exist in each of the related dimension table or a centralized table of a star schema. Store the fact table type's different measure, such measure, which is directly connected to the table and measure dimensions that are not related to the dimension table [5].

Table dimensions are where the data is not associated with the related fact tables placed in a multidimensional table. Table dimensions are also called small table (minor table), because usually smaller and hold descriptive data that reflect the dimensions of a business. Dimension table is a table that contains a category with a summary of the detailed data can be reported, such as benefit statements can be reported on the fact table as a dimension of time (such as per month or per year) [6].

3.2 Star Schema

Star schema is a logical structure that has a fact table in the middle, which consists of factual data, and surrounded by dimension tables containing reference data. The scheme exploits the characteristics of factual data such as the fact that generated by the events that turned up on time and does not change the past. Star schema can be used to accelerate the performance of queries with information de-normalization reference to a single dimension table [1].



3.3 ETL (Extraction, Transformation, Loading)

ETL is the process of searching the data, integrating, and put the data into a data warehouse [3]. This process consists of three stages are:

3.3.1 Extraction

The first step of the ETL process is the process of retrieving data from one or more operating systems as a data source (usually taken from the OLTP system, but can also be a source of data outside of the database system). Most data warehouse projects combine data from different sources. In essence, the process of extraction is the process of disassembly and cleaning of the extracted data to get a pattern or structure of the desired data.

3.3.2 Transformation

The process of cleaning up the data has been taken on the process of extracting the data so that it conforms to the structure of the data warehouse or data mart.

3.3.3 Loading

This is the final step in the ETL process. The process of entering data into the end target, in this case is a data warehouse or data mart. Data derived from the transformation process. Once the data generated from the transformation process in accordance with the desired conditions in a data warehouse or data mart, the loading process will run. The data from the staging area will be moved into the data warehouse or data mart.

4. PROPOSED METHOD

4.1 The Framework

Creating a framework that contains data warehouse model sales and supply of goods. In practice, this study was conducted with the following methodology:

4.1.1 Method of Analysis

In this study, the method of analysis used Fact Finding Technique is a method that includes:

- Field Survey

This method is used to get real writing materials on the company in question, conduct research and interviews with the heads of the warehouse with the company's data warehouse application design. The results of field surveys, among others, information about the company's background, organizational structure, duties, authority of each piece, the systems that are running, the database is used, and information about the problems faced by companies today.

4.1.2 Analysis of Survey Results

This method is used to address and provide solutions to related problems that occur within the company. The analysis used is the analysis of the data and analysis needs and information.

4.1.3 Design Methods

In this research report, according to the design method [6] there are nine stages in building a data warehouse using nine-step methodology.

4.1.4 Choosing the Process

At this stage an election process that refers to the subject matter required by the data mart. Data mart is a part of the data warehouse that supports report generation and analysis of data on a unit, section or operations in a company.

4.1.5 Choosing the Grain

The selection process to determine the grain is what is described by the records in the fact table. Fact table is a table that contains numbers and history data which are generated key is unique because it is a collection of primary keys and foreign keys that exist in each of the related dimension tables, while the dimension table is a table that contains categories with detailed summaries of data that can be reported, as gains in the fact table as a dimension of time (per month, per quarter, per year). At this stage determine the identification of a fact table and dimensions of the fact table.

4.1.6 Identifying and confirming the dimension

At this stage make a set of dimensions required to answer all questions posed by the fact table. Set the dimensions of the data mart makes it easier to understand and use.

4.1.7 Choosing the facts

At this stage select the fact table that can imply all the grain that is used in the data mart.

4.1.8 Storing Pre-calculation in the Fact Table

At this stage after the fact table chosen, each fact table should be re-examined to determine whether there is a pre- specified calculation.

4.1.9 Rounding Out the Dimension Table

At this stage, re-examination of the dimension table and add a text description that describes the dimension tables, as well as determine the dimension attribute hierarchy to simplify the analysis process.

4.1.10 Choosing the Duration of the Database

At this stage the database to determine the period of time a few years back.

4.1.11 Track Changes of Dimension Slowly

At this stage the observed dimensions of the dimension tables. There are three basic types of dimensional changes slowly are rewritten dimension attribute changes, dimensional attribute changes resulted in the creation of a new dimension, dimensional attribute changes resulted in an alternative attribute is made, so between the old and new attributes accessed together.

4.1.12 Deciding the Query Priorities and the Query Mode

At this stage, create a physical design consideration. Physical database design issues are critically affecting the perception of the end user to the data mart.

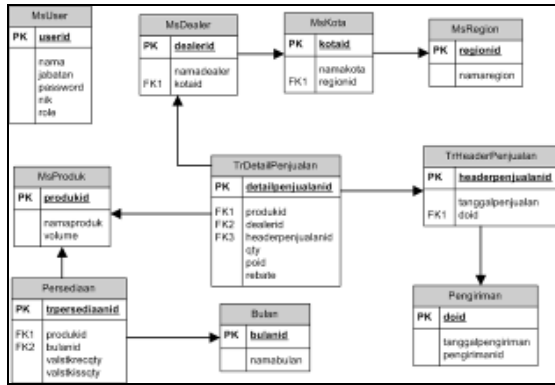


Figure 1: Entity Relationship Diagram

5. EXPERIMENTAL RESULT

We have questionnaire for user and use several questions for this experiment. Experiments performed on the website without lag and give the expected result.

The method used in the drafting of the proposed data warehouse at PT. Delta Djakarta Tbk is a nine-step methodology.

5.1 Determining Process

Determine the business processes that will be used in making the data warehouse. These processes are obtained from the analysis of the running system and required by the company. Business processes that have been defined are stock and sale. The processes above were determined from the results of the survey at the request of the company and the company.

5.2 Choosing the Grain

The next step is to determine the grain to the dimension tables and fact tables that will be

used in the data warehouse. Grain for the fact table are inventory fact (FactPersediaan), this fact shows that there are changes in product inventory in the company; sales fact (FactPenjualan), this fact contains product sales transactions by the company to the dealer.

5.3 The Dimensions of the Existing Identification

After the selection of the grain needs to be done next is to identify dimensions associated with a fact. Dimensions that exist in this process is the time of sale, time inventory, shipping, months, dealers, product, region and city.

5.4 Choose the Facts

If already determine the dimensions involved. The next step is to ensure that the fact table can be obtained in the process of grain. Facts on this company, among others, are the fact of sale, delivery, and inventory.

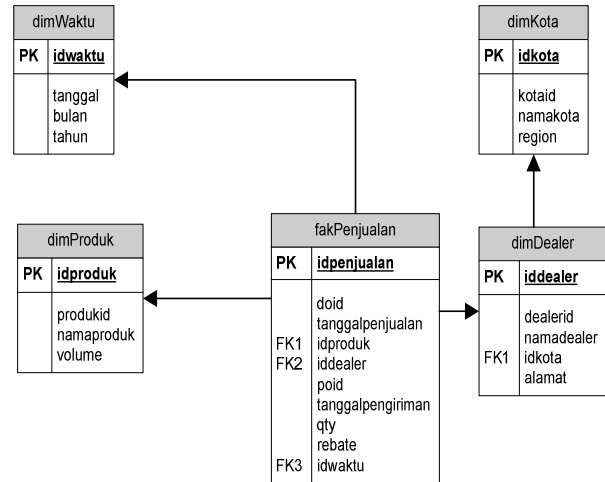


Figure 2: Scheme of the sales fact

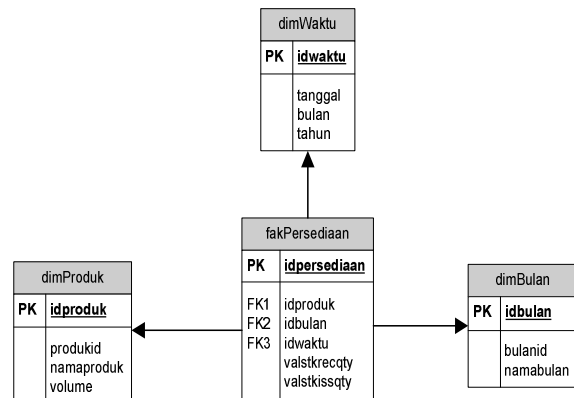




Figure 3: Scheme of the supply of goods fact

The dimensions tables are:

Table 1: DimWaktu

Attribute	Data Type	Length
IdWaktu	Int	11
Tanggal	Datetime	
Bulan	Int	11
Tahun	Int	11

Table 2: DimProduk

Attribute	Data Type	Length
IdProduk	Int	11
ProdukId	Char	4
NamaProduk	Varchar	100
Volume	Int	11

Table 3: DimDealer

Attribute	Data Type	Length
IdDealer	Int	11
DealerId	Char	6
NamaDealer	Varchar	255
AlamatDealer	Varchar	255
IdKota	Int	3

Table 4: DimKota

Attribute	Data Type	Length
IdKota	Int	11
KotaId	Char	3
NamaKota	Varchar	25
Region	Char	255

Table 5 : DimBulan

Attribute	Data Type	Length
IdBulan	Int	11
BulanId	Int	11
NamaBulan	Varchar	25

5.5 Choose the Duration of the Database

At this stage the specified duration of the data to be stored in the data warehouse. In practice, the duration of the database specified in the data warehouse for PT. Delta Djakarta is over the next five years. This is done to increase the dynamism of the analysis of the data warehouse is done.

Data warehouse that will be made have the duration of the data for five years from 2006 until 2010. Database used data drawn from the data is moved into the OLTP and OLAP data through Transform System (DTS). This results in the fact tables and dimension tables to prepare sales reports and inventory reports.

5.6 Track Slowly Change the Dimension Table

At this stage, in the process of updating the records in the dimension tables that have been defined for the data that is modified by the OLTP system.

5.7 Specify the Query Priorities and Query Modes

In this step, predicted to make the physical design. The most important thing in the physical design affects the perception of end-user data mart is a sequence of commands that exist in the physical and the availability of the summary table. Behind all of this there are additional physical design that affects the administration, backups, indexing performance, and security.

5.8 Backup and Recovery

This component is where backing up the data from the data warehouse has been completed transformed. With the backup of data, if there is damage to the data, the company still has data backup. PT. Delta Djakarta performs data backup so complete daily work hours. Backup (data storage) is stored in the external CD.

Data Recovery aims to restore the data to the last state before damage occurs. So if there is a

problem that eliminates the data. Data can be recovered from the CD.

5.9 Web Application Prototype Data Warehouse



Figure 4: Data warehouse Report Screen

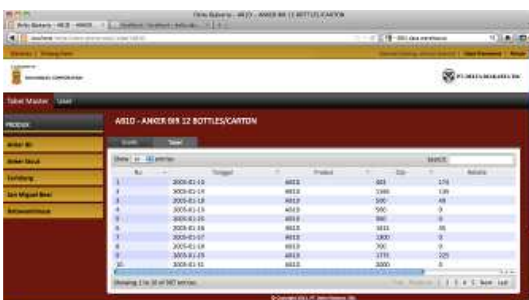


Figure 5: Data warehouse Table Screen

6. CONCLUSION

Based on the analysis and research conducted, it can be concluded as follows : database operating contain data that are detailed and transformed into a data warehouse, through a nine step methodology produced two fact table is a fact table sales (FactPenjualan) and inventory fact tables (FactPersediaan) and the dimension tables are time dimension table (DimWaktu), product dimension table (DimProduk), dealer dimension table (DimDealer), city dimension table (DimKota), month dimensions table (DimBulan).

In order to obtain an integrated data, historical character, up-to-date, interactive, and can be manipulated according to the needs of the executive (stakeholders), data more summarized compared with operational data there previously. The resulting data warehouse applications can help the executive (stakeholders) in obtaining the multidimensional report, where the report can be directly obtained by direct customized according to the desire of the executive (stakeholders) as a decision maker.

Some suggestions are proposed for the development of a data warehouse to fit the needs of information are application of sales and supply of goods model of data warehouse which has been designed to be further expanded its use into the application of data mining, executive information systems, and decision support system. In order to use the data warehouse is not only in the field of sales and inventory only, because the actual data warehouse can be further developed in various fields that are useful for analysis, such as the production, distribution, and financial fields. In anticipation of expected growth data and the greater the increase, the need for computer data storage media need to be considered.

ACKNOWLEDGMENT

The author thanks to Bina Nusantara University. This work was supported in part by a grant from Directorate of Research, Bina Nusantara University.

REFERENCES:

- [1] Connolly, T.M. and C.E.Begg. (2002). Database System:A Practical Approach to Design,Implementation, and Management.(3th edition). USA: Addison Wesley.
- [2] Data Warehouse. (2008). Data Warehouse. Retrieved 14-20-2012 from <http://datawarehouse4u.info>
- [3] Gustiarahman, Irfan. (2006). Data Warehouse. Retrieved 14-10-2011 from <http://myhut.org/public/datawarehouse.doc>
- [4] Heizer, Jay and B.Render. (2001). Operations Management. (6th edition). New Jersey: Prentice Hall.
- [5] Hoffer, J.A., M.B.Prescott, and F.R.McFadden. (2005). Modern Database Management. (7th edition). New Jersey: Prentice Hall.
- [6] Inmon, W.H. (2002). Building the Data Warehouse.(3th edition). New York: John Wiley.
- [7] Kimball, Ralph and M.Ross. (2002). The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling. (2nd edition). New York: John Wiley & Sons