Jurnal Ilmiah Komputer dan Informatika (KOMPUTA)
Edisi...Volume.... Bulan 20..ISSN :2089-9033

SOFTWARE DEVELOPMENT DATA MART IN GP SHOES

Roni Sulaeman (10109455)

Teknik Informatika – Universitas Komputer Indonesia
Jl. Dipatiukur 112-114 Bandung
Email : ronisulaeman85@yahoo.com

ABSTRACK

GP SHOES is a private company whose main activity producing men's shoes and women's shoes.The company is in the trust as a supplier for other companies. Currently in presenting a strategic update of production or production stratragis reports required by the production manager is still manual and requires a long time in the search process data. Surely it would hamper decision-making in production by the production manager. The result would be a loss for the company.

Based on the existing problems in the production division in GP SHOES, hence the need to build a software data marts to facilitate the decision making the production manager to obtain strategic information quickly so that it can be used long-term planning. Data Mart can provide information quickly, easy and detail used for data analysis and report generation can also support the production of strategic briefly or have a period of time in the form of tables and graphs. Development of a data mart application system using SSIS (SQL Server Integration Service) and for the construction of a data mart using object analysis.

Based on the results of blackbox testing and beta, it could be concluded that the data mart is able to present a strategic information quickly and succinctly in analyzing strategic information, and can facilitate in making the report strategic information needed by the production manager at GP SHOES.

Keyword : *Data Mart , **Constellation Skema**, SSIS , OLTP, ETL, OLAP.*

1. *Introduction*

GP SHOES is a private company whose main activity is producing shoes for men's and women's

the company was incorporated in 1898 and located in Jl. Gunungpuntang km 28, Kp. Kebontunggul RT 03 RW 03 Ds. Campakamulya Kec.Cimaung Kab. Bandung. Products manufactured by GP SHOES has a high quality and economical prices in order to meet market demand.

GP SHOES did a lot of production is generated every day. With so many manufactured products, then the data from the production of more and more. Data output from the project entered into the existing system to be analyzed by the production manager and serve as useful information for the company, the company is currently experiencing problems in presenting a production of strategic information or reports required by the production stratrgis production manager.

In the presentation of the information is still done manually and requires a long time in the search process data. Surely it would hamper decision-making in production by the production manager. The result would be a loss for the company.

Investigation on the problems of the production at the GP'S SHOES, it is necessary to build a data mart software to facilitate in decision-making parties production manager to get information quickly that strategic planning can be used long term. Data mart can present information in a quick, easy and detail that is used to analyze the data and can also support making the strategic production report in summary or have a period of time in the form of tables and graphs.

Data Mart is part of the data warehouse, which is a collection of data subject oriented, integrated, have a period of time, and can not be updated and can support the production manager for decision making, and can assist in making the final report and analysis of data on production division [1].

1.1 Data Warehouse

The data warehouse is a subject-oriented data set, integrated, can not be updated, has the dimension of time, which is used to support management decision-making processes and business intelligence [4]. The data warehouse has characteristics, as follows :

1. *Subject Oriented*

*Data warehouse subject oriented data warehouse means designed to analyze data based on certain subjects within the organization, rather than on the particular application or function. The data warehouse is organized around the main subjects of the company (customers, products and sales), this is because the needs of the data warehouse for storing data that is supporting a decision.*Terintegrasi (*Integrated)*

Data Warehouse can store data coming from separate sources into a format that is consistent and integrated with each other.Rentang Waktu (Time-variant)

all data in the data warehouse can be said to be accurate or valid at any given time. To view the time interval used to measure the accuracy of a data warehouse, , We can use the way include:

- *The simplest way is to present the data warehouse at a certain time range, such as between 5 to 10 years into the future.*
- *The second way, using variations / differences in time are included in data warehouse either implicitly or explicitly, an explicit the element of time in days, weeks, months and others. Implicitly for example, when the data is duplicated at each end of the month, or quarterly. The element of time will remain implicit in the data.*
- *The third way, the time variation presented data warehouse through a long series of snapshots. Snapshot is a view of a specific portion of the data corresponding user desires of all the data that is read-only.*

2. *Non-Volatile*

Data warehouses can not be updated in real time but on-referesh of the operating system on a regular basis. The new data are being added to the database itself. The database is constantly receiving and storing new data, and then combined with previous data.

1.2 Data Mart

Data Mart is part of the data warehouse that supports the creation of report and analysis on a unit, section or operational at a persahaan. Data mart are often used to provide information to the functional segments organisasia. [1].

There are four tasks that can be performed by the data mart, four tasks are as follows:

1. *Preparing report*

   Preparing report is one of the data mart to the most commonly used. By using a simple query obtained reports per day, per month, per year, or whenever desired time period.
2. *On-Line Analytical Processing (OLAP)*

   With the data mart, all the information both detail and summary results needed in the analysis of easily obtained. OLAP is a concept of multidimensional data and allows the user to analyze the data in detail, without typing any SQL commands. Another facility is the roll-up and drill-down. Drill-down is the ability to see details of the information and the roll-up is just the opposite.
3. *Data Mining*

   Data mining is the process of digging (mining) knowledge and new information from a large number of data in the data mart.
4. *The process of executive information*

Data marts can make a summary of important information with the goal of making business decisions, without the need to explore the entire data. By using a data mart of all reports have been summarized and can also find out all the details in full, thus simplifying the decision-making process.

1.2.1 Model Dimensionaling

Dimensional model of the data mart consists of the fact tables and dimension tables. Fact table is a table that contains a collection of primary key foreign key contained in each dimension table, while the dimension table is a table that contains detailed data that describes a foreign key contained in the fact table.

There are several models of the scheme contained in the modeling data marts, the star schema, snowflake schema, and constellation schemes. Explanation of each model are as follows :

1. *Star Schema*

This scheme follows the shape of a star, where there is one fact table in the center of a star with several surrounding dimension tables. All associated with the dimension tables to the fact table. The fact table has several primary key in the dimension table. Here is an example of a star schema can be seen in Image 1.1.

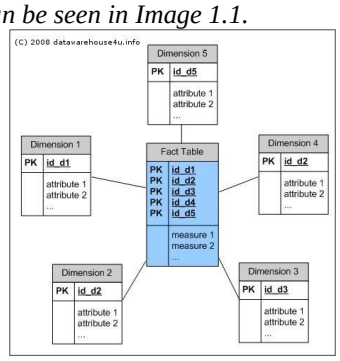


Image 1.1 Star Schema

2. *Snowflake Schema*

According to Connolly and Begg [1], Snowflake Schema is an extension of a star schema with an additional dimension tables that are not

directly related to the fact tables. The dimension tables associated with another dimension tables. Here is an example of the snowball scheme can be seen in Image 1.2.

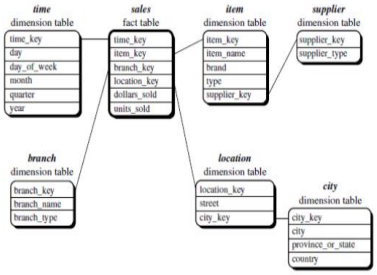


Image 1.2 Snowflake schema

3. *Skema Constellation*

Constellation scheme is a multidimensional schema that contains more than one table to the fact that sharing table dimensions.
Here is an example constellation scheme can be seen in Image 1.3.

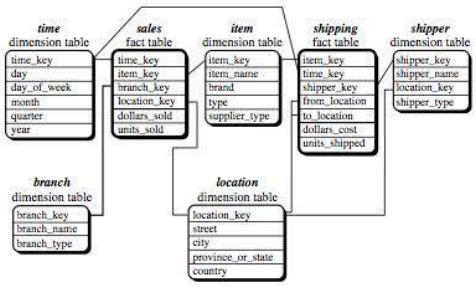


Image 1.3 Skema Constellation

1.2.2 Troubleshooting Data Mart

Troubleshooting methods that are used in the manufacture of a data mart on GP SHOES are as follows:



Image 1.4 Tahapan data mart [3]

1.2.3 Busnies Requiremen Defenition

Analyzing business processes and all of the needs that exist in GP SHOES in making the data mart.

1.2.3.1 Analysis of the data source

Analysis of the data source is the process of analyzing existing data sources in GP SHOES production. The data source is made up of several documents can be seen Table below:

Tabel 1.1 Sumber Data

No	Data	Definisi
1	*Customer*	*This data contains customer data owned GP SHOES*
2	*Order*	*This data shows the shoes of the customer order data*
3	*Production*	*This data contains data products (shoes) are produced from raw materials into finished goods*
4	*Shoes*	*This data contains data shoes that have been produced by GP SHOES*
5	*Size*	*This data contains the data size of the shoes manufactured by GP SHOES*
6	*Colour*	*This data contains the color data of shoes on GP SHOES*
7	*The use of raw materials*	*This data contains data usage of raw materials that have been used by GP SHOES*

1.2.3.2 Analisis OLTP GP SHOES

In this study, the data source used is by using OLTP contained in Gp SHOES. The following diagram OLTP GP SHOES relations can be seen in Image 1.5

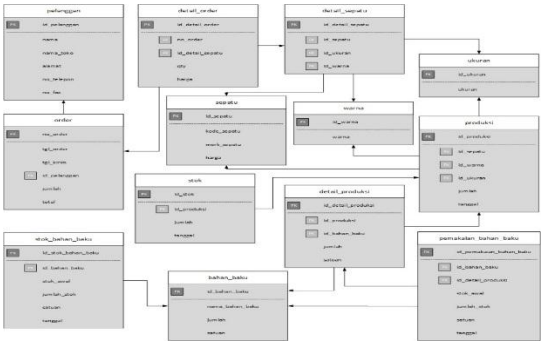


Image 1.5 OLTP GP SHOES

1.2.3.3 Information Needs Analysis

Analysis of information needs is the stage to analyze what is needed by GP SHOES for data mart to be built. The information will be presented in detail. Based on interviews with production managers GP SHOES, information is needed, among others:

1. *Information production quantities of shoes every year, every month and every date.*
2. *Information shoe production number based on the size of shoes every year, every month and every date.*

3. *Information shoe production number based brand shoes every year, every month and every date.*
4. *Information shoe production quantities by color every year, every month and every date.*
5. *Information production quantities of shoes by brand and size every year, every month and every date.*

1.2.*4 Model Dimensionaling*

Modeling data into multidimensional data based on the results obtained from the Business Requirement Definition.

1.2.4.1 Analysts Dimension and Business Facts

1	***Strategic Information Needs***	*nformation production quantities of shoes every year, every month and every date.*
	table Facts	*Fact_produksi*
	Tabel Dimensi	*1. Dim_sepatu* *2. Dim_waktu*
2	***Strategic Information Needs***	*Information shoe production number based on the size of shoes every year, every month and every date*
	table Facts	*Fact_produksi*
	table Dimensions	*1. Dim_ukuran* *2. Dim_waktu*
3	***Strategic Information Needs***	*Information shoe production number based brand shoes every year, every month and every date.*
	table Facts	*Fact_produksi*
	table Dimensions	*1. Dim_sepatu* *2. Dim_waktu*
4	***Strategic Information Needs***	*Information shoe production quantities by color every year, every month and every date.*
	table Facts	*Fact_produksi*
	table Dimensions	*1. Dim_warna* *2. Dim_waktu*
5	***Strategic Information Needs***	*Information production quantities of shoes by brand and size every year, every month and every date*

Based on the above explanation can be concluded that in the construction of a data mart using some fact and dimension tables, it can be seen the model scheme used in the construction of a data mart can be seen in Image 1.6.

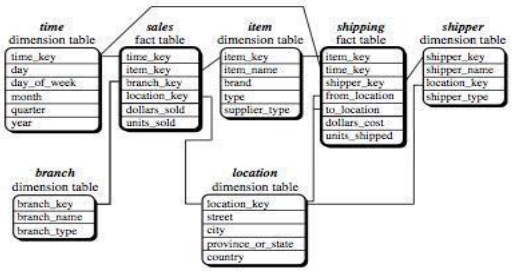


Image 1.6 Skema Constellation

1.2.5 Physican Design

This stage is the stage of the physical design of the data mart. Such as hardware and software you need, how much memory is required, the establishment of the partition if required, and others.

a. The software required to run the data mart as follows :

1. *DBMS SQL Server 2012 as database.*

b. While the hardware neede to run the data mart are as follows :

1. *Processor : Intel Core 2 Duo, 2.0 GHz*
2. *Memory : RAM 1 GB*
3. *Harddisk : 256 GB*
4. *VGA : 128 MB*

1.2.6 Data Stagging Design

Designing a data staging consists of three main stages or commonly called ETL (Extract, Transform, and Load) which is the process of converting data from OLTP database into the data mart.

a. *Extract*

This process is the selection of data from the data source for the manufacture of data mart, which is the product table, the production table, the table of raw materials, table stock out, and table stock production as well as tables that are not used for the manufacture of data marts, namely tables detail production and table stock entry because it is not needed in the information needsi. The attributes that exist in the table will be extracted no change increase or decrease the attributes in the table will be extracted no change increase or decrease its attributes, it still remains the same as the source data. The process of extracting data from the data source into the data mart is as follows:

Tabel 1.2 Extract

No	Nama Tabel	Field
1	Tabel Pelanggan	id_pelanggan
		Nama
		nama_toko
2	Tabel Order	no_order
		tgl_order
		tgl_kirim
		id_pelanggan
		Jumlah
		Total
3	Tabel produksi	id_produksi
		id_sepatu
		id_warna
		id_ukuran
		Jumlah
		Tanggal
4	Tabel ukuran	id_ukuran
		Ukuran
5	Tabel warna	id_warna
		Warna

b. *Transform*

The process of transformation is done consists of two processes, that is :

1. *Cleaning*

Cleaning process cleans unnecessary data from tables in the extract, which removes unused fields. Here is a field name is omitted in the process of cleaning.

a) Cleaning tabel order

In order table does not require field tgl_kirim and a total that will send the order table will be used as the fact table. In order table does not require a date field and a total that will send the order table will be used as the fact table. Cleaning process in order table field removed because no_order field, id_kirim and Total are not used to the process of analyzing the data order. For more details in the process of cleaning the table order can be seen in Tabel 1.3 Cleaning Tabel Order

Tabel Order	
No	**Field**
1	no_order
2	tgl_order
3	tgl_kirim
4	Id_pelanggan
5	Jumlah
6	Total

→

Tabel Order	
No	**Field**
1	Tgl_order
2	Id_pelanggan
3	Jumlah

2. Conditioning

Conditioning process at this table is to change field tgl_order into tabel dimensi waktu with primary key id_waktu. For more details on the conditioning process can be seen in the production Tabel 1.4.

Tabel 1.4 Conditioning Tabel Order

Tabel Order	
No	**Field**
1	tgl_order
2	id_pelanggan
3	Jumlah

→

Fact_Order	
No	**Field**
1	id_waktu
2	id_pelanggan
3	Jumlah

Tabel Order	
tgl_order	date

→

Dim_waktu	
id_waktu	Integer
Tanggal	Integer
Bulan	Integer
nama_bulan	nvarchar(50)
Tahun	Integer
full_date	Date

c. *Loading*

When the data is extracted and transformed, then the data is entered into the data mart. The process of loading the application data marts will be performed automatically after the process is complete transform. The technique used is the update. This process will immediately update the data mart without changing existing data.

1.2.7 OLAP dan Reporting Tools

Manage the data in the data marts into a multidimensional data based on the model that will be shown to the user for decision making.

1. *Analisis Menggunakan OLAP*

In this layer, the data collection from the data mart to make the output in the form of reports and used for data analysis with OLAP. OLAP analysis process used is a roll-up and drill-down and slicing and dicing for both processes helps in filtering based on the dimensions.

a. *Roll-Up*

Roll-Up is a process where we want to see the data globally. For example, displaying the number of products produced per month. Roll-up can display information on the number of products produced by the period of a month into a number of products produced per year.

Tahun	Bulan	DUSTIN SHOES	JK COLLECTION	SCORPION SHOES	UD.CIPTA KARY...
2013	Agustus	989	0	0	0
	Desember	3614	0	0	0
	November	553	0	0	214
	Oktober	778	231	0	841
	September	1491	0	0	0
2013 Total		7425	231	0	1055
2014	Agustus	35	0	0	0
	Desember	599	0	0	0
	Februari	1724	205	26	0
	Januari	10	291	157	0
	Juli	1587	0	0	0
	Juni	300	0	0	0
	Maret	2284	0	0	0
	Mei	1699	45	0	0
	November	297	0	0	0
	Oktober	0	0	0	240
	September	2170	0	0	1970

Into

Tahun	Bulan	DUSTIN SHOES	JK COLLECTION	SCORPION SHOES	UD.CIPTA KARY...
2013		7425	231	0	1055
2014		10705	541	183	2210

b. *Drill-Down*

Drill-Down is the inverse of the roll-up, which we want to see the data in more detail. For example, displaying the number of products produced per year. Drill-down to show information on the number of products produced by periods per year to the number of products produced per month.

Tahun	Bulan	DUSTIN SHOES	JK COLLECTION	SCORPION SHOES	UD.CIPTA KARY...
2013		7425	231	0	1055
2014		10705	541	183	2210

Into

Tahun	Bulan	DUSTIN SHOES	JK COLLECTION	SCORPION SHOES	UD.CIPTA KARY...
2013	Agustus	989	0	0	0
	Desember	3614	0	0	0
	November	553	0	0	214
	Oktober	778	231	0	841
	September	1491	0	0	0
2013 Total		7425	231	0	1055
2014	Agustus	35	0	0	0
	Desember	599	0	0	0
	Februari	1724	205	26	0
	Januari	10	291	157	0
	Juli	1587	0	0	0
	Juni	300	0	0	0
	Maret	2284	0	0	0
	Mei	1699	45	0	0
	November	297	0	0	0
	Oktober	0	0	0	240
	September	2170	0	0	1970

c. *Slicing and Dicing*

Slicing and dicing is the process of taking pieces of the cube by specific values in one or more dimensions. For example, to see the number of products produced by year and month.

Slicing :

Tahun	Bulan	DUSTIN SHOES	JK COLLECTION	SCORPION SHOES	UD.CIPTA KARY...
2013	Agustus	989	0	0	0
	Desember	3614	0	0	0
	November	553	0	0	214
	Oktober	778	231	0	841
	September	1491	0	0	0
2013 Total		7425	231	0	1055
2014	Agustus	35	0	0	0
	Desember	599	0	0	0
	Februari	1724	205	26	0
	Januari	10	291	157	0
	Juli	1587	0	0	0
	Juni	300	0	0	0
	Maret	2284	0	0	0
	Mei	1699	45	0	0
	November	297	0	0	0
	Oktober	0	0	0	240
	September	2170	0	0	1970

Into

Tahun	Bulan	DUSTIN SHOES	JK COLLECTION
2013	Agustus	989	0
	Desember	3614	0
	November	553	0
	Oktober	778	231
	September	1491	0
2013 Total		7425	231
2014	Agustus	35	0
	Desember	599	0
	Februari	1724	205
	Januari	10	291
	Juli	1587	0
	Juni	300	0
	Maret	2284	0
	Mei	1699	45
	November	297	0
	September	2170	0
2014 Total		10705	541

Dicing :

Dicing is the opposite of slicing.

1.2.8 Deployment

Slicing and dicing is the process of taking a cut Operation data marts and reporting tools that is so.

2. Design Data Mart Application

The design of the data mart can be seen in the following image:

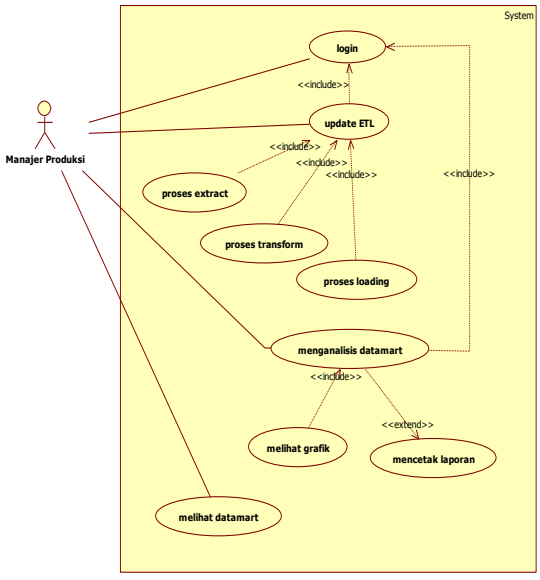


Image 1.7 Use Case Diagram

2.1 Class Diagram designed in the construction of a data mart in GP SHOES can be seen in the following figure:

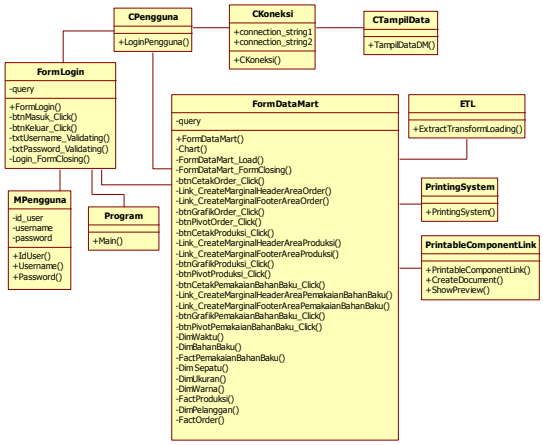


Image 1.8 Class Diagram

2.1 Menu Structure Design Data Mart

Here is the design of the software menu structure of the data mart to be built can be seen In Image1.8.

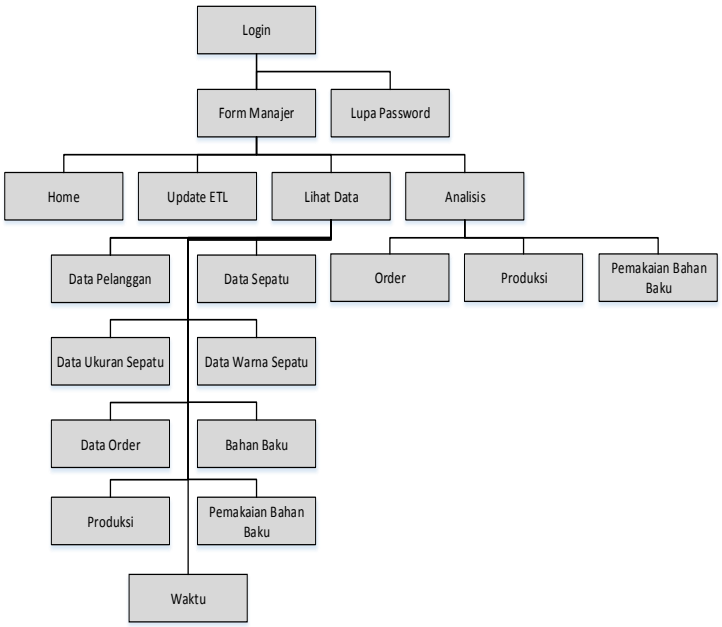


Image 1.8 Struktur Menu Data Mart

3. CONCLUSION

Based on the results of testing on the application analsisi and data marts are built, it can be concluded that:

1. *Application data marts on GP SHOES can provide strategic information that is concise and precise and can speed up the process of analyzing for decision making by the production manager.*

2. *Application of data marts with a GP SHOES can facilitate at the production division in preparing reports on the strategic information needed by the production manager.*

4. BIBLIOGRAPHY

[1] Andrew Cristian Tooy, Bandung, 2013.
[2] M. Y. Pusadan, Rancang Bangun Data Warehouse, Yogyakarta: Graha Ilmu, 2013.
[3] R. Kimball dan M. Ross, The Data Warehouse Toolkit, Indianapolis: John Wiley & Sons, Inc, 2013.
[4] W. Inmon, Building Data Warehouse, Indiana: Wiley Publishing, 2005.